International Journal of Interactive Multimedia and Artificial Intelligence

September 2018, Vol V, Number 2, ISSN: 1989-1660

Education breeds confidence.
Confidence breeds hope.
Hope breeds peace.
Confucius

Special Issue on Big Data and Open Education

http://ijimai.unir.net

INTERNATIONAL JOURNAL OF INTERACTIVE MULTIMEDIA AND ARTIFICIAL INTELLIGENCE
ISSN: 1989-1660 –VOL. 5, NUMBER 2

Editor's Note

One of the most well-known requirements in educational settings is the need to know what happens in a course, lesson plan or full academic programme. That is true for any type of education but in particular for Open Education with the multiple dimensions of openness (Stracke, 2018). On the one hand, educators (i.e. teachers, professors, tutors, etc.) and practitioners in Open Education need to reshape the course plan according to the actual features of the learners (e.g., learning styles, motivation, performance, et cetera) and therefore they require real-time analytical information to supervise, assess, adapt and offer feedback to the learners. On the other hand, Open Education offers specific opportunities through online learning using Open Educational Resources (OER) and providing Massive Open Online Courses (MOOCs) (Corbi & Burgos, 2015). The online environments and platforms provide huge amount of data on all activities (a huge Excel sheet, labelled as Big Data). More importantly, Open Education with open teaching and learning is now commonly shaped by a learner-centred approach that pushes the learners to be the driver of their own learning. That is, learners require awareness to self-assess their progress along the course and make decisions for their next steps. In short, Open Education is now an always changing process that requires effective support for the decision making process by the educators and the learners.

In addition, Open Education at any time requires a big deal of flexibility, as an overall strategy for achieving learning quality and success: flexible learning means to get knowledge, retrieve information and provide feedback at any time, from anywhere; flexible teaching requires to assess from multiple locations and devices and to provide coaching in a large variety of formats; flexible academic management, to combine personal objectives with group goals along with institutional vision and accreditation requirements; flexible content authoring, to integrate open educational materials from the best sources, no matter if proprietary or not; flexible policies, to match competences with official credit recognition from those open sources.

To this extent, the emergence of big data in the open educational arena pushes the research of data mining and analytic techniques in order to describe and understand facts and processes in the course context. Learning Analytics is a relatively recent research field that provides techniques and methods to extract information from learning scenarios, to process the data for automatic discovery of information that goes beyond the raw data, and to report back the findings to the participants of the learning process (i.e. learners, educators) (Picciano, 2014). As a general approach, the human judgement plays a central role in the sense-making process of Learning Analytics systems, while the automatized discovery is a tool to accomplish this goal.

According to the International Conference on Learning Analytics and Knowledge, Learning Analytics refers to "the measurement, collection, analysis and reporting of data about learners and their contexts, for [the] purposes of understanding and optimizing learning and the environments in which it occurs" (Siemens, 2012). It is one of the fastest growing field in technology enhanced learning research (Baker & Yacef, 2009; Ferguson, 2012; Romero & Ventura, 2007). McGrath (2010) and Romero et al. (2008) provide sample use cases of how challenges of tracking learner activities are met with data mining techniques in the context of Learning Management Systems (LMSs). Jivet et al. (2017 and 2018) present a literature review of Learning Analytics dashboards and Scheffel et al. (2017) propose an evaluation framework for Learning Analytics dashboards.

Research is very active in the Learning Analytics field, but there is yet a need for a real impact on daily activities of the end-users. Learners, educators and academic managers require a more flexible and open paradigm to integrate the best of materials, ideas, resources, strategies and any other component of the learning path. This special issue on big data and open education combines a number of excellent outcomes in the field, which can provide a real impact in both, current and future trends and research directions in the research area and the related application contexts. The special issue focuses on the full process of Learning Analytics in Open Education: from data collection methods, through innovative use of analysis techniques in the educational world, and up to information representation methods, with a clear stress on improving a flexible and personalised approach to open and distance settings, for every single shareholder in the learning process.

Analytics techniques are used for the early identification of learners at risk, for score prediction, and as a straightforward way to help learners self-assess their performance on a course. This problem is addressed by a well-known area called 'academic analytics' (Campbell et al., 2007), of which there are several examples in the literature. Macfadyen & Dawson (2010) present an 'early warning system' for educators that mines data from the LMS; Munoz-Organero et al. (2010) establish a relationship among usage patterns of LMS and learner motivation; Romero-Zaldivar et al. (2012) analyse the correlation between learner involvement on a course and the score they obtain; and Swan (2016) closes the gap between analytics, learning and visualisation.

Visual analytics techniques also use graphical representations for synthesising information and deriving insights from large amounts of dynamic, ambiguous and often conflicting data; they are used to detect the expected and discover the unexpected (Keim et al., 2008). They also increase learner retention (De Freitas, 2015) and boost the open paradigm to education (El-Assady et al., 2018). The idea behind visual analytics is to let a computer program filter and pre-process the data, arrange it visually and then let the user perform an interpretation. Given that the field of visualisation is so relevant for supporting educators and learners, Santos et al. (2013) and Bodily et al. (2018) discuss the concept of 'learning dashboards' and identify its empirical evaluation as a research challenge. De Laet et al. (2018) present a framework for the creation of learning dashboards, considering different roles in the learning process and allowing for their integration with third-party systems.

Beyond the logical technical risks and challenges of working with real-time information, process speed and information accuracy, however, practical daily implementation of these visualisation techniques remains a challenge. The practical adoption by non-technical educators of such dashboards requires closer integration with the educational methodology, the learning scenario and the lecturer's profile. Although there is great potential to improve the relation between learner, educator, scenario, institution and other potential practitioners and stakeholders, the field of visual learning analytics requires a clear connection with the end user (educators as well as learners) from basic to expert level, so that they become widely adopted and increasingly useful to the educational community.

The expected impact of Learning Analytics strategies ranges from the promotion of the learners' self-reflection, to the reshape of institutional educational strategies, through the support to educators. It copes with many various areas such as OER and MOOCs, formal and informal learning, face-to-face contexts, blended learning and distance scenarios. In addition, the impact on the actual learning situation highly depends on when the discovered information is presented to

DOI: 10.9781/ijimai.2018.07.005

the practitioners. If the information is analyzed and provided after the course, it allows for a refinement of the educational strategy in future courses (summative approach). On the contrary, when the findings are provided during the course, then the course participants can lively adapt resources and efforts in order to get their expectations accomplished (formative approach).

Research is very active in the fields of Learning Analytics and Open Education, but there is yet a need for a real impact on daily activities of the end users and on improving the learning quality (Stracke, 2018). Learners, educators and academic managers require a more flexible and open paradigm to integrate the best of materials, ideas, resources, strategies and any other component of the learning path. This special issue on Big Data and Open Education combines a number of excellent outcomes in the field, which can provide a real impact in both, current and future trends and research directions in the research area and the related application contexts. The special issue focuses on the full process, from data collection methods, through innovative use of analysis techniques in the educational world, and up to information representation methods, with a clear stress on improving a flexible and personalised approach to open and distance settings, for every single shareholder in the learning process.

In summary, Open Education can benefit from Big Data and Learning Analytics used and analysed in the right way as well as Open Education can be the enabler for a broader implementation and acceptance of Big Data and Learning Analytics, again if realized in the correct way. Recent research shows the upcoming relevance of Big Data in Open Education through decision support systems aiming at learners and educators, usually representing the information with information visualization techniques and dashboards. There are a lot of on-going research projects related Learning Analytics and Open Education that are producing interesting results to improve learning quality and to achieve societal impact. In this special issue, we present just a selection of fine papers focused on analytics (Moreno et al.), an intelligent assistant (Lodhi et al.), higher education (Hamoud et al.), vocational education (de Lange et al.), employment (Peñalvo et al.), gamification (González-González et al.), augmented reality (Medina et al.) and Open Educational Resources (Idrissi et al). We hope that the reader enjoys as much as we did, as editors of this double-blind, peer-reviewed compilation.

Prof. Dr. Daniel Burgos

Prof. Dr. Roumen Nikolov

Assoc. Prof. Dr. Christian M. Stracke

References

[1] Baker, R. S. J. D., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining, 1(1), 3–17.

[2] Corbi, A., & Burgos, D. (2015). Semi-automated correction tools for mathematics-based exercises in MOOC environments. International Journal of Interactive Multimedia and Artificial Intelligence, 3(3), 89-95.

[3] Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic Analytics. Educause Review, 42(October), 40–57.

[4] Donovan, J. O., Smyth, B., Gretarsson, B., Bostandjiev, S., & H, T. (2008). PeerChooser : Visual Interactive Recommendation. Computer Science and Informatics, 1085–1088.

[5] Ferguson, R. (2012). The State of Learning Analytics in 2012 : A Review and Future Challenges a review and future challenges.

[6] Jivet, I., Scheffel, M., Specht, M., & Drachsler, H. (2018). License to evaluate: Preparing learning analytics dashboards for educational practice. In Proceedings of International Conference on Learning Analytics and Knowledge (LAK '18). (in print, 10 pages). DOI: 10.1145/3170358.3170421

[7] Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017, September). Awareness is not enough: pitfalls of learning analytics dashboards in the educational practice. In European Conference on Technology Enhanced Learning (pp. 82-96). Springer, Cham.

[8] Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., & Melançon, G. (2008). Visual Analytics: Definition, Process, and Challenges. In A. Kerren, J. T. Stasko, J.-D. Fekete, & C. North (Eds.), Information Visualization (pp. 154–175). Berlin, Heidelberg: Springer Berlin Heidelberg.

[9] Larusson, J. A., & Alterman, R. (2009). Visualizing student activity in a wiki-mediated co-blogging exercise. Proceedings of the 27th international conference extended abstracts on Human factors in computing systems CHI EA 09, 4093.

[10] Leony, D., Pardo, A., De la Fuente Valentín, L., De Castro, D. S., & Kloos, C. D. (2012). GLASS. Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12 (p. 162). New York, New York, USA: ACM Press.

[11] Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. Computers & Education, 54(2), 588–599.

[12] Mazza, R., & Dimitrova, V. (2004). Visualising student tracking data to support instructors in web-based distance education. Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters (pp. 154–161). ACM.

[13] Munoz-Organero, M., Munoz-Merino, P. J., & Kloos, C. D. (2010). Student Behavior and Interaction Patterns With an LMS as Motivation Predictors in E-Learning Settings. IEEE Transactions on Education, 53(3), 463–470.

[14] Picciano, A. G. (2014). Big data and learning analytics in blended learning environments: benefits and concerns. International Journal of Interactive Multimedia and Artificial Intelligence, 2(7), 35-43.

[15] Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. Expert Systems with Applications, 33(1), 135–146.

[16] Romero, Cristóbal, & Ventura, S. (2011). Preface to the special issue on data mining for personalised educational systems. User Modeling and User-Adapted Interaction, 21(1-2), 1–3.

[17] Romero, Cristóbal, Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. Computers & Education, 51(1), 368–384.

[18] Romero Zaldívar, V. A., Pardo, A., Burgos, D., & Delgado Kloos, C. (2012). Monitoring Student Progress Using Virtual Appliances : A Case Study. Computers & Education, 58(4), 1058–1067.

[19] Santos, J. L., Verbert, K., Govaerts, S., & Duval, E. (2013). Addressing learner issues with StepUp! Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13 (p. 14). New York, New York, USA: ACM Press.

[20] Scheffel, M., Drachsler, H., Toisoul, C., Ternier, S., & Specht, M. (2017). The Proof of the Pudding: Examining Validity and Reliability of the Evaluation Framework for Learning Analytics. In E. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), Data Driven Approaches in Digital Education: Proceedings of the 12th European Conference on Technology Enhanced Learning, LNCS 10474 (pp. 194–208). Berlin, Heidelberg: Springer.

[21] Siemens, G. (2012). Learning analytics: envisioning a research discipline and a domain of practice. Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12 (p. 4). New York, New York, USA: ACM Press.

[22] Stracke, C. M. (2018). Como a Educação Aberta pode melhorar a qualidade de aprendizagem e produzir impacto em alunos, organizações e na sociedade? In T. Amiel (Ed.), Utopias and Dystopias in Education. Sao Paulo: UNICAMP. (in print, 37 pages)

[23] Verpoorten, D., Westera, W., & Specht, M. (2011). A first approach to "Learning Dashboards" in formal learning contexts.

TABLE OF CONTENTS

OPEN ACCESS JOURNAL

ISSN: 1989-1660

Comparison of Clustering Algorithms for Learning Analytics with Educational Datasets

Alvaro Martínez Navarro[1], Pablo Moreno-Ger[2] *

[1] Departamento de Ingeniería de Sistemas, Universidad Mariana, Pasto (Colombia)
[2] Escuela Superior de Ingeniería y Tecnología, Universidad Internacional de La Rioja, Logroño (España)

Received 3 November 2017 | Accepted 22 January 2018 | Published 9 February 2018

unir LA UNIVERSIDAD EN INTERNET

Abstract

Learning Analytics is becoming a key tool for the analysis and improvement of digital education processes, and its potential benefit grows with the size of the student cohorts generating data. In the context of Open Education, the potentially massive student cohorts and the global audience represent a great opportunity for significant analyses and breakthroughs in the field of learning analytics. However, these potentially huge datasets require proper analysis techniques, and different algorithms, tools and approaches may perform better in this specific context. In this work, we compare different clustering algorithms using an educational dataset. We start by identifying the most relevant algorithms in Learning Analytics and benchmark them to determine, according to internal validation and stability measurements, which algorithms perform better. We analyzed seven algorithms, and determined that K-means and PAM were the best performers among partition algorithms, and DIANA was the best performer among hierarchical algorithms.

Keywords

Clustering Methods, Computer Languages, Data Analysis, Engineering Students, Performance Evaluation, Unsupervised Learning.

DOI: 10.9781/ijimai.2018.02.003

I. Introduction

Since the turn of the Century, researchers have been studying clustering methods and comparing them from different perspectives. These algorithms were the core of an emerging data mining discipline, which would soon explode along with the popularity of big data approaches in all fields. And when these ideas were applied to digital education, the field of Learning Analytics was born.

But the core of these approaches is still the use of adequate clustering algorithms for each scenario, and this problem has received a fair share of attention. Berkhin contrasted theoretically different algorithms [1], and indicated how to perform the most typical evaluations, data preparation and measurements. In [2], the authors studied 216 articles written between 2000 and 2011, classifying the literature in three axes (knowledge types, analysis types, and architecture types) and exploring the different context where such techniques may be used. Remarkably, the study highlighted the potential applications of data mining techniques in social sciences, psychology, cognitive sciences and human behavior, which is very relevant for this specific work. In turn, other studies such as [3] provided solid grounds for defining and scoping clustering, discussing aspects such as variable selection and similarity measurements, and provided a theoretical foundation of grouping methods and applications. Other authors have studied the theoretical limitations and potential pitfalls of these techniques [4].

All in all, the theoretical foundation of clustering algorithms is solid and has been the object of detailed studies. In terms of experimental analysis, we can find publications as early as in the 80s. In [5] the author proposed specific analysis methods, and also studied the performance impact of different perturbations. In [6] the authors compared five grouping algorithms and used four different supervised automatic learning algorithms to analyze their performance. In [7] three algorithms were measured with four cluster validation indexes, using synthetic and real datasets. Other relevant works have conducted formal tests to determine the most appropriate data mining algorithms for specific fields such as classification [8] or text mining from RSS sources [9]. However, few experimental studies focused on analyzing performance using specifically educational datasets.

In turn, Learning Analytics (LA) research ranges from theoretical essays on the potential impact of LA in education [10] to very focused studies on how it is useful for establishing personalized feedback to improve academic performance [11]. There are also works proposing dynamic models for data analysis of educational datasets [12] or proposing the use of different statistical algorithms to rank academic performance [13]. Regarding the evolution of the state of the art, different works have studied stakeholders, benefits and challenges [14]-[15] or differentiated types of educational settings, tasks and outputs [16]. Other authors have studied specific quality indicators to assess the impact of LA in education [17] or studied in depth the foundations of LA in terms of data mining techniques [18]. All these efforts can be characterized for their use of clustering algorithms as an analysis technique, which is the focus of this work, but typically focus on one or two algorithms at most, or just in the educational and social implications, rather than focusing specifically in experimental comparison of the performance of different algorithms when applied in educational settings.

And while the challenges of small-scale and relatively clean educational datasets lie mostly on how to identify the best visualizations or practical uses of the data, the emergence of open education formats is yielding increasingly complex and noisy datasets, imposing non-

* Corresponding author.

E-mail addresses: amartinez@umariana.edu.co (Alvaro Martínez Navarro), pablo.moreno@unir.net (Pablo Moreno-Ger).

trivial burdens on the data mining algorithms applied to make sense of how students are interacting with these open materials.

At the crossroads of these lines of research, our goal is to contribute experimental validation of the performance of different clustering techniques when specifically applied to educational datasets, thus providing a more solid foundation for further works focusing on practical aspects rather than back-office performance.

To achieve this goal, we have conducted a practical experiment using a real-world dataset provided by Universidad Mariana in Colombia, benchmarking different algorithms and configurations in terms of internal validations and stability measurements.

II. Materials and Methods

The experimental design is quite straightforward. We started with a literature review to select a representative set of clustering algorithms. Then, we organized a workflow for testing each algorithm and selected specific measurements for comparison, and finally we applied this workflow to all algorithms targeting an educational dataset provided by Universidad Mariana. This section details each of these steps.

A. Selecting the Algorithms to Be Benchmarked

The specific selection of algorithms was conducted after performing a literature review, with a heavy influence of related works from other fields ([1], [4], [16]) and trying to provide a wide perspective of the potential approaches.

The final selection of algorithms is summarized in Table I.

B. Experimental Platform

We employed different platforms and tools to create our experimental pipeline. We started with raw and cross-referenced data available on an Oracle 10g database server. We extracted different listings and used Microsoft Excel to review, perform basic cleaning (including anonymization) and saving as CSV (comma-separated values) files.

All statistical analyses and clustering algorithms were applied using the opensource platform R, for which we created our test scripts using the R Studio graphical interface. The platform was also used to create the different visualizations that helped in this study and that are included in this article. In Table II we provide a summary of the different libraries that we used for the experiment.

TABLE I. Algorithm Selection

NAME	Description
K-means	It builds the clusters with similarity measure, each cluster is represented by a centroid, which is the average of its elements. It works only with numerical data and its affected by outliers [1].
CLARA	It builds the clusters without all the data, instead, sample them at random and then the centers are calculated with the PAM algorithm. It works with numerical data [1].
PAM	It is a robust version of K-means. The process of grouping is done through the search of a sequence of objects called medoids. Its main characteristic is the reduction of the average dissimilarity of the observations to its nearest observation [1].
FANNY	It calculates the clustering in a diffuse way in a number K clusters.
Hierarchical	It uses previous clusters to form the new clusters. It's of two types: divisive (top-down) and agglomerative (bottom-up). It can be applied with any type of variables and it is not necessary to establish a priori the number of clusters to be formed. It is sensitive to noise and outliers.
AGNES	Hierarchical algorithm of agglomerative type, it uses the Euclidean distance as similarity measure.
DIANA	Hierarchical algorithm of divisive type, it uses the Euclidean distance as similarity measure.

C. Benchmarking Performance

In order to benchmark the performance of each algorithm, we focused on the facilities provided by the *clValid* library presented in [19] to measure internal and stability validations.

Internal validations are computed using intrinsic information from the datasets to assess the quality of the resulting clusters. We used the three main internal validation measurements offered by the validation library: *connectivity* (which provides a value in the $[0, \infty)$ range where lower is better), *silhouette width* (which proves values in the $(-1,1)$ range where higher is better), and *Dunn index* (which provides a value in the $[0, \infty)$ range where higher is better).

In terms of stability measurement, we again selected the main measures from [19], all of them focused on inspecting each cluster and sequentially removing internal columns and checking whether the cluster remains valid. We employed APN (average proportion of non-overlap), AD (average distance between measurements), ADM (average distance between means) and FOM (figure of merit, focused on the average intra-cluster variance of the observations). All of them take values in the $[0, \infty)$ range where higher is better except APN, with values in the $[0,1]$ range with preferred results close to zero.

TABLE II. R Libraries

Library	Description
cluster	It allows clustering analysis by implementing hierarchical and partition algorithms. Details in: https://cran.r-project.org/web/packages/cluster/index.html
ggplot2	It builds visualizations using the information of the data meaning. Details in: https://cran.r-project.org/web/packages/ggplot2/index.html
factoextra	It offers different easy-to-use functions to extract and visualize the results of multivariate analyzes, it simplifies clustering analysis and its graphical representations. Details in: https://cran.r-project.org/web/packages/factoextra/index.html
readr	It provides fast and friendly mechanisms to read files in csv, tsv and fwf formats. Details in: https://cran.r-project.org/web/packages/readr/index.html
RColorBrewer	It provides color schemes to be used with various types of graphic representations. Details in: https://cran.r-project.org/web/packages/RColorBrewer/index.html
gplots	It offers several tools to draw processed data. Details in: https://cran.r-project.org/web/packages/gplots/index.html
d3heatmap	It creates interactive heat maps that can be viewed in different environments such as R console and R studio. Details in: https://cran.r-project.org/web/packages/d3heatmap/index.html
stats	It offers different statistical functions for data analysis. Details in: https://stat.ethz.ch/R-manual/R-devel/library/stats/html/00Index.html
NbClust	It provides 30 indexes to determine the optimal number of clusters in a dataset and it offers the best clustering scheme of different results. Details in: https://cran.r-project.org/web/packages/NbClust/
clValid	It provides functions to evaluate clustering results in biological and statistical way. Details in: https://cran.r-project.org/web/packages/clValid/index.html
pvclust	It is an implementation of bootstrap resampling to evaluate the uncertainty in hierarchical clustering. Details in: https://cran.r-project.org/web/packages/pvclust/index.html

D. Preparing the Dataset

One particularity of Learning Analytics approaches is the focus on inspecting academic data generated at specific institutions for specific courses, since comparisons across platforms, institutions and even individual courses may be challenging due to significant differences in usage patterns [10]

We therefore chose to focus on a specific student profile within the host institution: Computer Engineering students enrolled in the first semester with the 2010-2016 timeframe (7 years).

The datasets were constructed with the main measurement of the average grade of all courses in the semester. The data was extracted from the host University's database system, using specific queries and basic data management (e.g. transposing) to create rows representing each specific student's aggregated data.

The scripts were also prepared to provide anonymization when required and removed non-essential personal data (e.g. personal addresses or phone numbers).

All the relevant R scripts as well as the dataset can be downloaded from https://goo.gl/oNHm2R.

III. Results

A. Pre-processing

The first step in our test pipeline was the conduction of relatively simple cleanup tasks, mostly focused on removing instances where some grades were missing (this may happen either due to an administrative issue or an error in the grade reporting system).

After this step, we analyzed for each variable different comparison statistics, averaging their maximum, minimum, average and mean values. We determined that their standard deviation and variances were minimal, therefore making it possible to work directly with the original data without requiring a previous normalization.

We also analyzed the dataset to validate its clusterability, using two approaches: first, we used the *Hopkins statistic* method from [20], with a resulting value of 0.2036606 (<0.5) which shows that the values are potentially groupable. We also validated this notion visually, by representing the tendency of the data to be grouped. This is achieved by (1) calculating the dissimilarities between all datapoints and storing them in a dissimilarity matrix according to their Euclidean distances, (2) sorting the matrix so that similar objects are closer together and (3) displaying the matrix to check the presence of high values along the diagonal of the matrix (Fig. 1).

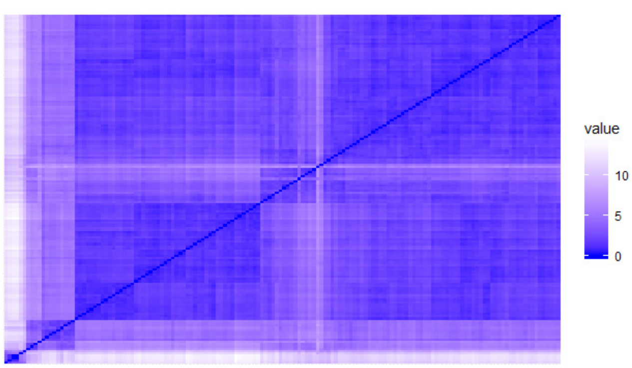


Fig. 1. Data clustering trend where we can observe dark blue rectangles aligned with the diagonal line, which can be interpreted as a potential clustering amenability of the data.

Therefore, we could ascertain the adequacy of our dataset for the goals of our experiment.

B. Optimal K Value

In order to determine the optimal number of clusters, we again used to separate and complementary methods. The first approach we used employed the *gap stastic* [21] targeting the K-means algorithm with a maximum value of K=10 and 10,000 Monte Carlo samples. The result can be observed in Fig. 2, yielding a proposal for one single cluster (or rather hinting that the data was not prone to clustering), although applying the *1-standard error* criterion K=5 may also be a valid candidate.

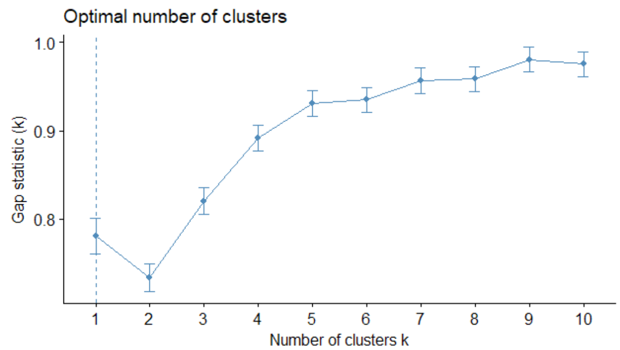


Fig. 2. Optimal K presented by the gap statistic.

Given this partially unsatisfactory result, we also looked at the nbClust R package, which analyses 30 separate indexed to determine the optimal K. We ran this analysis for values of K in the [2,10] range, with full clustering and all indexes included. As observed in Fig. 3, this yielded an optimal number of clusters of K=3.

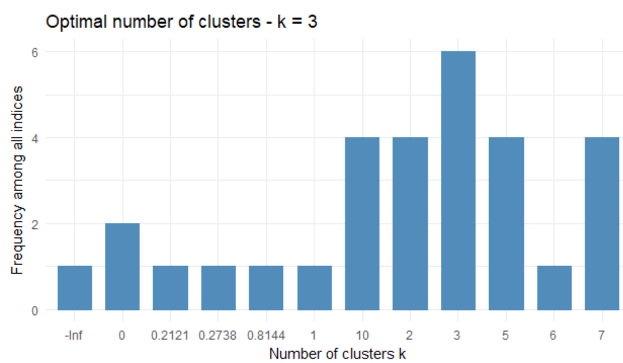


Fig. 3. Optimal K reported by nbClust.

C. Algorithm Execution with Optimal K Value

All seven algorithms were run on the data using K=3 in order to study their behavior and performance.

Regarding the set of partition algorithms, Fig. 4a shows the dispersion and cluster charts produce by the K-means, CLARA and PAM algorithms (all three produced the same output), while Fig. 4b shows the results for the FANNY algorithm, which displayed a significantly worse performance.

These results were validated through a silhouette inspection, which measures the adequacy of each observation for each cluster representing the average distance between groups. Fig. 5 shows the results of these inspections. The K-means, CLARA and PAM algorithms yielded an average silhouette width of 0.55 while the FANNY algorithm yielded 0.29. This indicates a good result for the first three algorithms, while the FANNY algorithm even had a cluster with negative average width, representing a large number of incorrectly assigned instances.

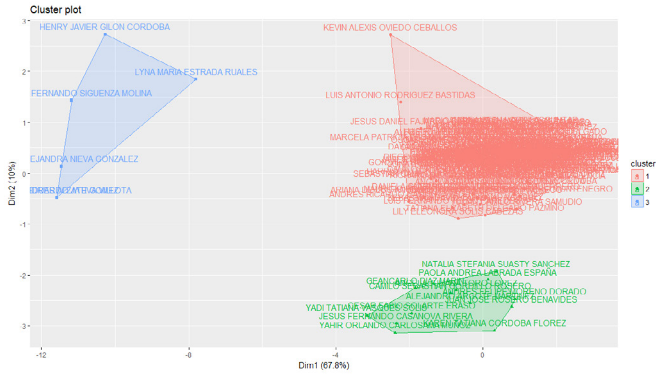


Fig. 4 a. Cluster plot for K-means, CLARA and PAM algorithms. All three yielded the same result.

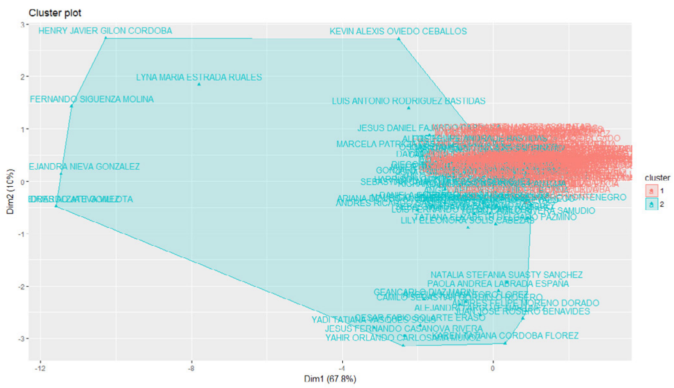


Fig. 4 b. Cluster plot for the FANNY algorithm.

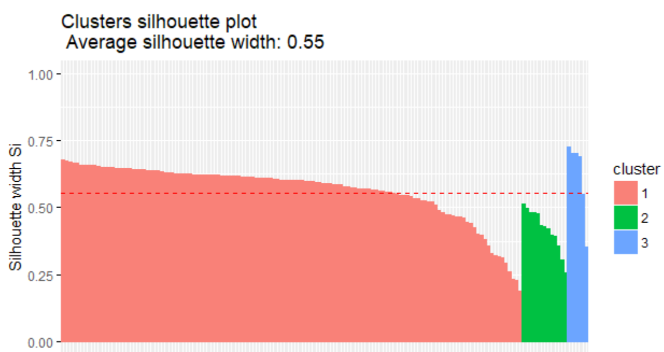


Fig. 5 a. Clusters silhouette plot for the K-means, CLARA and PAM algorithms. All observations are >0 with an average of 0.55.

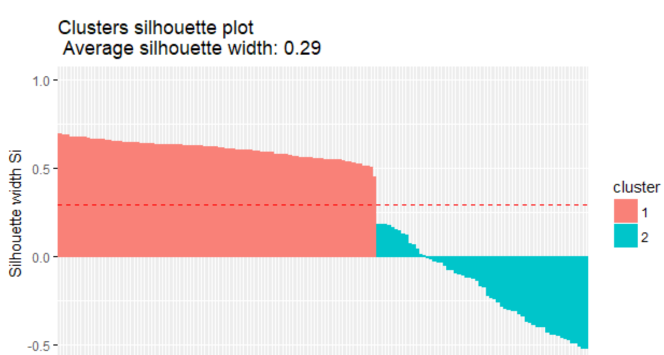


Fig. 5 b. Clusters silhouette plot for the FANNY algorithm. In cluster 2 (cyan color) there are negative values.

We took the same approach for our selection of hierarchical algorithms. Fig. 6 show the cluster plots for the hierarchical and AGNES algorithms (which were equal) and for the DIANA algorithm (which presented issues with one of the clusters).

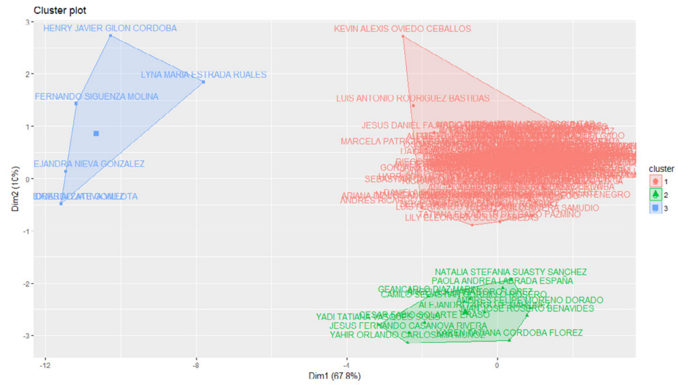


Fig. 6 a. Cluster plot for the hierarchical and AGNES algorithms. Both yielded the same result.

Fig. 6 b. Cluster plot for the DIANA algorithm.

Similarly to the other family, the clusters yielded by the hierarchical variants were validated through silhouette inspection, using a different visualization due to concerns with the very small cluster yielded by the DIANA algorithm.

The silhouettes can be observed in Fig. 7. The first two algorithms yielded the same result while, remarkably, the DIANA algorithm had produced three clusters, all with positive silhouette values, meaning that all values are well assigned, even if the value for cluster #3 (the small blue one in Fig. 6b) is relatively low (0.27) although the overall value is 0.54, very close to the other two algorithms.

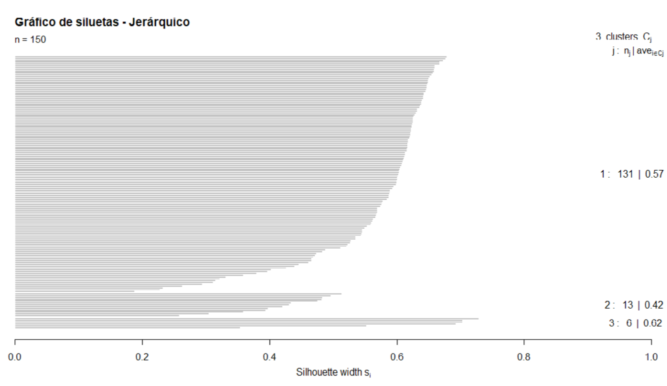


Fig. 7 a. Clusters silhouette plot for the hierarchical and AGNES algorithms.

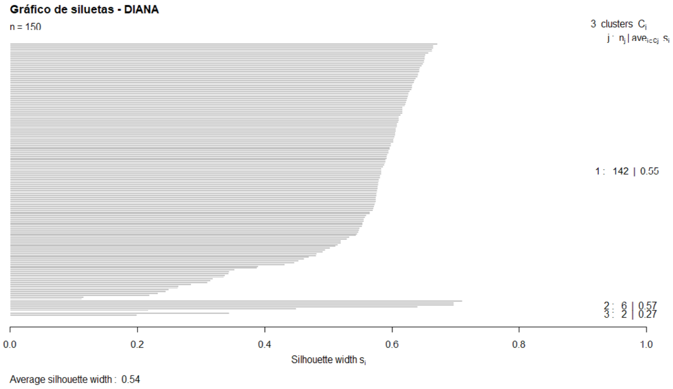


Fig. 7 b. Clusters silhouette plot for the DIANA algorithm.

D. Algorithm Execution Doubling the Optimal K Value

While the previous observations already present some interesting insights, the decision of going with a specific K value was a concern, especially given that one of our tests suggested potential gains for relatively high K values.

We therefore repeated the process using a higher value, K=6 (doubling the previous value) in order to check both the consistency of the previous results and the effect of increasing K in general.

In Fig. 8 we can observe the performance of the clustering algorithms once we double the K value. FANNY basically maintains the same average (and poor) performance for K=3 and K=6, while the other algorithms actually decreased their performance, yielding small but significant errors in most cases. In this sense, PAM and CLARA took a significant performance hit, with 11 and 10 wrongly classified observations respectively. In turn, the classic K-means algorithm presented four negative values after increasing the K value.

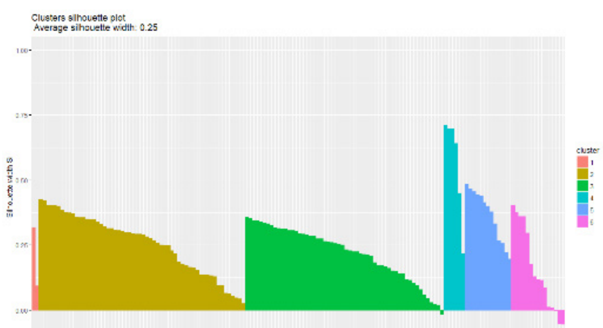


Fig. 8 a. Clusters silhouette plot for K-means with K=6.

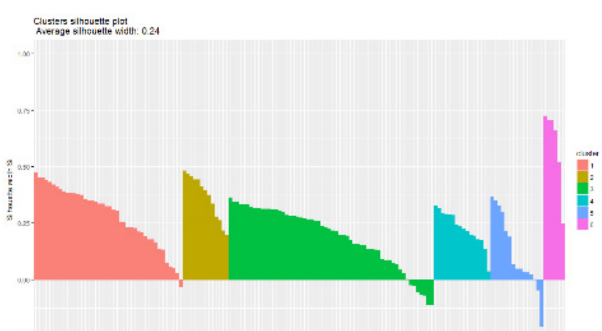


Fig. 8 b. Cluster silhouette plot for CLARA with K=6.

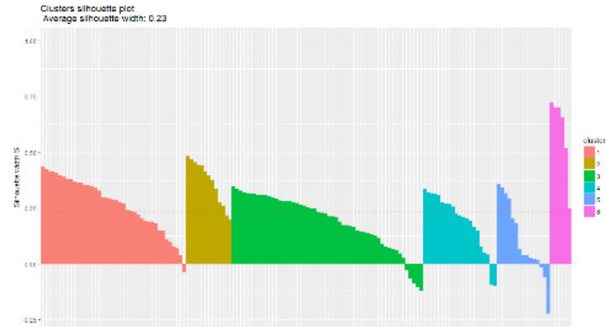


Fig. 8 c. Cluster silhouette plot for PAM with K=6.

Regarding hierarchical algorithms, their cluster silhouette plots (Fig. 9) show that the hierarchical and AGNES algorithms still show the same behavior but doubling the K value impacted their performance negatively. Not only they assigned more instances incorrectly, but they also reduced the average silhouette value from 54% to 24%. In turn, the DIANA algorithm also took a performance hit, but did not experience any incorrect classifications.

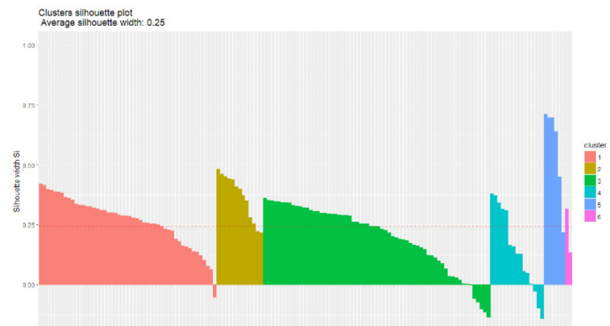


Fig. 9 a. Cluster silhouette plot for hierarchical with K=6.

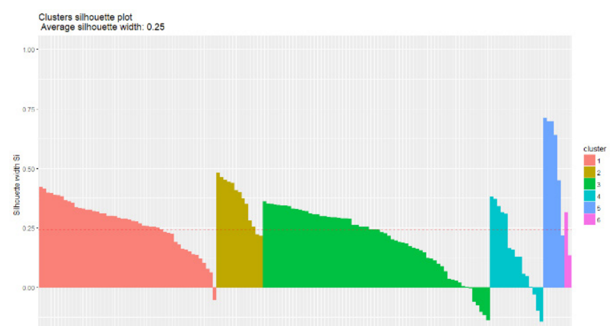


Fig. 9 b. Cluster silhouette plot for AGNES with K=6.

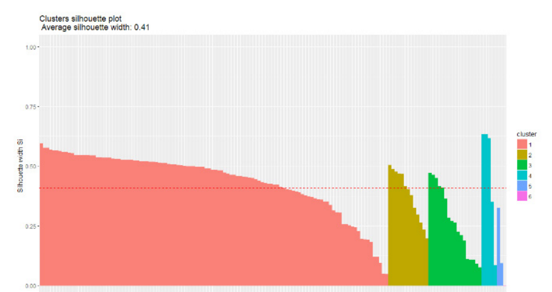


Fig. 9 c. Cluster silhouette plot for DIANA with K=6.

E. Internal Validation and Stability Measurement

As stated in previous section, in order to assess the performance of the different algorithms we performed a comparison using internal validation and stability measurements. Fig. 10 displays the R output with all the internal validation scores for our selection of algorithms, for varying K values from 3 to 6, while Fig. 11 displays the output for stability scores.

```
Clustering Methods:
 kmeans clara pam fanny hierarchical agnes diana

Cluster sizes:
 3 4 5 6

Validation Measures:
                                  3       4       5       6

kmeans       Connectivity    5.7167 11.1667 32.6063 38.1222
             Dunn            0.3721  0.3721  0.1034  0.2010
             Silhouette      0.5549  0.5476  0.4209  0.3849
clara        Connectivity    5.7167 54.1635 70.5667 69.2286
             Dunn            0.3721  0.0485  0.0494  0.0787
             Silhouette      0.5549  0.2151  0.2301  0.2435
pam          Connectivity    5.7167 58.4603 75.7325 76.1421
             Dunn            0.3721  0.0485  0.0485  0.0781
             Silhouette      0.5549  0.2168  0.2262  0.2345
fanny        Connectivity        NA      NA      NA      NA
             Dunn                NA      NA      NA      NA
             Silhouette          NA      NA      NA      NA
hierarchical Connectivity    8.6151 11.9484 13.7079 16.6579
             Dunn            0.3883  0.3883  0.4338  0.4338
             Silhouette      0.5450  0.5377  0.5513  0.5524
agnes        Connectivity    8.6151 11.9484 13.7079 16.6579
             Dunn            0.3883  0.3883  0.4338  0.4338
             Silhouette      0.5450  0.5377  0.5513  0.5524
diana        Connectivity    8.6151 10.3746 28.6325 31.9659
             Dunn            0.3883  0.4338  0.1209  0.1466
             Silhouette      0.5450  0.5586  0.4156  0.4083

Optimal Scores:

             Score  Method       Clusters
Connectivity 5.7167 kmeans       3
Dunn         0.4338 hierarchical 5
Silhouette   0.5586 diana        4
```

Fig. 10. Internal validation scores for all algorithms.

```
Clustering Methods:
 kmeans clara pam fanny hierarchical agnes dia

Cluster sizes:
 3 4 5 6

Validation Measures:
                        3      4      5      6

kmeans       APN   0.0350 0.1462 0.0524 0.0345
             AD    2.4473 2.4391 2.2072 2.0718
             ADM   0.2036 0.5720 0.4537 0.3029
             FOM   0.6958 0.6866 0.6907 0.6690
clara        APN   0.0758 0.1308 0.1661 0.1674
             AD    2.4425 2.1941 2.0620 1.9967
             ADM   0.2482 0.3520 0.4425 0.5323
             FOM   0.6917 0.6630 0.6674 0.6680
pam          APN   0.0476 0.1499 0.1706 0.2969
             AD    2.4014 2.2172 2.0254 1.9934
             ADM   0.1323 0.4477 0.3421 0.5667
             FOM   0.6957 0.6676 0.6585 0.6578
fanny        APN       NA     NA     NA     NA
             AD        NA     NA     NA     NA
             ADM       NA     NA     NA     NA
             FOM       NA     NA     NA     NA
hierarchical APN   0.0187 0.0525 0.0049 0.0071
             AD    2.7104 2.6614 2.2709 2.2316
             ADM   0.1080 0.2569 0.1111 0.1171
             FOM   0.7016 0.6912 0.6787 0.6715
agnes        APN   0.0187 0.0525 0.0049 0.0071
             AD    2.7104 2.6614 2.2709 2.2316
             ADM   0.1080 0.2569 0.1111 0.1171
             FOM   0.7016 0.6912 0.6787 0.6715
diana        APN   0.0350 0.0265 0.0176 0.0161
             AD    2.7106 2.3175 2.2031 2.0415
             ADM   0.1729 0.1330 0.4379 0.1834
             FOM   0.7015 0.6873 0.6765 0.6598

Optimal Scores:

    Score  Method       Clusters
APN 0.0049 hierarchical 5
AD  1.9934 pam          6
ADM 0.1080 hierarchical 3
FOM 0.6578 pam          6
```

Fig. 11. Stability scores for all algorithms.

Regarding partition algorithms, for K=3, K-means, CLARA and PAM tie as the best performers, due to their lower connectivity score (5.7167), best silhouette coefficient (0.5549) and best Dunn index (0.3721). However, this tie disappeared when we increased K values, with K-means maintaining better performance while CLARA and PAM quickly degraded their numbers. In turn, the FANNY algorithm yields no output given its inability to generate measurable clusters. In terms of stability, PAM achieved the best score with K=6 for AD and FOM measurements, and also performed excellently with ADM with K=3. K-means achieved the best APN score with K=6.

Focusing on hierarchical algorithms, hierarchical achieved the best connectivity and Dunn scores for K=3 and K=5 respectively, while DIANA achieved the best score in terms of correct instance assignments when K=4. In terms of stability scores, hierarchical again displays the best scores for K=3 and K=5 for APN and ADM measurements, while DIANA achieves optimal values for AD and FOM when K=6.

IV. Discussion

One of the most relevant (although reasonably expected) observations is that no algorithm is a clear and obvious winner across all measurements and potential K values.

In terms of internal validation, K-means, CLARA and PAM achieved the best overall scores with K=3, although CLARA and PAM experienced a worst degradation as the K value increased. However, the hierarchical and AGENS algorithms also achieve very significant Dunn scores when K=4.

If we focus on K=3 (our selected optimal value), the worst performers in connectivity were the three hierarchical algorithms, although they achieved better Dunn index scores. However, they also presented more incorrect assignments, and therefore can be considered worst performers overall.

However, as we increased the K value, partition algorithms degraded their performance quickly, while hierarchical algorithms remained more stable and actually improved some scores.

In terms of stability, again there is no single algorithm that achieves the best score in all four measurements. PAM exhibited good AP and FOM behavior at K=6, hierarchical achieved very good APN with K=5 and very good ADM with K=3.

Again, if we focus on K=3, the worst performer in terms of APN and ADM was CLARA, while DIANA and AGNES performed poorly in AD and FOM respectively.

The pattern of degradation as we increased the K value exhibited by partition algorithms was also apparent when looking at stability measurements, yielding a consistent conclusion of the better behavior of hierarchical algorithms for higher number of clusters.

The FANNY algorithm failed to produce significant clusters and was therefore deemed as poorly fit for our specific dataset.

V. Conclusion and Future Works

Open education is bound to push the boundaries of how we analyze our educational datasets. And as the scope of our research actions Learning Analytics becomes more and more specialized, the specific underlying techniques, including the selection of a particular clustering algorithm, are bound to receive less attention than appropriate.

This study aims to provide researchers with insights into how the different algorithms exhibit different performance patterns depending on specific measurements and variation in K values, especially when the dataset is highly driven by a set of grades in different courses.

This is achieved through a detailed and highly practical experiment, selecting the most prominent algorithms identified in the literature and analyzing them using an assortment of assessment tools and an educational dataset from a higher education institution.

Among the seven clustering algorithms selected, we measured which algorithms performed better at an experimentally determined K value (K=3) and henceforth how they changed their performance if we increased this number.

During the experimental work we highlighted different relevant observations, from which we can distil some specific insights:

- Among partition algorithms, K-means and PAM were the best performers overall. The former achieved the best results in terms of internal validation (especially as we increased the K value) while the latter performed better in terms of stability.
- Among hierarchical algorithms, DIANA and hierarchical were the best performers, with a similar variation: the former achieved better internal validation scores, while the latter achieved better stability scores.
- Student grades in the sample dataset were highly groupable, as corroborated by the Hopkins statistic, a result that we expect would be extrapolated to other educational datasets, especially in higher education, where students tend to form grade patterns with ease. In lower education levels, the breadth of topics may introduce additional noise as students may have greater affinities for specific courses.
- In terms of determining the optimal K value, the Gap statistic was not really helpful, suggesting one single cluster even though the performance of the clustering algorithms for higher K values was rapidly apparent.
- Increasing the K value improved the performance of many algorithms in most metrics, although the number of errors also increased, and this improvement should be taken in context.

This work, however, also has specific limitations. Firstly, the preparation of the dataset was performed through an aggressive cleanup of the data, discarding all instances where any piece of information was missing. This resulted in a clean dataset 44% smaller than the original one. Given that this was an official dataset provided by the host University, it is to be expected to get similar noise levels in other real (non-synthetic) datasets, and better data cleanup techniques could be required. Making sure that we do not loose significant information while cleaning up data remains a significant open line of research.

In addition, most validations were performed trough different variations of Euclidean distance measurements, ignoring other approaches that may provide additional insights. This invites the potential expansion of this experiment either by including new measurements or by testing alternative educational datasets (or both).

Finally, we expect that the comparisons performed in this work will be helpful for future researchers looking into how to select the best algorithms for performing clustering analysis of educational datasets in higher education.

Further than the specific scores achieved by the different algorithms, we believe that this work adds value by identifying performance patterns that can be used as a base in future research.

Acknowledgment

This work was partially funded by Universidad Internacional de La Rioja (UNIR, http://www.unir.net) through the Research Institute for Innovation & Technology in Education (UNIR iTED, http://ited.unir.net) and the IBM-UNIR Chair on Data Science in Education.

Thanks to Universidad Mariana - Pasto - Colombia for their collaboration in this research, especially to Iván Bastidas for his help in obtaining the data for this experiment.

References

[1] P. Berkhin, "A Survey of Clustering Data Mining Techniques", *Grouping Multidimensional Data. Springer*, no. c, pp. 25–71, 2006.

[2] S. H. Liao, P. H. Chu, and P. Y. Hsiao, "Data mining techniques and applications - A decade review from 2000 to 2011", *Expert Systems with Applications*, vol. 39, no. 12, pp. 11303–11311, 2012.

[3] Q. Hen, "A Review of Clustering Algorithms as Applied in IR", *Technical Report IUIUCLIS--1999/6+IRG*, pp. 1–33, 1999.

[4] G. Sreenivasulu, S. V. Raju, and N. S. Rao, "Review of Clustering Techniques", *Proceedings of the International Conference on Data Engineering and Communication Technology Advances in Intelligent Systems and Computing 468*, pp. 523–535, 2017.

[5] G. W. Milligan and M. C. Cooper, "Methodology Review: Clustering Methods", *Applied Psychological Measurement*, vol. 11, no. 4, pp. 329–354, 1987.

[6] L. Candillier, I. Tellier, F. Torre, and *O.* Bousquet, "Cascade evaluation of clustering algorithm", *Springer Verlag. 17th European conference on machine learning*, pp. 574–581, 2010.

[7] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1650–1654, 2002.

[8] N. Settouti, M. Bechar, and M. Chikh, "Statistical Comparisons of the Top 10 Algorithms in Data Mining for Classification Task", *International Journal Of Interactive Multimedia And Artificial Intelligence*, vol. 4, no. 1, pp. 46–51, 2016.

[9] A. M. Jalil, I. Hafidi and L. Alami, "Comparative Study of Clustering Algorithms in Text Mining Context", *International Journal Of Interactive Multimedia And Artificial Intelligence*, vol. 3, no. 7, pp. 42–45, 2016.

[10] K. Swan, "Learning Analytics and the Shape of Things To Come", *The Quarterly Review of Distance Education*, vol. 17, no. 3, pp. 5–12, 2016.

[11] Q. Nguyen, D. T. Tempelaar, B. Rienties, and G. Bas, "What learning analytics-based prediction models tell us about feedback preferences of students", *The Quarterly Review of Distance Education*, vol. 17, no. 3, pp. 13–33, 2016.

[12] S. De Freitas *et al.*, "Foundations of dynamic learning analytics: Using university student data to increase retention", *British Journal of Educational Technology*, vol. 46, no. 6, pp. 1175–1188, 2015.

[13] O. J. Oyelade, O. O. Oladipupo, and I. C. Obagbuwa, "Application of k Means Clustering algorithm for prediction of Students' Academic Performance", *International Journal of Computer Science and Information Security*, vol. 7, no. 1, pp. 292–295, 2010.

[14] J. A. Reyes, "The Skinny on Big Data in Education: Learning analytics simplified", *TechTrends*, vol. 59, no. 2, pp. 75–80, 2015.

[15] R. Klamma, "Community Learning Analytics – Challenges and Opportunities", *Advances in Web-Based Learning – ICWL 2013. Lecture Notes in Computer Science. Springer*, vol. 8167, pp. 284–293, 2013.

[16] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art", *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, vol. 40, pp. 601–618, 2010.

[17] M. Scheffel, H. Drachsler, S. Stoyanov, and M. Specht, "Quality Indicators for Learning Analytics", *Journal of Educational Technology & Society*, vol. 17, no. 4, pp. 124–140, 2014.

[18] R. S. J. d. Baker and G. Siemens, "Learning analytics and educational data mining: Towards Communication and Collaboration", *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12*, pp. 252-254, 2012.

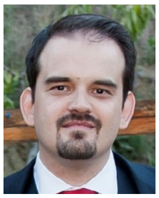
[19] G. Brock, V. Pihur, S. Datta, and S. Datta, "clValid : An R Package for Cluster Validation", *Journal Of Statistical Software*, vol. 25, no. 4, pp. 1–28, 2008.

[20] A. Banerjee and R. N. Davé, "Validating clusters using the Hopkins statistic", *Proceedings of 2004 IEEE International Conference on Fuzzy Systems*, pp. 149–153, 2004.

[21] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic", *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 2. pp. 411–423, 2001.

Alvaro Martínez Navarro

Alvaro Martínez Navarro was born in Pasto, Colombia in 1974. Computer Engineer from Universidad Mariana since 1998. He received his master degree from Universidad de Nariño in Teaching of Higher Education. Candidate to Master in Analysis and Visualization of Massive Data / Visual Analytics & Big Data from Universidad Internacional de La Rioja. He is an Associate Professor at the Universidad Mariana and he is member of the research group GISMAR (B ranking). His research area is analysis and visualization of massive data, learning analytics and higher education in engineering.



Pablo Moreno-Ger

Dr, Pablo Moreno-Ger was born in Madrid, Spain, in 1981 and got his computer engineering degree from Universidad Complutense de Madrid in 2004. He received his PhD. in computer science in 2007 in the Department of Software Engineering and Artificial Intelligence at Universidad Complutense de Madrid. He is the Vice-Dean for Academic Affairs at the School of Engineering and Technology at UNIR. Previously he worked at Universidad Complutense de Madrid, where he was a member of the e-UCM research group and Vicedean for Innovation at the School of Computer Engineering. He has a strong research record in Technology-Enhanced Learning, AI, Machine Learning and Learning Analytics, and has published over 150 academic articles in these topics. He currently holds the IBM Chair on Data Science in Education (http://research.unir.net/ibmchair/) and is a member of the IAR Research Group in AI and Robotics (http://gruposinvestigacion.unir.net/ia/).

StuA: An Intelligent Student Assistant

Pooja Lodhi, Omji Mishra, Shikha Jain*, Vasvi Bajaj

Department of Computer Science, Jaypee Institute of Information Technology, Noida (India)

Received 7 September 2017 | Accepted 23 January 2018 | Published 16 February 2018

Abstract

With advanced innovation in digital technology, demand for virtual assistants is arising which can assist a person and at the same time, minimize the need for interaction with the human. Acknowledging the requirement, we propose an interactive and intelligent student assistant, StuA, which can help new-comer in a college who are hesitant in interacting with the seniors as they fear of being ragged. StuA is capable of answering all types of queries of a new-comer related to academics, examinations, library, hostel and extra curriculum activities. The model is designed using CLIPS which allows inferring using forward chaining. Nevertheless, a generalized algorithm for backward chaining for CLIPS is also implemented. Validation of the proposed model is presented in five steps which show that the model is complete and consistent with 99.16% accuracy of the knowledge model. Moreover, the backward chaining algorithm is found to be 100% accurate.

Keywords

Artificial Intelligence, Backward Chaining, CLIPS, Expert System, Rule-based System, Virtual Assistant.

DOI: 10.9781/ijimai.2018.02.008

I. Introduction

Artificial Intelligence is the science of making computers perceive their environment in a similar way as a human does and takes actions [1]. Artificial Intelligence is not a cognitive process rather it has an array of separate components of intelligence like learning, problem-solving, perception, and language understanding. Using artificial Intelligence, researchers are creating various systems that can mimic decision -making abilities of human based on past experiences and understand human thoughts. Such systems consist of a knowledge-base, inference engine, and user-interface. The knowledge-base has facts i.e., past experiences whereas inference engine has a set of rules that can infer a given situation while the user interface is used to interact with the user. Knowledge-base represents the knowledge that is stored by a model and is used to conclude new rules and check inconsistencies. The development of most of the virtual and intelligent software systems is implemented in CLIPS(C Language Integrated Production System, first released by NASA in 1988) [2-3] programming environment. CLIPS is a programming tool which is designed to ease the development of software that can model human knowledge. CLIPS program is used because it is flexible, expandable and has low cost. It is the only expert system shell for which it is claimed that the shell has been certified correct [27].

Virtual Assistants refer to software that can provide required information or perform work for human. There are mainly two types of virtual assistants: rule-based and stochastic. When there is lack of data, expert knowledge is used to design rule-based assistants also known as expert systems. While in presence of already tested large and valid information, stochastic virtual assistant can be designed using various machine learning techniques. A number of virtual assistants exist in different domains. A number of rule-based assistants are working in different domains such as counseling [4-7], prediction [8-10], diagnosis [11-19], design [20-21, 30], e-learning [28-29] and recommendations [31]. Some of the stochastic assistants are Chatbots, question answering systems and assistants like Apple's Siri, Google's Google Assistant, Amazon's Alexa, Microsoft's Cortana and Facebook's M.

* Corresponding author.

E-mail addresses: shi_81@rediffmail.com

All these virtual assistants are designed with an aim to replace a human being and assist the people around. These systems are doing really well in their domain. In the world of technology, people are more comfortable in communicating with a computer or any other electronic device than a human being. Sometimes people hesitate to interact and ask queries to a human being as either they want to keep it secret or they feel shy. Sometimes they want somebody's help but are not able to reach a right person to seek a help. This motivated us to design an interactive and intelligent student assistant situated in a close environment. It helps students who are new to the college environment. When a newcomer or a fresher enters in a college, he/she may not be familiar with the environment and seeks for the guidance from some experienced person like senior or teacher. But at the same time, some of them fear of being ragged or hesitate to ask. Hence, the new-comers remain unaware of the college environment. In such a scenario, the proposed student assistant, StuA, helps the newcomers and tries to provide the required information in a safe and sound manner to the students. It helps in familiarizing them with the rules and regulations of the college in a comfortable manner. StuA is equipped with the unique features which allow the user to ask questions on WHAT IS, WHERE, WHEN, WHAT HAPPEN format. Moreover, it facilitates real-time, low-cost expert-level assistance 24X7, unlike a human. The system based on its previously gained knowledge and beliefs is able to provide answers to most of the queries. The set of questions is not limited to only the prefixed questions. Nevertheless, to the best of our knowledge, no such virtual intelligent assistant exists till date. Moreover, it can be further customized to handle the queries of a new person in any environment such as a new office, organization, reception of hotels, hospitals, schools and malls.

StuA is a rule-based virtual assistant in the domain of education. The reason of designing it as rule-based is the lack of documented information. The rule-base system is designed with the help of experts and general information in this domain. The implementation has been done in CLIPS. The main limitation of CLIPS is that it supports only forward chaining and does not allow backward chaining. As a result, many of the queries cannot be processed. Till date, no generic extension of CLIPS with backward chaining exists. A few researchers [22-23] have tried to add this functionality but for the specific domains only. So, here, we propose a generic model of backward chaining also, to extend the functionality of CLIPS so as to provide better inference mechanism. The whole model is designed using Java as the user interface. When java is integrated with CLIPS it provides the flexibility to redefine the output and with help of java, we can assert only relevant facts reducing the overhead of searching through irrelevant facts. It also aids the user by providing a friendly user-interface.So now, only the relevant rules are fired and the filtered and processed output is displayed.

The paper is organized as follows. In Section 2, the related work is discussed followed by the detail explanation of the proposed model (Section 3). Section 4 presents the implementation detail of the proposed model and Section 5 discusses the test results and validation of the proposed model. The Section 6 concludes the paper with limitations and provides direction for future work.

II. Related Works

Till date several systems are designed in various domains [4-21]. One of the expert systems in education domain is developed to help students in selecting the best branch who are planning to take admissions in engineering; this system is called Student Counseling System (SCS). Here they have used certainty factor to provide basis to their judgment. They have only chosen engineering branch as their domain for their counseling system. This expert system answers the queries in the form of 'yes' or 'no' only using forward chaining [4]. This system is rule-based system which uses CLIPS to answer the queries.

Another domain is diagnosing some of the eye diseases. It provides the expert guidance on eye diseases. The disadvantage of the system is that it is not taking the symptoms as input from the user. Rather, the expert system prompts menu and user has to select from it and, based on the disease, user has to answer queries in yes/no form [11]. Moreover, it is able to answer only the queries that can be derived either from the facts or from forward chaining. Various expert systems are being designed in the field of medical sciences. One of them is plant disease diagnosis system. They have used two methods to diagnose plant disease. One is Step by Step description and other is Graphical Representation System. The limitation of step by step description is the user is not allowed to enter his query, rather he has to select from a list of options. Hence search query is limited. Graphical Representation System provides more meaningful results when searched with keyword disease. Here the user has to only enter keywords. The limitation is he cannot enter the full description of disease [13]. The system proposed was the rule-based system which uses forward reasoning and pattern matching to answer the queries. Nevertheless, they concluded that expert systems is one of the successful methods which helped and supported users in making the right decisions in scenario where they have lack of knowledge.

Another expert system in the row deals with Diagnosis of Neuromuscular Disorders [15]. The proposed system is implemented using JESS. Here, the user is presented with a list of questionnaires about symptoms and possible treatments are suggested according to disorder. They designed different rules for different disorder on basis of knowledge acquired from experts. The limitation of this work is that it can detect only Cerebral Palsy, Multiple Sclerosis, Muscular Dystrophy and Parkinson's disease. Even other small diseases can't be detected using this system. Another research work deals with the work of diagnosis of rice plant disease. This expert system is also developed using JESS [16]. They have used SQL to store the data and extract information stored.

Similarly, there are different research works available in the field of system that assists people known as virtual assistants. Another example in domain of education is an expert system which aimed to assist Student in there major selection process. It is a prototype of an advisory expert system which can assist new and incoming students to select suitable majors they can apply to and the most convenient institutions they can attend. In this research work, user has to select from different courses and subjects and after confirmation final output is shown. They have designed a rule-based system in which they majorly defined three broad categories of rules to assist the students. The drawback of their work is that they have not tested their work on real-world cases; they have used fabricated test-cases to simulate their results [5]. A system PAS (Postgraduate Advisory System) is proposed which enables the students to select and get a plan for each semester without the need to consult advisors [6].The proposed system was different from other systems as it responded taking into account students thesis fields. The system is limited to one department; its rules are only defined for one department. Another expert system that is developed worked in Car Failure domain. Here, the author has categorized problems broadly into three categories. The drawback of the system is that user has to choose from a limited number of options [14]. He cannot input his own query.

TABLE I. Existing Rule-Based Systems

Domain	Paper	Category	Type	Criticism
Education	SCS [4]	Question Answering	Rule-Based	Only answer query in form of yes or no.
	Student Major Selection [5]	Assistant	Rule-Based	Application is not tested on real-world cases; fabricated test-cases used to simulate the results.
	PAS [6]	Assistant	Rule-Based	Limited to one department.
Medical Sciences	Eye Diseases Diagnosis [11]	Question Answering	Rule-Based	Symptoms are not taken as input from the user. User has to select from a list of symptoms designed by the expert. And it only answers query in form of yes or no.
	Plant Disease Diagnosis System [13]	Question Answering	Rule-Based	User is restricted to select from limited set of queries.
	Diagnosis Of Neuromuscular Disorders [15]	Question Answering	Rule-Based	This system only detects a limited number of disorders.
	Diagnosis Of Rice Plant Disease [16]	Question Answering	Rule-Based	-
Industry	Car Failure [14]	Assistant	Rule-Based	User is restricted to select from limited set of queries.

Moreover, there are virtual assistant such as Siri, Cortana, and Google assistant which are using some machine learning techniques to learn the procedure of doing the task. Cortana mostly answers general questions for which it pulls information from Bing. Whereas, Siri assists in all kind of works in a device like calling people, sending text messages, setting reminder, etc.

In conclusion, the virtual assistants can be designed in two ways: rule-based and machine learning based. For the domains, where data is not available in the form of corpus, it is difficult to apply some machine learning technique for decision making. In such cases, expert knowledge is required to design the systems (also known as expert systems). Most of the expert systems are using CLIPS and hence they are able to answer the queries using forward chaining only. They provide a fixed set of queries and user has to select one out of them only to get the answer which restricts the functionality of the system. Table I presents the summary of existing rule-based systems.

III. Proposed Model

Recognizing the need of a virtual assistant for the new-comers in a college, we propose an interactive and intelligent student assistant, StuA, which can help the new students in familiarizing them with the new environment, rules and regulations. It is a rule-based assistant because of lack of documented information. Furthermore, it tries to answer the queries of the students using inference mechanism without fixing the set of queries and hence providing a broader and unrestricted platform to resolve the doubts. The user is allowed to ask anything related to the domain and he need not pick the question from the already framed list. The domain knowledge is stored in the form of knowledge-base drilled by knowledge engineers. The proposed assistant responds to the question not only using forward chaining. Rather, the CLIPS tool is extended to empower the inference mechanism of the tool with backward chaining algorithm as well. The extended version of CLIPS with backward chaining is capable to handle most of the queries such as

Inference rule: Student break rule => student has to pay fine

Query: when student has to pay fine?

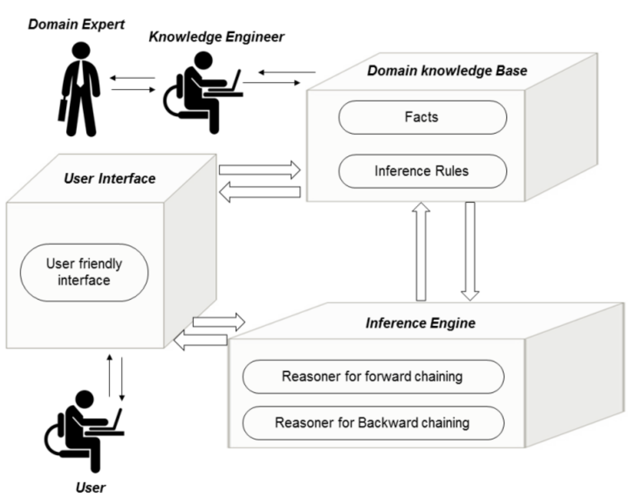


Fig. 1. Overall architecture of the proposed model, StuA.

The overall architecture of the proposed model is presented in Fig. 1. In the model, the knowledge engineer designed the knowledge-base in collaboration with the domain expert. The knowledge is stored in the form of facts and inference rules. The inference engine works with forward chaining as well as backward chaining. A user-friendly interface is provided to the user where he interrogates. To simplify the process, some drop downs are provided. The user asks for a query using the user interface. It is then treated by the inference engine. At this stage, first, the type of the query is analyzed. Depending upon the type of the query inference mechanism is carried out either through facts directly, or through forward chaining or the backward chaining using the inference rules. The result is then displayed to the user. Unlike CLIPS, only the relevant rules are fired and the filtered and processed output is displayed. The filtering and processing of output are being carried out by integrating CLIPS with java.

A. Domain Knowledge-Base

Domain specific knowledge is considered as bottleneck information in building knowledge-based systems. The model is trained with minimum possible knowledge. It often encountered missing data in running new test examples thus using inference mechanism it is able to answer them. Here, Information is handcrafting which is the simplest way to put knowledge into program. The focus is mostly on gathering knowledge from human experts, college website and through feedback from students.

One of the key points while constructing a virtual assistant is transparency i.e., making the system understandable despite the complexity of task. This is because:

- The system improves through consecutive development, which requires thorough understanding of earlier versions.
- The system improves through criticism from the people who are not familiar with its implementation details.
- The system uses its own learning methods for solving the problems.

After acquiring knowledge from human experts, it maps the knowledge so that it can be used in the program. Knowledge is basically in the form of sentences which is then broken into subject, object and predicate so that they could be programmed in the knowledge-base.

B. Inference Mechanism

A survey was conducted to observe the type of query any newcomer could ask and four major classes of questions is identified. The first class is "TRUE/FALSE" which is used to tell whether an asked fact is true or not. For example "Student can return book on Monday ", such statement is either true or false. The second class of question is "WHAT-IS or WHERE". This class means that these questions are related to some atomic knowledge about the environment. For example, we have some information like "LRC is learning resource center. The minimum CGPA is 4", these are some of the atomic information which cannot be further broken. So the query can be "what-is minimum CGPA?" and this can be directly extracted from the information available. Similarly, the third class is "WHAT-HAPPEN". This class identifies the consequences of some situation or helps to know the outcome of some event. Suppose we have some fact like "if student fails supplementary examination then they get back in that year", in this fact query could be, "what happens if student fails supplementary examination?" This also led to the identification of the fourth class of question that is "WHEN" which is meant to know the cause for some event. For example, we have a fact that "if attendance is less than 60%, students get debarred." In such case, the user can ask the query that "when student get debarred?" So, for all these situations of an environment, the queries are divided into various classes. All the queries can be broadly classified into these four classes only. However, some queries can also be resolved in one level inferencing while others may require deep reasoning. Both are handled in the model.

The proposed model has a rule and fact based inference system. In this model, the TRUE/FALSE class is directly handled by CLIPS. The information is stored as fact in Knowledge Base (KB) and to answer that the fact is true or not, KB is searched. The information handled by "WHAT-IS or WHERE" class is atomic knowledge and can be inferred directly thus they are stored as facts in the knowledge-base of our model. The "WHAT-HAPPEN and WHEN" class has the information stored in the form of rules. In our model, the rules are typically structured as antecedent with their consequents. The inference engine examines the type of class to which the query belongs and then responds accordingly by executing the corresponding consequents or by searching for the correct antecedent. Forward chaining approach is used to infer queries of WHAT-HAPPEN class as in this class we want to know the consequent of a situation. Forward chaining is an inferencing method. It uses the available data and inference rules to extract more data until it is able to find the goal. This method is also called Data-driven as the data determine which rules should be selected and used. For example, the following information is present in the KB:

Initial fact:

Grade is F

Rules:

If (grade is F) then (student fails subject).

If (student fails subject) then (has to take supplementary exam)

If we want to know "what-happen-if grade is F". For this, we need to search for the consequent with the help of forward chaining. Similarly, backward chaining approach (explained in section III.C) is used to answer the query of WHEN class. Here, the goal is to decide that can we infer the fact "when student has to take supplementary exam" from the initial facts or not? With the help of backward reasoning, it tries to prove the goal (i.e. student has to take the supplementary exam). Hence, we have to show that antecedent (i.e. student fails subject) can be proved. Now, this becomes our new sub goal and so on. This continues until the initial fact as the antecedent is found.

To accommodate all these types of queries, an inference engine is proposed which is made up of query processor and different types of processing models as shown in Fig. 2. This inference engine does all the inferencing. It selects correct processing model for a given query. There are four models for query processing i.e. True False (TF) processor, Fact processor, Conclusion processor and Backward Chaining (BC) processor. The TF processor answers query for the TRUE/FALSE class of query. The Fact processor extracts the fact from the belief base to provide an answer. It handles the WHAT-IS and WHERE class of queries. The Conclusion processor infers a fact from the knowledge-base on the basis of the query provided to the inference engine. This processor uses forward chaining for the processing and thus handles the WHAT-HAPPEN class of queries. The BC processor uses the backward chaining approach to infer the facts for answering the WHEN class of queries.

The model is implemented with help of CLIPS and JAVA. CLIPS tool is integrated with JAVA with help of CLIPSJNI [18]. This integration helped in overcoming some limitation of CLIPS. First of all, it suppressed the triggering of all rules. Secondly, it also suppressed the assertion of irrelevant facts. As a result, the response time is also improved and only the relevant information is shown to the user. Moreover, CLIPS doesn't have a feature of backward chaining also. So, using JAVA, CLIPS is extended with a generic form of backward chaining which can provide a first level explanation of the queries.

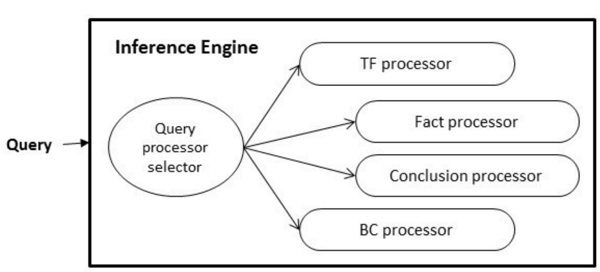


Fig. 2. Inference Engine of Proposed Model.

C. Extended CLIPS Tool

Backward chaining is an inference methodology in which a conclusion is available and moving backward to find the base facts supporting the conclusion is possible. CLIPS is a public domain software tool which is widely used for building intelligent systems [2-3]. It combines the programming paradigms of procedural, object oriented and logical languages, but it does not support backward chaining [24-25]. Backward chaining is important as it is a faster approach and it is more suitable for goal driven queries.

There are various queries that require backward chaining. For example, if someone asks "when they can get debarred" then there exists a rule "if attendance less than 60% then get debarred" in the knowledge-base. So here, in this case, given the consequent, antecedents should be determined. Such type of queries is handled with help of backward chaining. It is observed that mostly "WHEN" type of queries can be answered using backward chaining and they are being processed by BC processor module of the proposed model.

Algorithm I: backward_chaining

```
Input: Query
Output: Response

Declare environment variable clips1, clips2
Load facts in clips2
Load backward chaining rules in clips1
Add query fact in clips1
fact_list← run clips1
if fact_list length ≤ 1
        ans← "nothing can be inferred"
end if
else
     connection_list← run clips1
     for each fact fiϵ fact_list do
          v← extract value from fi
          new_query← create query from v
          append v in bc_factlist
          fl2← run clips2
          if fl2 ≠ ϕ then
               append v in delete_fact_list
          end if
     end for
end else

for each Di ϵ delete_fact_list do
     if bc_factlist contains Di then
          remove from bc_factlist and clips1
     end if
end for

settingValue (first element of connection_list)

for each Fi ϵ connection_list do
     if correct_flag of Fi is True then
          ans← fact connected to Fi
     end if
end for
```

Algorithm II: settingValue(connection_list element)

Input:connection_list
Output: correct_flag and check_flag for all elements in connection_list are set accordingly

Declare correct_flag← true and check_flag← false
if check_flag of Fi is true then
return
end if
If correct_flag of Fi is false then
check_flag of Fi ← true
end if
for each Fi ϵ connected fact of Fi in connection_list do
if bc_factlist not contains connected fact then
H ← Search for connected fact of Fi in connection_list
if H = ϕ then
correct_flag← false and check_flag← true
break loop
end if
else
settingValue (H)
end else
end if
end for

In this algorithm, firstly, a set of basic facts is created that exists in our knowledge-base with help of bc_factlist variable. Also, a list named connection_list is used to maintain the relationship between different base and derived facts. In settingValue() function recursive call is used to successfully infer a set of facts for a given query. This is done by searching for a given fact in the base facts list. For a derived fact, a recursive call is made. Through this recursive call, the facts deriving the other facts are being searched in the fact-base list.

IV. Implementation

The proposed model, StuA, is implemented using JAVA and CLIPSJNI. A knowledge-base is created with various facts and rules related to a specific domain i.e. the college environment. It takes the query as an input. The query is provided as an input to inference engine which distinguishes the query class and sends it to the appropriate processor. The processor answers the query by forming a simple sentence using a first level of natural language generation. The user interface is shown in Fig 3(a)-3(d).

Fig. 3(a) shows the simulation of TRUE/FALSE class. It uses the TF processor to process the query. The TF processor searches if the fact asked exists in the knowledge-base or not accordingly answer the query. In Fig. 3(b), the query of WHAT-IS or WHERE class is being processed by Fact processor. It searches for facts matching the given information and extracts the rest of the information from there. It further uses that extracted information to answer. The WHAT-HAPPEN class of query is performed by conclusion processor as shown in Fig. 3(c). It processes the query by finding all the relevant facts that could be inferred. It uses forward chaining for this process. It finally sorts the inferred facts and answers the query. Fig. 3(d) shows the simulations related to WHEN class. This class of queries is processed using the BC processor in which the proposed backward chaining method is implemented. This method searches for the initial facts in the fact-base which could be reached from the goal statement and stores the first level information related to the particular goal. This first level information is processed to form a sentence if an initial fact is successfully found. Table II lists some questions posed to the virtual assistant, StuA.

TABLE II.
Sample Question Answers Processed By The Proposed Model, StuA

Question asked by user	Answers given by StuA
Where is LT1	Ground floor have LT1
What is min CGPA	Min. CGPA is 4.0
What happen student return book late in LRC	Student has to pay fine
When student get debarred	Attendance less than 60
What happen student fails supplementary exam	Student fails supplementary exam implies student get back , have to take summer semester, backlog
How can one access study material	Connect to college intranet and use the login details

V. Verification and Validation

The testing process of any automated tool includes verification and validation. Verification checks the completeness of the knowledge-base. Validation checks the correctness of the knowledge-base in terms of consistency. As suggested by Wentworth et al. [27] and Ghasem & Alizadeh [26], firstly the logical completeness and logical consistency of the knowledge-base is checked. Then, the knowledge model is validated followed by the validation of semantic consistency of knowledge items. Finally, the backward chaining algorithm integrated with CLIPS is validated.

A. Logical Completeness

Logical completeness means the expert system produces some conclusion for all inputs. This can be done by following the below mentioned steps [27]:

1. Constructing a logical formula that represents conditions under which the system is complete; this logical formula will be called the completeness formula in conjunctive normal form.
2. Eliminate ORs containing logical opposites or all possible values of a variable.
3. If the resulting logical expression is TRUE, the system is complete.

We checked the completeness of all the subsystems as specified above and found them COMPLETE. For illustration, Study_Material_Accees subsystem completeness check is shown below:

To show the completeness of Study_Material_Accees subsystemof Knowledge Base, firstly construct the logical formula:
Study_Material_Accees = YES
OR Study_Material_Accees = NO
Expressing in terms of input conditions:
(Connected_internal_Network= YES
andHave_Login_Access = YES)
OR
(Connected_internal_Network= No
OR Have_Login_Access = No)
Writing this in conjunctive normal form gives
(Connected_internal_Network= YES
OR Connected_internal_Network= No
OR Have_Login_Access = No)
AND
(Have_Login_Access = YES
OR Connected_internal_Network= No

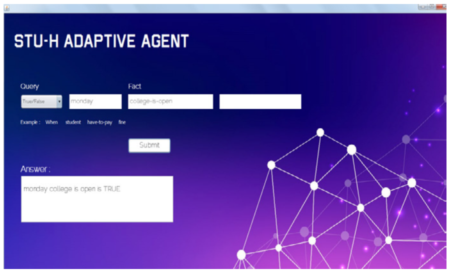


Fig. 3(a). Simulation of TF processor.

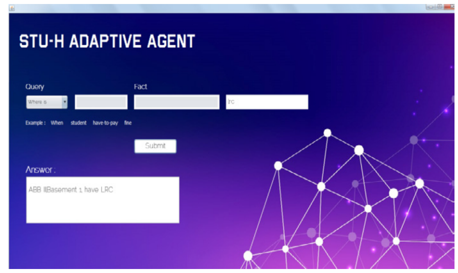


Fig. 3(b). Simulation of Fact processor.

Fig. 3(c). Simulation of Conclusion processor.

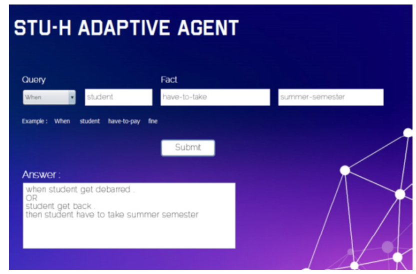


Fig. 3(d). Simulation of BC processor.

OR Have_Login_Access = No)

The first term is TRUE because yes and no are the only possible values for Connected_internal_Network. Likewise the second term is also TRUE. Therefore, the formula expressing completeness of Study_Material_Accees is TRUE, and hence the subsystem is complete.

B. Logical Consistency

Logical consistency means for all inputs, the knowledge base produces a consistent set of conclusions, i.e., that for each set of possible inputs, all the conclusions can be true at the same time. To establish consistency, the user must do the following [27]:

1. Construct a logical formula that represents conditions under which consistency fails; this logical formula will be called the consistency formula. Write this formula in disjunctive normal form.
2. Eliminate ANDs containing logical opposites or other contradictory sets of conjuncts.
3. If the left hand side of the resulting logical expression is FALSE, the system is consistent.

We checked the consistency of all the subsystems as specified above and found them CONSISTENT. For illustration, Study_Material_Accees subsystem consistency check is shown below:

To *show the consistency of Study_Material_Accees subsystem of Knowledge Base, firstly construct the logical formula. The only set of inconsistent conclusions is*

Study_Material_Accees = YES

AND Study_Material_Accees = NO

Expressing in terms of input conditions:

(Connected_internal_Network= YES

andHave_Login_Access = YES)

AND

(Connected_internal_Network= No

OR Have_Login_Access = No)

Writing this in conjunctive normal form gives

(Connected_internal_Network= YES

AND Have_Login_Access = YES

AND Connected_internal_Network= No)

OR

(Connected_internal_Network= YES

AND Have_Login_Access = YES

AND Have_Login_Access = No)

*The first term is FALSE because yes and no are contradictory values for Connected_internal_Network. Likewise the second term is also FALSE. Therefore, the formula expressing consistency of Study_Material_Accees is FALSE, and hen*ce the subsystem is consistent.

C. Knowledge Models Completeness Check

Logical completeness and consistency are necessary but not sufficient for a knowledge model to be complete. It should be *semantically complete* as well, i.e., it must base its decisions on all information considered to be relevant by the expert [27].

One of the ways to check the completeness of a knowledge model is to create a knowledge model with a single expert and review the knowledge model with other experts who are not connected with the development of the model. Following the same, the knowledge model is created by knowledge engineers and for completeness check, we selected 70 final year students (46 males and 24 females) of the college as they seem to be the best suited experts in this domain. They were asked to use the tool for one whole day and check for the correctness of the answers given by the virtual assistant. All the 70 students, in total, posed 836 questions to the assistant and found 829 answers correct. All of them were satisfied with the working of the tool and 99.16% accuracy is reported.

D. Validating the Semantic Consistency of Underlying Knowledge Items

Even if the expert knowledge has been properly encoded into an expert system knowledge-base, the KB will probably produce errors if the underlying expert knowledge is wrong. Therefore, it is important to validate the expert knowledge behind the knowledge-base. This can be done by checking the confidence level of experts.

The basic method for validating a knowledge item is [27]:

- Ask a panel of experts whether it is true or false.
- Tally the TRUE/FALSE answers.
- Analyze the results statistically.

This test is conducted on the same population of experts. All the 70 students were asked to answer 35 question (excerpts shown in Table III) as yes or no. All the experts agreed upon all the 35 questions unanimously. The confidence level is computed using the formula

$$\text{Confidence Level} = 1 - (1 / 2^{**}N)$$

Where N is the number of experts. In this experimentation, 100% confidence is reported.

TABLE III. Excerpts from True/False Questionnaires

Questions	**T/F**
If student has not given T1 or T2, he can give makeup exam	True
If student is debarred, he can give supplement exam	False
If student is debarred, he can appear for summer semester	True
If student fails, he can give supplement exam	True
If CGPA is less than 4, student get year back	True
If students returns book late in LRC, he need not to pay fine	False
100% weight-age is given in makeup exam	False

E. Backward Chaining Algorithm Validation

As a next step of validation, the model is simulated for the pre-simulated domains [22] and [23]. In this step, mainly backward chaining module was validated. A few existing models have implemented the domain specific backward chaining. To test our generic model, we used those simulations. Table IV shows the testing results of the proposed model for backward chaining design in CLIPS with the help of JAVA. In table IV, the output of some of the pre-simulated examples of backward chaining is compared with our proposed model. Through this testing, it is concluded that the proposed model works perfectly in various scenarios. 100% correctness for all the examples is achieved and with an advantage of providing single level explanations.

VI. Conclusion

People sometimes hesitate to interact with a stranger and ask their queries in a new environment. A virtual assistant provides a solution for this. In this paper, we proposed an interactive and intelligent student assistant, StuA, situated in a specific domain (i.e. the college environment) where it is capable of answering all types of queries of a new-comer to make him/her familiar with the new environment. It facilitates real-time, low-cost expert-level assistance with 24X7 availability. The model is designed using CLIPS as it allows inferencing. Further, it provides a one-level explanation of queries, which is its advantage over other existing models. We have worked on the limitation of CLIPS by proposing and implementing a generic model of backward chaining. The model is validated on various pre-simulated examples which gave 100% correctness. Further, successful checks are performed for logical completeness and consistency of the knowledge-base. Further, semantic completeness is also checked with the help of 70 domain experts and found the accuracy of 99.16%

The proposed system is restricted to the domain of college and

TABLE IV. Comparison of Existing Specific Model with the Proposed Generic Model

Pre-simulated scenarios Results	**Results generated by the proposed model**
Defining defrule: BCC__pushBacktrack3a TRUE CLIPS> (load "test1.clp") Defining defrule: start +j Defining defrule: rule1 +j TRUE CLIPS> (reset) CLIPS> (run) Goal go home Achieved CLIPS> _	Backward chain file created Enter query? - go-home - facts got matched are : (<Fact-1> <Fact-2>) Facts that got matched after infrencing .... 2 rv = "-~1/-2" when finish work . then go-home BUILD SUCCESSFUL (total time: 2 seconds)
TRUE CLIPS> (load "test2.clp") Defining defrule: start +j Defining defrule: rule1 +j TRUE CLIPS> (run) Goal go home Failed CLIPS> _	run: Backward chain file created Enter query? - go-home - facts got matched are : (<Fact-1> <Fact-2>) Facts that got matched after infrencing .... 2 rv = "-~1/-2" Don't know when go-home
CLIPS> (reset) CLIPS> (run) move from start to 1 move from 1 to 2 move from 2 to 3 move from 3 to 4 move crate climb get eat Goal monkey full Achieved CLIPS>	run: Backward chain file created Enter query? - monkey-full - facts got matched are : (<Fact-1> <Fact-2> <Fact-3> <Fact-4> <Fact-5> Facts that got matched after infrencing .... 15 rv = "-~1/-2-~2/-3-4-~3/-5-6-4-~6/-5-7-4-~5/-8-9-~8/-10-11-~10/-12-13- when bananas location withMonkey. then monkey-full BUILD SUCCESSFUL (total time: 10 seconds)
<== f-159 (goal is stationary) Is your animal attached permanently to an object? (yes no) no ==> f-160 (variable stationary no) <== f-56 (question stationary is "Is your animal attached permanently to an object?") FIRE 64 remove-rule-no-match: f-160,f-158 <== f-158 (rule (if stationary is yes) (then family is stationary)) FIRE 65 rule-satisfied: f-160,f-157 <== f-157 (rule (if stationary is no) (then type.animal is jellyfish)) ==> f-161 (variable type.animal jellyfish) FIRE 66 goal-satified: f-1,f-161,f-129 <== f-1 (goal is type.animal) I think your animal is a jellyfish FIRE 67 propagate-goal: f-133,f-67	Backward chain file created Enter query? type.animal is jellyfish facts got matched are : (<Fact-1> <Fact-2> <Fact-3> <Fact-4> Facts that got matched after infrencing .... 10 rv = "-~1/-2-3-~2/-4-5-~4/-6-7-~6/-8-9-~8/-10" when order is cells & stationary is no. then type.animal is jellyfish BUILD SUCCESSFUL (total time: 21 seconds)

academics which can be customized to various other domains such as at reception of hotels, hospitals, and offices or in schools, malls or in an organization. The proposed model of backward chaining can further be extended to incorporate loops and other higher level programming components. This paper opens lots of future possibilities. A simpler query-writing mechanism can be designed, which might be achieved through Natural Learning Processing (NLP). Nevertheless, it can be made more adaptive by adding learning. With learning and continuous user interaction, it can become more efficient in answering queries.

References

[1] E. A. Feigenbaum and P. McCorduck, The Fifth Generation, Addison-Wesley, 1984.

[2] J. C. Giarratano, G. Riley, Expert Systems: Principles and Programming, 4th Ed, Course Technology, 2004.

[3] J.C. Giarranto, CLIPS User's Guide, Version 6.22, 1998.

[4] T. Sharma, N. Tiwari and K. Shah, "Student counselling system: A Rule-based expert system based on Certainty factor and backward chaining approach," International Journal of Application or Innovation in Engineering and Management, vol. 2, no. 1, pp. 371-375, 2013.

[5] Y. Bouaiachi, M. Khaldi and A. Azmani, "A Prototype Expert System for Academic Orientation and Student Major Selection," International Journal of Scientific & Engineering Research, vol. 5, no. 11, pp. 25-28, 2014.

[6] Al-Ghamdi, S. Al-Ghuribi, A. Fadel, F. Al-Aswadi and T. AL-Ruhaili, "An Expert System for Advising Postgraduate Students," International Journal of Computer Science and Information Technologies, vol. 3, no. 3, pp. 4529-4532, 2012.

[7] S. S. A. Samy, M. Baraka and A. Baraka, "A Proposed Expert System For Guiding Freshman Students In Selecting A Major In Al-Azhar University, Gaza", Journal of Theoretical and Applied Information Technology, vol. 4, no. 9, pp. 889-893, 2008.

[8] M. Lazzari and P. Salvaneschi, "MISTRAL: an Expert System for the management of warnings from automatic monitoring systems of dams," Proceedings of the 2nd Specialist Seminar on integrated CAD of earthquake resistant buildings and civil structures, 1993.

[9] D. Selva, "Rule-based system architecting of Earth observation satellite systems", PhD dissertation, Massachusetts Institute of Technology, Cambridge, MA, US, 2012.

[10] D. Selva, B. Cameron, and E. F. Crawley, "A rule-based method for scalable and traceable evaluation of system architectures", Research in Engineering Design, vol. 25, no. 4, pp. 325–349, 2014.

[11] S. S. A. Naser, A. Z. A. Ola, "An expert System for Diagnosing eye disease using CLIPS," Journal of Theoretical And Applied Information Technology, vol. 4, no. 10, pp. 923-930, 2008.

[12] Hofmeister, M. Alan, Althouse, R.B. Likins, M. Morgan, D.P. Ferrara, J.M. Jenson, W.R. Rollins and Evelyn, "SMH.PAL: an Expert System for identifying treatment procedures for students with severe disabilities," Exceptional Children, vol. 61, no. 2, pp. 174-181, 1994.

[13] S. S. A. Naser, K. Kashkash and M. Fayyad, "Developing an Expert System for Plant Disease Diagnosis," Journal of Theoretical and Applied Information Technology, vol. 1, no. 2, pp. 78-85, 2008.

[14] A.T. Al-Taani, "An Expert System for Car Failure Diagnosis," Proceedings of World Academy of Science, Engineering and Technology, vol. 7, pp. 457-460, 2005.

[15] R. Borgohain, S. Sanyal, "Rule-based Expert System for Diagnosis of Neuromuscular Disorders," International Journal of Advanced Networking and Applications, vol. 4, no. 1, pp. 1509-1513, 2012.

[16] K. Robindro, S. K. Sarma, "JESS based expert system architecture for diagnosis of rice plant disease: design and prototype development," 4th International Conference on Intelligent Systems, Modelling and Simulation, pp. 674-676, 2013.

[17] B. M. Salem, M. Roushdy and R. A. Hod, "A case based expert system for supporting diagnosis of heart diseases," AIML Journal, vol. 5, no. 1, pp. 33-39, 2005.

[18] S. A. Samy, R. Al-Dahdooh, A. Mushtaha and M. El-Naffar, "Knowledge Management in ESMDA: Expert System for Medical Diagnostic Assistance," AIML Journal, 2010.

[19] S. A. Samy and S. Almursheidi, "A Knowledge Based System for Neck Pain Diagnosis," World Wide Journal of Multidisciplinary Research and Development(WWJMRD), vol. 2, no. 4, pp. 12-18, 2016.

[20] Y. A. Nada, "A Novel Expert System for Building House Cost Estimation: Design, Implementation, and Evaluation," International Journal of Advanced Computer Science and Applications, vol. 4, no. 8, pp. 50-55, 2013.

[21] N. L. Griffin, F.D. Lewis, "A Rule-based Inference Engine Which is Optimal and VLSI Implementable," IEEE International Workshop on Tools for Artificial Intelligence, VA, pp. 246-251, 1989.

[22] J. C. Giarratano, CLIPS user's guide, Artificial Intelligence Section, Johnson Space Center, 1987.

[23] Huyck, http://www.cwa.mdx.ac.uk/bis2040/backchain/backchain.html, dated May, 2017.

[24] S.H. Liao, "Expert system methodologies and applications—a decade review from 1995 to 2004," Expert Systems with Applications an International Journal, vol. 28, no. 1, pp. 93-103, 2005.

[25] Kishan, V. Chadha and C. Maini, "A Review of Development and Applications of Expert," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 2, no. 10, pp. 319-325, 2012.

[26] Armin Ghasem Azar, Zohreh Mohammad Alizadeh, "A Basic Proof Method for the Verification, Validation and Evaluation of Expert Systems," Journal of Information Systems and Telecommunication, vol. 1, no. 1, pp. 27-31, 2013.

[27] James A Wentworth, Rodger Knaus, Hamid Aougab, "Verification, Validation, and Evaluation of Expert Systems", An FHWA Handbook, Volume 1, 1997.

[28] Bhattacharya, S., and S. Nath. "Intelligent e-Learning Systems: An Educational Paradigm Shift," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 4, no. 2, pp. 83-88, 2016.

[29] Secades, V. A., and O. Arranz. "Big Data & eLearning: A Binomial to the Future of the Knowledge Society," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 3, no. 6, pp. 29-33, 2016.

[30] Jain, S. and Asawa, K., "Formalization of Event Perception and Event Appraisal Process," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 3, no. 3, pp. 69-73, 2015.

[31] Choudhar, R.S., Kukreja, R., Jain, N. and Jain, S., "Personality and Education Mining based Job Advisory System," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 2, no. 7, pp. 30-34, 2014.

Pooja Lodhi

Pooja Lodhi is pursing M.Tech (Computer Science) from Jaypee Institute of Information Technology (JIIT), Noida, India. Her area of interest includes Artificial Intelligence, Agent Modelling, Expert Systems and Machine Learning.

Omji Mishra

Omji Mishra is pursing M.Tech (Computer Science) from Jaypee Institute of Information Technology (JIIT), Noida, India. His area of interest includes Expert Systems, Soft Computing and Machine Learning.

Vasvi Bajaj

Vasvi Bajaj is a undergraduate student at Jaypee Institute of Information Technology (JIIT), Noida, India. Her area of interest includes Machine Learning and Artificial Intelligence.

Shikha Jain

Shikha Jain is presently working with Jaypee Institute of Information Technology (JIIT), Noida, India as Assistant Professor. She has more than fourteen years of research and academic experience. She has received her PhD in computer science from Jaypee Institute of Information Technology, Noida, India, in the area of "Emotion modelling for an autonomous agent and its programming". She has published a number of research papers in the renowned journals and conferences. Her research area includes Affective Computing, Emotion Modelling, Cognitive Affective Architectures, Agent Modelling, Machine Learning, Soft Computing, Algorithm Design and Automata theory. She is also member of technical program committee of several conferences.

Predicting Student Performance in Higher Education Institutions Using Decision Tree Analysis

Alaa Khalaf Hamoud, Ali Salah Hashim, Wid Aqeel Awadh *

University of Basra, Basra (Iraq)

Received 3 November 2017 | Accepted 22 January 2018 | Published 9 February 2018

Abstract

The overall success of educational institutions can be measured by the success of its students. Providing factors that increase success rate and reduce the failure of students is profoundly helpful to educational organizations. Data mining is the best solution to finding hidden patterns and giving suggestions that enhance the performance of students. This paper presents a model based on decision tree algorithms and suggests the best algorithm based on performance. Three built classifiers (J48, Random Tree and REPTree) were used in this model with the questionnaires filled in by students. The survey consists of 60 questions that cover the fields, such as health, social activity, relationships, and academic performance, most related to and affect the performance of students. A total of 161 questionnaires were collected. The Weka 3.8 tool was used to construct this model. Finally, the J48 algorithm was considered as the best algorithm based on its performance compared with the Random Tree and RepTree algorithms.

Keywords

Prediction, Students' Success, Decision Tree, J48, Random Tree, REPTree, Weka.

DOI: 10.9781/ijimai.2018.02.004

I. Introduction

Recently, the increasing volume of data and using them to increase students' performance is one the major challenges of higher education institutions (HEIs). Higher education institutions are generally interested in the success of students during their study [1, 2]. Higher education institutions, in a lifetime of teaching, have a large data set of student information stored in their databases. However, storage is not a problem. Handling data, extracting relevant patterns, and discovering knowledge stored in the massive database are tremendously difficult. Accordingly, data mining can be considered as a very promising tool to attain these objectives [3].

Data mining is used to detect and extract relevant as well as worthy information from a very huge volume of data. Currently, this process has acquired a great deal of attention as well as concerns from the information industry and society. This technique is also receiving significant attention in data analysis as well as it been recognized as a newly emerging tool for analysis [4]. Predicting student performance is a very beneficial in providing students as well as lecturers with the necessary assistance in the learning process. Predicting the possible outcomes of the learning process on the basis of the results of prediction can help an organizations change the outcome of new students and performance by adjusting the factors that contributed to the past performance [5].

Predicting of student performance also aids educational planners and decision makers and administrators to adequately plan for changes in student population in any direction (i.e., find factors which increase performance or decrease it). However, coming up with a manual set of rules required to predict student performance is commonly difficult. For these reasons, there is an extreme need to achieve these goals [6]. Particularly, data mining tools became very popular among researchers and users because of their ease of use and availability such as the following tools (most of them are available online) Microsoft Excel, SPSS, Weka, Protege (a knowledge acquisition system) and Rapid Miner. A number of these tools (e.g., MS Excel-based) are freely available to HEI professors with which they can benefit with their existing knowledge of the Excel application. The accessibility, availability, ease of use and understandability are the most reasons for the inclusion these tools in the research. Weka is chosen as a data mining analysis tool for supporting conclusion because of its highly readable and understandable results [7].

We aim to analyze collective student information through a questionnaire (based on LimeSurvey and Google Form) as well as classify the collected data to predict and categorize student performance. We also seek to elucidate the different factors that affect student performance (success and failure rates) in relation to other variables in the data set of students by applying decision tree algorithms. The work seeking to find the possibility of depending decision tree algorithms' results in support academic decisions related to improve academic performance and give a proper road map for students and lecturers. Therefore, this paper uses multiple classification and prediction methods to confirm and verify the results with multiple Weka classifiers. We select the best result in terms of accuracy and precision based on performance results. This newly discovered knowledge can help students as well as instructors in carrying out enhanced educational quality. Identifying possible underperformers at the beginning of the semester/year and increasing the attention allotted for them will aid their educational process as well as improve students' grades. In fact, underperformers, as well as good performers, can benefit from this study by exerting extra effort

* Corresponding author.

E-mail addresses: alaa7alaf@gmail.com (Alaa Khalaf Hamoud), alishashim2009@gmail.com (Ali Salah Hashim), umzainali@gmail.com (Wid Aqeel Awadh).

in conducting quality projects and research through obtaining help and attention from their instructors.

The paper organized as follow: in section 2, the recent related works to academic performance are presented; section 3 explains some proposed decision tree algorithms that used in the model (J48, REPTree and Random Tree). Section 4 lists and explains the proposed model while the final section explains the conclusion points that concluded from the model.

II. Related Work

Natek Srečko, and Moti Zwilling [8] focused on data mining for small data sets of student and answered the research questions by comparing two data mining tools. The best model was chosen based on results evaluation. They compared their prediction of student performance for the academic year 2012/2013 using the attribute: ‘‘Final grade’’, with the actual Final grade of the same students. After that, the authors chose the best parameters for filling the data set from sample data mining techniques. The conclusions of the work were very promising and encouraging for HEIs in terms of integrating data mining tools as very important part of their knowledge management systems in higher education.

Alaa Khalaf Hamoud [9] constructed a model on the basis of an experimental data set on Portuguese students from two courses (Mathematics; 395 instances) and Portuguese (Portuguese language course; 659 instances). Paulo Cortez and Alice Silva, from the University of Minho, Portugal, collected and analyzed the data. Three decision tree algorithms (J48, RepTree, and Hoeffding Tree (VFDT)) were applied to this work. The results confirmed that the J48 algorithm was most suitable for classifying and predicting the willingness of students to complete higher education and success in their courses.

Mishra Tripti, Dharminder Kumar, and Sangeeta Gupta [10] used different classification algorithms to construct a prediction model of performance on the basis of the academic and social integration as well as various emotional students’ skills which other studies, thus far, have not considered. The J48 (implementation of C4.5) and Random Tree algorithms have been applied to students’ records of colleges affiliated with Guru Gobind Singh Indraprastha University to predict third semester performance. Clearly, Random Tree was more accurate compared with J48 in performance prediction.

Muluken Alemu Yehuala [11] built models and tested them using a sample data set of 11,873 regular undergraduate students. WEKA application software was used for the building model and analyzing it. The results offered supportive and constructive recommendations to academic decision makers in universities to enhance supporting and making decisions. Furthermore, the results will aid in rebuilding and modification process of curriculum structure to improve student academic performance. Based on previous research results and findings, students are able to decide about their preferred field of study before they register and enroll. These findings verify that the Ethiopian Higher Education Entrance Certificate Examination (EHEECE) results, number of courses per semester, gender and number of students in the class, as well as the majority of study are the consedrable factors that affect student performance. Accordingly, based on the results, the level of student success can be controlled. Therefore, preventing educational institutions from incurring serious financial strains is possible.

III. Decision Tree

A decision tree is one of data mining classification technique which used to build a top down tree like model based on the attributes of a given data set. Moreover, a decision tree is a predictive modeling technique that used to predicate, classify, or categorize given data objects on the basis of a previously generated model using a training data set with the same features (attributes) [12].

The structure of the generated tree includes a root node as well as internal and leaf (terminal) nodes. The first or root node is the first top node which has no incoming nodes and one or more outgoing edges. An internal or middle node has one incoming edge and one or more outgoing edges, where each internal node denotes a test on an attribute and each edge represents a result of the test. Finally, the leaf node represents the final suggested (predicted attribute (label) of a data object [13].

The decision tree classification technique is performed in two stages [14]: tree building and pruning. Tree building stage follows top to down manner. During this stage, the tree is recursively partitioned until the data items belong to the same class label. This stage is very tedious and consumes a lot of computation processes since the training data set is repeatedly reprocessed. Tree pruning stage is done in bottom up manner. This practice is performed to improve the prediction and classification accuracy results of the algorithm by minimizing over fitting (noise or considerable detail in the training data set). Over fitting may result misclassification error in the decision tree algorithm. Decision tree offers many advantages to data mining, some of which are the following:

- Decision tree can be clearly understood by the analyst and any end user.
- Decision tree can handle different kinds of input data, namely, nominal, numeric, and textual.
- Decision tree can process erroneous data sets or missing or uncompleted values.
- Decision tree has a high level of performance with a minimal amount of effort and time.
- Decision tree can working in data mining applications over a variety of platforms [15].

In this study, three decision tree algorithms were used on collective student data, namely, J48, Random Tree, and RepTree.

A. J48

J48 is used for both of classification and prediction operations. For classification, J48 was chosen (on the basis of the C4.5 algorithm from machine learning) given that this algorithm is one of the most used in Weka tools that offers stability among the precision, speed, and interpretability of results. In addition this algorithm classifies data in the form of a decision tree with which we can easily identify weak students. Classification learning as a part of Educational Data Mining (EDM) was also implemented to predict the performance of students [16].

B. REPTree

The reduced-error pruning (REP) tree as a decision tree-learning algorithm can be considered a fast classifier based on the principles of computing information gain with entropy and minimizing the error arising from variance [17]. REPTree generates many of trees and applies regression tree logic in changing the iterations. Subsequently, the algorithm selects the best one from all spawned trees. Based on variance and result information, the algorithm builds a regression decision tree. Further, this algorithm prunes the trees based on using back-fitting method and reduced-error pruning. As in C4.5, this algorithm can also works with missing or uncompleting values by splitting the identical instances into pieces [18].

C. Random Tree

The random tree algorithm selects a test on the basis of a specific number of random features at each node without pruning. Commonly, Random trees refer to those randomly built and have nothing to do with machine learning. The merit of building a random tree is the efficiency of training and minimal memory requirements. To create a random decision tree, Random Tree algorithm uses only one pass over the data [17, 19].

IV. Model

Fig. 1 shows the construction process of the model, which passes through four steps. First, building a questionnaire is considered a part of the data preprocessing step. This step includes preparing the data set for evaluation, cleaning data, converting data range, and creating a derived column (Failed) based on another column (Number of Failed Courses). The derived column (Failed) is created on the basis of a simple equation: If (Number of Failed Courses > 0), then Failed = 'F'. Else Failed = 'P'; F = abbreviation of Failed, P = abbreviation of Passed. The column Failed is considered the goal class of the model.

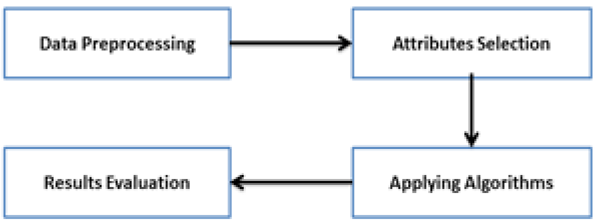


Fig. 1. Model Construction Diagram.

A. Data Preprocessing

1) Data Collection

The questionnaires that were built on Google Forms and an open source application (LimeSurvey) were used conduct a survey on students from the College of Computer Science and Information Technology (CSIT), University of Basrah. The first questionnaire (built Lime Survey) was used to locally collect data from CSIT, whereas Google Forms was used to collect data over the Internet. A total of 161 questionnaires were completed after combining the CSV files from Google Forms and the LimeSurvey questionnaire. The research sample (161 answers) represents an acceptable sample CSIT population with a 10% margin of error for the result [20].

Table I shows the description of all answers to the questionnaire. Question possible value was shortened and converted from a nominal to a numeric type for ease of use and understanding. Response values of questions (Q18-Q60) are of the form {1; 2; 3; 4; 5} where 1; 2; 3; 4; 5 represents the answers "Strongly Disagree'', "Disagree", "Neutral", "Agree", "Strongly Agree" respectively.

The initial step in data preprocessing is preparing data by removing rows with empty values and converting data for evaluation as well as processing. A total of 8 rows with empty values are under one or more columns. After removing these rows, we obtain a total 151 answers. Subsequently, row values are converted for data processing in the Weka 3.8 tool with its built-in classifiers.

TABLE I. Questionnaire Description

Class	Variable	Description	Possible Value
Demographic Data	Q1	Department	IS, CS
	Q2	Age	18,19,20, >20
	Q3	Stage,	1,2,3,4
	Q4	Gender	Female Male
	Q5	Address	In Basra Out of Basra
Social Information	Q6	You're Status?	Married Single
	Q7	Are you working?	YES, NO
	Q8	Are you live with your parents?	YES,NO
	Q9	Are you parent a live?	YES, NO
	Q10	Are your father working?	YES, NO
	Q11	Are your mother working?	YES,NO
Academic Information	Q12	No. of Fail Courses	0,1,2, >2
	Q13	No. of Absence Days	0,1-5, 5-10, >10
	Q14	No. of Credits	<12 , 13-17, >17
	Q15	GPA	<60,61-70,71-80, > 80
	Q16	No. of Complete Credits ,	<36 , 36-71, 72-107, >107
	Q17	No. of Years of Study	1,2,3,> 3
Study Skill	Q18	I can pick out the important point in the material that I read easily.	1,2,3,4,5
	Q19	I write notes and later I use them for preparing for tests.	1,2,3,4,5
	Q20	I schedule my time carefully for study.	1,2,3,4,5
	Q21	I am cool, very calm, and collected during the exam process.	1,2,3,4,5
Motivation	Q22	Getting a low test score does not make me feel that I am a failure.	1,2,3,4,5
	Q23	I made my own decision to go to get a college education	
	Q24	I study even when less important things distract me.	1,2,3,4,5
	Q25	I am excited about the courses I take.	1,2,3,4,5
	Q26	I have a very clear idea of the benefits which I expect to get from my education.	1,2,3,4,5
Personal Relationship	Q27	I develop and maintain supportive relationships with others.	1,2,3,4,5
	Q28	I control my upsets or anger without blaming others.	1,2,3,4,5
	Q29	I have ability for making friendships and create worthy relationships in any new place.	1,2,3,4,5
	Q30	I am open with people whom I don't especially like in order to learn from them.	1,2,3,4,5

Health	Q31	I have enough energy for studying and still fully enjoying other areas of my life.	1,2,3,4,5
	Q32	I regularly exercise.	1,2,3,4,5
	Q33	My emotional health supports my ability to learn.	1,2,3,4,5
	Q34	I control the alcohol and drugs I put into my body.	1,2,3,4,5
Time Management	Q35	I can use a computer effectively to enhance my success in school.	1,2,3,4,5
	Q36	I plan for each single day of the week.	1,2,3,4,5
	Q37	I plan my day by deciding what is important first.	1,2,3,4,5
	Q38	I don't study hard on the night before exam because of good planning.	1,2,3,4,5
	Q39	I plan times regularly for doing fun things.	1,2,3,4,5
Money Management	Q40	I plan for spending or saving my money, and I control of my personal finances.	1,2,3,4,5
	Q41	I have a clear view of the financial resources available to me to support my education.	1,2,3,4,5
	Q42	I can make money.	1,2,3,4,5
	Q43	My education supports my long range financial goals.	1,2,3,4,5
	Q44	I have an emergency savings plan which I follow as needed.	1,2,3,4,5
Personal Purpose	Q45	I see education as something I will be do through all of my life.	1,2,3,4,5
	Q46	I know that I am responsible for my own education.	1,2,3,4,5
	Q47	I responsible of the quality of my life.	1,2,3,4,5
	Q48	Ready to face challenges even where I'm not sure how to meet them.	1,2,3,4,5
	Q49	I have a clear idea about the best career plans for me	1,2,3,4,5
Career Planning	Q50	I know a lot about myself in relation to what kind of profession would be best for me.	1,2,3,4,5
	Q51	I have worked in the profession areas that interest me, or at least seen for myself what they require.	1,2,3,4,5
	Q52	I know the potential qualities that my employers will expect of me.	1,2,3,4,5
	Q53	I know education choices and schedules that I must follow in order to reach my career goal.	1,2,3,4,5
Resource Needs	Q54	I have enough money to complete my education and still meet all of my other financial commitments.	1,2,3,4,5
	Q55	My relationships with friends, family, and others support my educational goals.	1,2,3,4,5
	Q56	The available resources to me are enough to meet my needs.	1,2,3,4,5
Self_Esteem	Q57	I do not hesitate to ask for help when needed from family members, friends, or other appropriate individuals.	1,2,3,4,5
	Q58	I actively try to improve myself for my own benefit and for others as well.	1,2,3,4,5
	Q59	I am able to manage time and other life demands effectively.	1,2,3,4,5
	Q60	I am sure that I have the ability and skills to do college-level work successfully.	1,2,3,4,5

2) Reliability

Reliability is a feature of data set used for characterizing the overall uniformity of a measure. A measure is said to have a high level of reliability if it gives similar results under consistent conditions. For example, the measurements of people's height and weight are generally highly reliable [21]. In statistics, the coefficient alpha method is the most frequently used for the measuring of internal uniformity that is used as an indicator of reliability for the dependent variable of the study. Based on [22], it can be said that the value 0.7 indicates satisfying internal consistency in reliability. Table II shows that the coefficient alpha is 0.85 for the scaled variables that contain 60 items and 161 respondents.

TABLE II. Questionnaire Reliability

Cronbach's alpha	No.of items	No. of respondents	% of respondent
0.85	60	161	100%

B. Attribute Selection

Finding the most correlated attributes (questions) to the final class (Failed) and how much they affect the final class is important. Significantly, this stage shows the average correlation of the attributes to the final class. In turn, this average will help us find the questions with low correlation and remove them to improve the accuracy of the results. Questions with high correlation can be considered as recommendation points for students and academic staff.

In this step, the filter CorrelationAttributeEval is used to evaluate the correlation between the class and other attributes. This step is crucial because we want to find the most closely related attributes that affect the class and ignore the less-related attributes from the model. The attribute evaluator algorithm (CorrelationAttributeEval) evaluates the worth of an attribute by measuring the correlation (Pearson's) between the attribute and the class.

TABLE III. Correlation Average of Questions

Seq.	Question	Average	Seq.	Question	Average
1	Q14	0.436	31	Q52	0.101
2	Q25	0.312	32	Q49	0.09
3	Q60	0.234	33	Q19	0.094
4	Q15	0.224	34	Q11	0.086
5	Q46	0.227	35	Q26	0.088
6	Q17	0.225	36	Q42	0.089
7	Q51	0.21	37	Q48	0.077
8	Q47	0.209	38	Q24	0.074
9	Q10	0.174	39	Q20	0.074
10	Q23	0.168	40	Q41	0.065
11	Q27	0.168	41	Q22	0.068
12	Q45	0.173	42	Q55	0.064
13	Q29	0.162	43	Q36	0.065
14	Q57	0.158	44	Q4	0.062
15	Q13	0.159	45	Q5	0.061
16	Q18	0.152	46	Q31	0.054
17	Q35	0.147	47	Q6	0.052
18	Q40	0.145	48	Q3	0.042
19	Q58	0.142	49	Q56	0.04
20	Q43	0.134	50	Q44	0.039
21	Q16	0.135	51	Q32	0.039
22	Q38	0.131	52	Q1	0.043
23	Q8	0.133	53	Q2	0.028
24	Q34	0.131	54	Q54	0.025
25	Q37	0.123	55	Q12	0.023
26	Q59	0.12	56	Q39	0.026
27	Q53	0.113	57	Q33	0.023
28	Q9	0.116	58	Q28	0.023
29	Q30	0.116	59	Q7	0.022
30	Q50	0.106	60	Q21	0.018

Table III shows the average correlation between questions/attributes and the final class with the mode of evaluation (10-fold cross validation) to ensure a high level of accuracy. The questions were arranged in an ascending order on the basis of the average rate of correlation to the final class. Questions with high average rates are most correlated to the final class. Later, the last twenty questions will be removed to increase the accuracy of the result.

C. Applying Algorithms

The Weka tool provides built-in algorithms that help in the application of different classifiers and obtain results in an easy and flexible process. Three algorithms will be used in this stage (J48, Random Tree, and RepTree) before (less correlated questions) and after the removal of attributes. The attribute removing process is profoundly helpful in discovering the effect of these attributes on performance and how they can increase or decrease accuracy.

Table IV shows details on performance before the removal of attributes from the data and further shows the three algorithms (J48, Random Tree and RepTree) according to their attributes, namely, True Positive (TP), False Positive (FP) rates, Precision, and Recall.

TABLE IV. Algorithm Performance Before the Removal of Attributes

Classifier		TP Rate	FP Rate	Precision	Recall
Without Attribute Filter	J48	0.529	0.485	0.539	0.529
	Random Tree	0.608	0.442	0.601	0.608
	Rep Tree	0.621	0.448	0.609	0.621
With Attribute Filter	J48	0.516	0.499	0.526	0.516
	RandomTree	0.641	0.373	0.646	0.641
	Rep Tree	0.634	0.440	0.623	0.634

The first row reveals the performance of the algorithms without an attribute filter (CorrelationAttributeEval). The second presents the performance after applying the attribute filter combined with the decision tree algorithm. A meta classifier (Attribute Selected Classifier) allows us to combine the classifier algorithm with the attribute evaluator and the search method to get highly accurate results.

The effectiveness of the attribute filter on the algorithms (Random Tree and RepTree) is evident. TP Rate, Precision, and Recall in both algorithms are increased, while FP Rate decreased after applying the attribute filter with algorithms (Random Tree and RepTree) while it decreased with the J48 algorithm.

TABLE V. Performance Details After the Removal of Attributes

Classifier		TP Rate	FP Rate	Precision	Recall
Without Attribute Filter	J48	0.634	0.409	0.629	0.634
	Random Tree	0.595	0.430	0.597	0.595
	Rep Tree	0.634	0.440	0.623	0.634
With Attribute Filter	J48	0.621	0.423	0.616	0.621
	Random Tree	0.614	0.412	0.615	0.614
	Rep Tree	0.601	0.488	0.583	0.601

D. Result Evaluation

Result evaluation is the final stage in the model construction process. Based on Tables IV and V, the accuracy of the J48 classifier after the removal of the less correlated attributes is apparently higher compared with that of the RepTree and Random tree classifiers. The TP rate attribute takes a value of 0.634, which is the highest value for this attribute, whereas Precision (refers to a positive predictive value) also gets the highest value at 0.629. Recall (refers to a TP rate) with a value of 0.634 and an FP rate with a value of 0.409 makes the RepTree classifier as the nominated algorithm for the model (see Fig. 2).

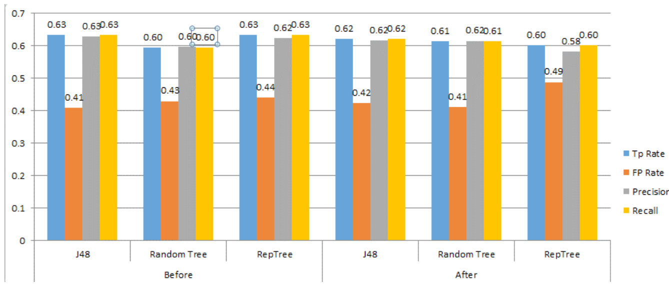


Fig. 2. Classifier Performance Chart.

Fig. 3 shows the result tree of the J48 algorithm, which is the best approach to show the most correlated factors/questions to the final class (failed). Each node in the tree is a question, and its branches are drawn on the basis of the answers. Each node can be considered a decision. The tree can be used also for predicting failure or success by answering questions in the nodes of the tree.

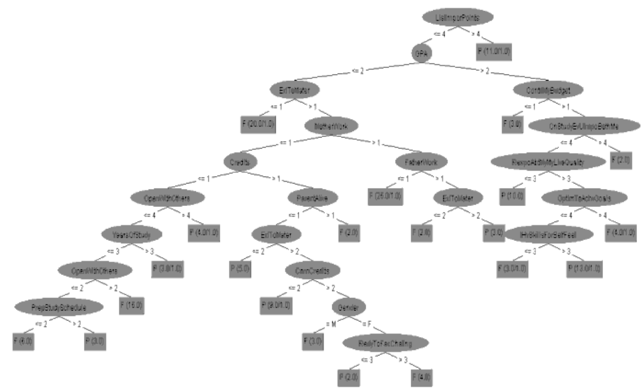


Fig. 3. J48 Classifier Tree.

V. Conclusion

This study aims to explore and test the process of applying the decision tree algorithms with questionnaire of students to seek the factors that affect student success/failure. Data mining algorithms and especially decision tree algorithms can be the best solution for predicting students' performance because they provide high accurate results and give road map for both academic stuff and students. Based on the results of the model, it can be said that a number of factors (attributes) can affect the accuracy of the result tree and overall student academic performance. Attributes such as Age, Work, Gender, Stage, and Status had less effect on student success, whereas GPA, Credits, List Important Notes, Father Work, and Fresh Food had the most significant effect on the final class. The attribute evaluator algorithms can be used to find closely related questions that negatively or positively affect the success of students. The questionnaire contains many unimportant questions that can be discovered by the data-mining algorithms. Large data set and number of attributes in the data set affects the accuracy of the decision tree. The model can be utilized by students and academic staff to decide which questions/answers will enhance academic performance and improve the success of institutions.

References

[1] A. M. Abaidullah, N. Ahmed, and E. Ali. "Identifying Hidden Patterns in Students" Feedback through Cluster Analysis." International Journal of Computer Theory and Engineering 7.1 (2015): pp. 16-20.

[2] M. Goyal and R. Vohra. "Applications of Data Mining in Higher Education." International journal of computer science 9.2 (2012): pp. 113-120.

[3] M. A. Yehuala. "Application of Data Mining Techniques for Student Success and Failure Prediction (The Case of Debre_Markos University)." International Journal of Scientific & Technology Research 4.4 (2015): pp. 91-94.

[4] K. M. Osei Bryson. "Towards Supporting Expert Evaluation of Clustering Results Using a Data Mining Process Model." Information Sciences 180.3 (2010): pp. 414-431.

[5] R. Bekele and M. Wolfgang. "A Bayesian Approach To Predict Performance of a Student (bapps): A Case With Ethiopian Students." Algorithms 22.23 (2005): 24.

[6] K. D. Kolo, S. A. Adepoju, and J. K. Alhassan. "A Decision Tree Approach for Predicting Students Academic Performance." International Journal of Education and Management Engineering (IJEME) 5.5 (2015): pp. 12-19.

[7] C. Romero, V. Sbastián, and G. Enrique. "Data mining in Course Management Systems: Moodle Case Study and Tutorial." Computers & Education 51.1 (2008): pp. 368-384.

[8] S. Natek and M. Zwilling. "Student Data Mining Solution Knowledge Management System Related To Higher Education Institutions." Expert systems with applications 41.14 (2014): 6400-6407.

[9] A. Kh. Hamoud. "Selection of Best Decision Tree Algorithm for Prediction and Classification of Students' Action". American International Journal of Research in Science, Technology, Engineering & Mathematics, 16(1), September- November, 2016, pp. 26-32. Available at SSRN: https://ssrn.com/abstract=3027086.

[10] T. Mishra, K. Dharminder and G. Sangeeta. "Mining Students' Data For Performance Prediction." Fourth International Conference on Advanced Computing and Communication Technologies. 2014.

[11] M. A. Yehuala. "Application of Data Mining Techniques for Student Success and Failure Prediction (The Case of Debre_Markos University)." International Journal of Scientific & Technology Research 4.4 (2015): pp. 91-94.

[12] Bhargava, G. Sharma, R. Bhargava and M. Mathuria, "Decision Tree Analysis on J48 Algorithm for Data Mining", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 6, June 2013.

[13] Amjad Abu Saa, "Educational Data Mining & Students' Performance Prediction." International Journal of Advanced Computer Science and Applications 7.5 (2016): pp.212-220.

[14] M. N. Anyanwu, and S. G. Shiva. "Comparative Analysis of Serial Decision Tree Classification Algorithms." International Journal of Computer Science and Security 3.3 (2009): pp. 230-240.

[15] L. Rokach and O. Maimon. "Data mining with decision trees: theory and applications." World scientific, 2014.

[16] K. N.Shah, et al. "Study of Data Mining in Higher Education-A Review." (IJCSIT) International Journal of Computer Science and Information Technologies 6.1 (2015): pp. 455-458.

[17] I. H. Witten, E. Frank, and M.A. Hall. "Data Mining: Practical Machine Learning Tools and Techniques." Morgan Kaufmann, 2016.

[18] S. K. Jayanthi and S. Sasikala. "REPTree Classifier For Identifying Link Spam in Web Search Engines." IJSC 3.2 (2013): pp. 498-505.

[19] G. Jagannathan, K. Pillaipakkamnatt, and R. N. Wright. "A practical Differentially Private Random Decision Tree Classifier." Data Mining Workshops, 2009. ICDMW'09. IEEE International Conference on. IEEE, 2009.

[20] G. D. Israel. "Determining Sample Size. "University of Florida IFAS extension." (2009).

[21] B. Carson. "The Transformative Power of Action Learning." Chief Learning Officer. Retrieved 2017, from: http://wial.org/executive-board/bea-carson-executive-board/.

[22] U. Sekaran, and R. Bougie. "Research Methods For Business: A Skill Building Approach." John Wiley & Sons, 2016.

Alaa Khalaf Hamoud

Alaa Khalaf Hamoud is a lecturer in Computer Information Systems, University of Basrah, Iraq. He received BSc degree from Computer Science Department, University of Basrh in 2008 with first ranking college student. He also received his MSc degree from the same department with first ranking department student. He participated in (seven months) IT administration course in TU berlin-Germany. His scientific interests are data mining, data warehousing.

Ali Salah

Ali Salah received M.sc in Information Technology from Tenaga University, (Malaysia) and B.sc in Computer Science from Basra University (Iraq). His current research interests includes cloud computing and data mining.

Wid Akeel

Wid Akeel received M.sc in Information Security (2012) from Basra University, (Iraq) and B.sc in Computer Science (2006) from Basra University (Iraq). Her current research interests includes information security, Steganography, image processing and data mining.

An Integrated Learning Analytics Approach for Virtual Vocational Training Centers

Peter de Lange*, Alexander Tobias Neumann, Petru Nicolaescu, Ralf Klamma

RWTH Aachen University, Aachen (Germany)

Received 3 November 2017 | Accepted 22 January 2018 | Published 15 February 2018

Abstract

Virtual training centers are hosted solutions for the implementation of training courses in the form of e.g. Webinars. Many existing centers neglect the informal and social dimension of vocational training as well as the legitimate business interests of training providers and companies sending their employees. In this paper, we present the virtual training center platform V3C that blends formal, certified virtual training courses with self-regulated and social learning in synchronous and asynchronous learning phases. We have developed an integrated learning analytics approach to collect, store, analyze and visualize data for different purposes like certification, interventions and gradual improvement of the platform. The results given here demonstrate the ability of the platform to deliver data for key performance indicators like learning outcomes and drop-out rates as well as the interplay between synchronous and asynchronous learning phases on a very large scale. Since the platform implementation is open source, results can be easily transferred and exploited in many contexts.

Keywords

Virtual Vocational Training, Learning Analytics, Digital Blended Learning.

DOI: 10.9781/ijimai.2018.02.006

I. Introduction

The digital transformation is affecting vocational training as any other business these days. A virtual vocational training center is a hosted solution for the implementation of training courses in the form of e.g. Webinars. There are many undeniable advantages coming with the utilization of virtual training centers, both for training providers and trainees. Training providers can save the rental for training facilities including the procurement of furniture, media technology and training materials. Furthermore, they can expand their business to other regions within their country or even across borders, e.g. in the European Union. Scaling their business is important for many training providers, in case their business is threatened by economic up- and downturns or by changing demands in job descriptions. They can enter other markets, e.g. translate electronic training materials, adjust materials to different job descriptions or can cooperate with other training providers in larger virtual training centers at low front-off costs. Trainees can save travel time and costs and take parts in training programs offered outside their living areas. However, the virtualization of training has also downsides, most important the loss of social interaction with other learners and the informal exchange of information before, during and after the training. So, the goal of a virtual training center is not only the formal implementation of training courses with the necessary training, testing and certification procedures but also a re-establishment of informal and social learning opportunities.

In this paper, we present the "Virtus Virtual VET Center" (V3C) concept and its implementation on a Web-based platform. It offers Webinar style training courses with the facilitation of trainers in a synchronous manner as well as self-regulated, asynchronous learning spaces, where learners can learn in a self-paced, socially-aware way together with other learners or alone. To evaluate the impact of the different learning modes on the learning outcomes we developed and applied several learning analytics methods, utilized and combined in a way to get insights into both synchronous and asynchronous learning modes. The platform itself uses different data collection methods based on the different interaction possibilities and also different data analysis and visualization procedures, combining them to an integrated holistic learning analytics approach for vocational training centers. The data and its visualization are accessible both for trainers and trainees, so that they can reflect on their learning progress.

The long-term goal of our research is to blend synchronous Webinar or MOOC style learning with informal social learning in vocational training in a flexible Web-based platform together with the fulfillment of business goals of vocational training centers as well as learning goals of small and medium sized companies. This blending is supported by using our integrated learning analytics approach, which makes it possible to track the learners' progress throughout the different learning phases and enables both learners and tutors to see differences and similarities in the learning outcomes.

The remainder of this contribution is structured as follows. Section 2 describes the concept and realization of the V3C platform. We then start Section 3 with stating our research questions, describing our data sources and the method of analysis we followed in our evaluation. The section is continued by presenting our results and their interpretation. Finally, Section 4 describes work related to this contribution and Section 5 concludes our paper and provides an outlook on future work.

* Corresponding author.

E-mail address: lange@dbis.rwth-aachen.de

II. A State-of-the-Art Virtual Vocational Training Center

A state-of-the-art virtual vocational training center has to combine a Learning Management Systems (LMS) with a Personal Learning Environments (PLE) [1]. The LMS has to be designed to be usable by all stakeholders responsible for course (content) creation, thus it has to have a very low technical entry barrier. The PLE should allow for both synchronous, class-style learning and asynchronous, self-regulated learning at the same time to cope with the requirements that vocational training brings. While classical teaching via Webinars is needed in this domain, vocational training needs to take additionally into account asynchronous learning. Since many students are working full time at different companies, they are connected to each other more in the form of a Community of Practice (CoP) [2]. Often, it is impossible for them to participate in each teaching session. To develop a platform that supports both learning modes, the platform needs to have a live video chat functionality as well as static content that is available all the time, like slidesets, videos and online assessment tools that provide feedback to both learners and tutors about the students' progress. Additionally, the system should track the learning progress of the students in form of data collections in the background. This data and its aggregations, clustering and evaluation allows for Learning Analytics (LA [3]) that can provide valuable insights for both students and tutors to either improve their learning or teaching (material) accordingly.

In the following, we present the realization of the "Virtus Virtual VET Center" (V3C), developed in the scope of the European Erasmus+ project VIRTUS. We first presented this platform in [4] and here we only want to give the reader a short overview of the platform's capabilities. We realized our platform as a hyper learning environment consisting of a LMS and a PLE.

Regarding the technical aspects, we used well-established Web development languages and protocols, such as PHP, JavaScript, Java RESTful microservices, HTML5, XMPP and WebRTC. The LMS allows for a modeling-like drag-and-drop design of course rooms (see Fig. 1). This eases the creation of courses, especially for non-technical learning designers, since it shows the learning room already in a "What-You-See-Is-What-You-Get" (WYSIWYG) fashion. Each course is divided first into several modules, which again are divided into multiple learning units.

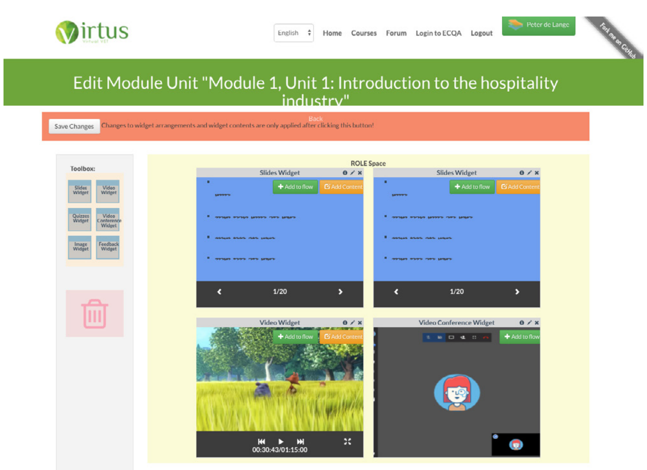


Fig. 1. The modeling view of the V3C platform.

The courses generated by the LMS of our platform are represented as "learning spaces", realized using the "Responsive Open Learning Environment" (ROLE) platform [5]. Fig. 2 shows such a learning space of a module of the course "Social Entrepreneurship". Each space represents a designed module, with its individual learning units being represented as "learning activities" of the space. This realizes a separation between the individual units which are part of a course, enables activity and progress tracking for individual units and allows for assessment via quizzes of the respective units' learning outcomes. V3C users can autonomously join spaces via the respective course unit in their LMS. Each course unit may consist of video chat, slide presentations, various multimedia content such as audio recordings, videos and images, and self-assessment quizzes. All realized software is Web-based and open source.

Fig. 2. A course room of the V3C platform.

Tutors and other learners can intervene in the learning process at any point via video or text chat that is available for each course. Since our target group for both learning designers and learners consists of people from Italy, Austria, Greece and Spain, the V3C is developed with extensive translation functionality, providing opportunities to offer and translate learning units into different languages. Finally, our platform is linked to the European Certification and Qualification Association (ECQA), which then conducts and assesses the final certified exams. For data protection, we use the OpenID Connect standard to feature a unified login for both the LMS and the PLE as well as for the certification process at the ECAQ.

The LMS of the V3C platform has a per-module analytics section where tutors can see usage statistics as well as progress reports of the course participants. While learners use the PLE, usage information is logged into a MariaDB database. We track user interactions as clicks on different HTML elements. We also implemented a routine that measures the time spent on the respective module unit as duration. This routine sends every minute a request that updates the time. All visualized information shown on the platform are real-time analytics. The analytics section is split up into three different views (participants, feedback and activity). The first view contains information about individual participant results in the learning module selected. It shows the total number of participants of the module and lists them. Since during the creation progress of a module and its units, the tutor has specified the ECVET points that are later granted to students passing the final test of the course, we weight these points with a factor, resulting in a minimum duration the user has to interact with the platform to have a "full completion rate" of the course. Another information to be observed in this view is the assessment of the module. At the end of each module, the user has to complete a quiz. By clicking on a participant, the tutor can inspect the monitored data, split up into the module's units. It now also shows the previously described data of each unit, as well as how much of the quiz was completed and how many correct answers were given. In the second view, tutors can see feedback of the learners. The feedback widget is optional for the learners and serves as a means for the learners, especially when using the platform in an asynchronous, self-regulated learning mode, e.g. to get in contact with the course designer to mention problems with the content. Of course, it can also be used to submit answers to given tasks by the tutor. The evaluation then has to be done manually by the tutors. The last view of the analytics section displays the "activity graph" of the module. All interactions tracked by the system and the overall activity of the module is displayed here. These visual analytics support enables tutors to track the activity of their modules over time, being able to see at what times students engage themselves most in the platform, which courses are more frequently visited etc. The graph is auto-scaled and shows the whole activity since the beginning of the module, but it also offers the option to select a desired time period. We show a screenshot of these graphs later in the evaluation section in Fig. 3.

III. Monitoring a Personal Learning Environment

In the following we describe the evaluation of the usage of the V3C platform. Part of what we present here, the questionnaire results, is based on the technical report created by the VIRTUS project consortium in [6].

A. Hypotheses

In order to evaluate our collected data and gain insightful conclusion regarding the usefulness of our approach, we formulated two research questions we wanted to answer in our evaluation.

RQ1: How well does platform immanent monitoring of student behavior reflect the learning process and how well does the data collection work?

This question for one aims at evaluating the effectiveness of our implemented solution regarding its technical capabilities of dealing with learner-generated amounts of big data. Second, it aims at identifying its capability to successfully make predictions regarding future assessment results, based on the learners' activity, with a special focus on at-risk students.

RQ2: How does asynchronous, self-regulated learning perform, compared to synchronous, Webinar-driven learning modes, if applied in the same platform?

Here, we want to compare the two contrary learning modes to evaluate their differences in final assessment results, drop-out rates and activity / engagement with the platform. Since our clustering of students into synchronous and asynchronous "evaluation groups" is done at random, our question is not aiming at categorizing students into different "types of learners", but at evaluating, what effect on the learning outcomes can be seen when replacing mandatory, tutor-supported synchronous learning sessions with additional time for self-regulated learning. Since our platform allows the evaluation of the two groups in the same learning rooms, we can achieve comparable results.

B. Data Sources

The data used in this evaluation was gathered from the platform immanent learning analytics mechanism, additionally complemented with the use of a questionnaire that addressed the general impressions participants had while using the platform. Our evaluation participants were either already workers in the tourism sector or planned to enter this sector. In fact, the courses offered by the platform concerned the two domains of (*Social Entrepreneurship* and *Tourism and Hospitality*). We had a total of 114 learners from four different countries, namely Austria, Italy, Greece and Spain. The evaluation period span two months. The synchronous learning phase was performed in the *Tourism and Hospitality* course and covered 18 units in four modules. Every module consisted of at least two units, with a maximum of up to five units. We were able to recruit 15 learners for the synchronous and 72 for the asynchronous evaluation. Each of the four vocational training Webinars spanned two hours and covered one module of the course, resulting in a four-day streak of consecutive synchronous evaluation sessions. The tutor facilitating the synchronous learning session spoke English and used the platform-immanent video conference widget. Starting with the beginning of the first synchronous evaluation session, learners who participated in this phase had two weeks to complete the course. The asynchronous learning phase began simultaneously with the synchronous learning phase, but here learners had 60 days to complete the course. Table I shows the number of learners who participated in each module. Finally, we invited the participants to fill out a questionnaire. We received 48 submissions here, without clustering them into the two evaluation clusters.

TABLE I. Participants per Module

Module	Synchronous Participants	Asynchronous Participants
SE1	0	42
SE2	4	50
SE3	0	37
SE4	0	37
SE5	3	34
TH1	15	72
TH2	14	58
TH3	15	55
TH4	14	54
TH5	9	51

(SE = Social Entrepreneurship, TH = Tourism & Hospitality)

C. Methods of Analysis

For collecting and public provision of our learning analytics, we used the MobSOS monitoring concept of our research group's Community Information System (CIS) platform [7]. The MobSOS Query Visualizer (MobSOS QV) was used to embed multiple measures together in a dashboard-like fashion, which we provide provided publicly on the platform. The stored queries for those visualizations are parameterized and can be used for every module and unit. For our analysis, we used the user activity and the scored result of the user. Currently, an activity is defined as any click interaction of the learner with the PLE, e.g. click on any button, watch a video or answering a quiz.

During the synchronous evaluation sessions, participants were allowed to use the whole platform. Although these sessions were conducted in English, participants also at this time were able to use the platform in their desired language. After each session, the tutor manually created an attendance list, which marked the attendants as part of the synchronous evaluation group, allowing us to cluster the participants into an asynchronous and a synchronous cluster for evaluation.

Both asynchronous and synchronous learners used the same learning rooms simultaneously. A module was marked as completed when the participant completed the quiz at the end of it. We only considered learners who have participated in each module of the course for the comparison of the scored result and the activity in the PLE. The scored result of a learner is the percentage of the correct questions of all modules. The activity is the percentage of all activities done by a user in the whole course compared to the activities done by the learner who has done the most activities in this course. We base the calculation of the drop-out rate by only taking those participants who finished all previous modules of a course, although our platform offers the possibility to only attend single modules of a course without the need to attend a previous module.

For further data which could not be derived by platform immanent monitoring, we follow the "observe, where possible; only survey, where inevitable" approach [7] by using additional five-level Likert items for yet missing, however relevant subjective factors. The questions of the survey were translated into the participants' native languages and were handed out in digital or printed format to the participants of both phases. Questions regarded both the platform as well as the teaching material, and we focused on those that deal with user satisfaction of the system since this is part of the success awareness. We had no means to link the questionnaire to individual persons, nor did we cluster them into asynchronous and synchronous participants' answers.

While using the platform, all participants were aware of the monitoring of their progress and behavior by our system. They had the possibility to always review their personal data we collected, such as quiz results and the time they spent in a learning space. Only the overview of all participants of a module was restricted to the tutor / learning designer as well as the visualization of the aggregated and anonymized learning room activity. This "Module Activity" view is also depicted in Fig. 3 of the next section.

D. Results

In this section, we present the results of our evaluation. It has to be noted that we only consider the course "Tourism and Hospitality" for analysis since it is the only course where we have both a synchronous and asynchronous evaluation, which enables the comparison between the two learning modes.

As already mentioned in the previous section, our platform features a visual analytics feature for tutors, teachers and learning designers to monitor the activity of their modules over time. Fig. 3 shows this view for all four modules of the course. It displays the absolute number of activities for each module over time. The timespan in this figure was chosen to match the sixty-day evaluation period. It can be seen that two peaks in every module exist, which correspond to the final assessment of both the synchronous evaluation phase (after 14 days) and the one of the asynchronous evaluation around six weeks later. As it can also be observed, there is a smaller, yet considerably high activity over the whole evaluation cycle, which is an indicator for continuous usage of the platform by the participants during our evaluation.

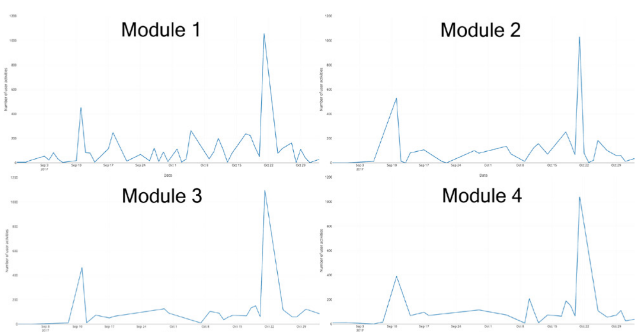


Fig. 3. Module activity over time.

Our next analytics compared the participants activity with their scored result in the assessment. Fig. 4 shows this comparison. Here, we used relative values for both activity and scored result to make them comparable, due to the different number of participants in both evaluation phases. We only considered those students that took part in the whole course. The Pearson correlation coefficients over all four modules were $p = 0.4593$ for the asynchronous phase and $p = 0.3589$ for the synchronous phase.

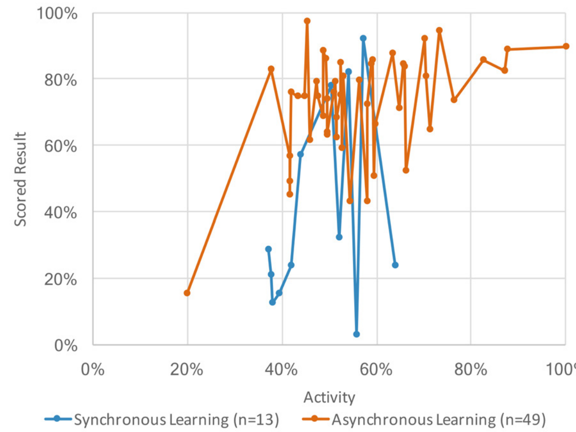


Fig. 4. Activities compared to scored result.

Fig. 5 gives an overview about the final aggregated results of the participants for both evaluations. Again, we use relative values here to make the results comparable and removed all prior drop-out participants from this statistic.

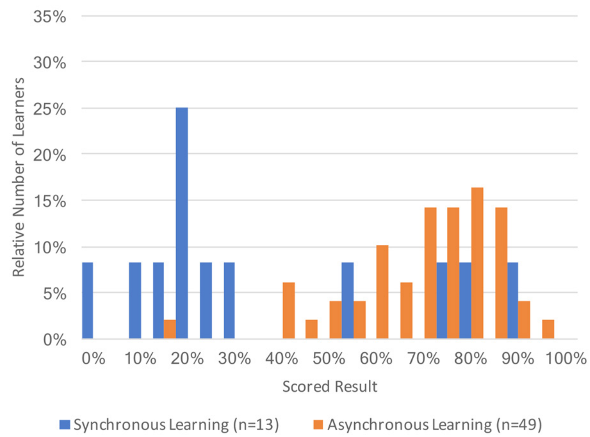


Fig. 5. Assessment results.

Our last analytics concerned the drop-out rate of both synchronous and asynchronous evaluation participants. Fig. 6 shows this statistic. The percentages are always relative to the base number of participants ($n = 15$ and $n = 72$). As it can be seen, here we have the complete participants of both evaluation groups of the evaluated course. When counting together the drop-outs of this figure, the number shrinks to the number of the prior three analytic measures ($n = 13$ and $n = 49$).

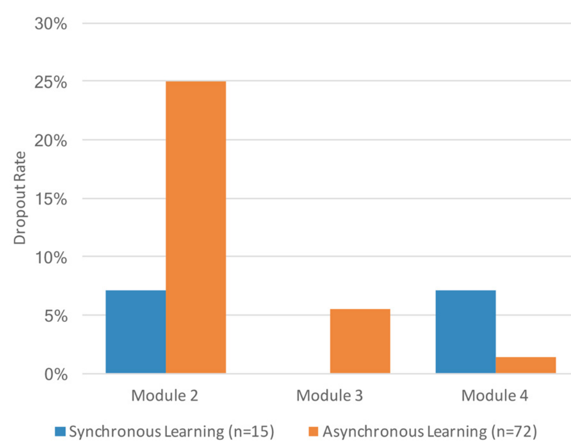


Fig. 6. Drop-out rate.

Finally, Fig. 7 shows the result of the questionnaire. We aggregated both synchronous and asynchronous participants together and also took into account learners from the other course ("Social Entrepreneurship") which we only used for asynchronous evaluation and thus have not taken into account for the prior analytics presented in this section.

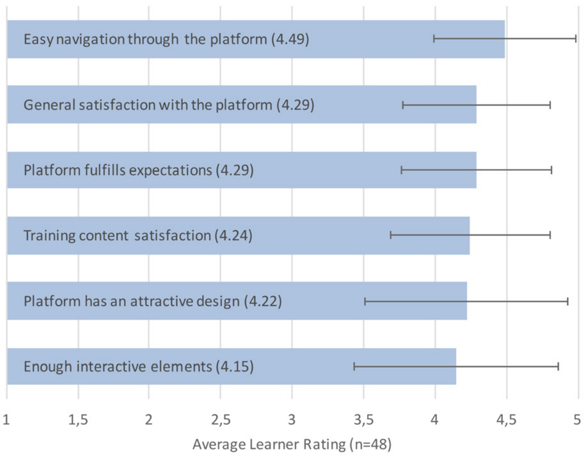


Fig. 7. Results of questionnaire.

E. Interpretation

With our evaluation, we wanted to answer two research questions. RQ1 regarded the effectiveness of our technical monitoring solution as well as its capability to make predictions regarding the learning outcome. For the first part, we can state that we were able to monitor all data that was produced by the learners in the evaluation and our platform was capable of handling both synchronous data streams of multiple people collaborative using the same learning room, as well as that it was robust against a long-term evaluation of two month for the asynchronous evaluation phase. This is also backed by Fig. 3, which shows a constant activity flow with the peaks clearly representing assessment deadline phases. For the second part of RQ1, we have to be a little more cautious with our interpretation. Fig. 4 shows a rather clear correlation between activity in a learning room and scored assessment results for participants of the asynchronous learning cluster. The Pearson correlation also confirms this with its value close to 0.5, which is considered as rather "correlation confirming" by most literature. For the asynchronous learning phase though, the relation is not that clear. The Pearson correlation again shows there is a slight correlation between both values, but the graph does not really allow for a clear prediction capability. Reasons for this might be the smaller subset of the asynchronous cluster. With a larger amount of participants here, the results might have been clearer. We also tried to evaluate the results on a per-module basis (instead of the four-module aggregated view), but these present similar results.

For RQ2, which asked how synchronous and asynchronous learning modes perform if applied in the same platform, first we can say that there is a clear trade-off between time provided to the participants and resulting assessment results. The cluster of asynchronous learners, which had about 75% more time to complete the assessments, scored about 33.38% better results than the cluster of synchronous learners. This can also be seen in Fig. 5. On the other hand, it has to be said that reducing the time needed to train people by two-third results also in a very high reduction of costs needed to be spend by employers to provide the training to their employees. Ultimately, it has to be decided by the facilitators of vocational training courses, if this trade-off between the time needed to train people and the expected results is worth of taking.

Regarding user acceptance of the platform, our questionnaire results (see also Fig. 7) indicate a high satisfaction of the users with the platform itself. Since all questions received similar high ratings, we will not discuss them in detail, but it has to be said that these ratings might also result from the domain we chose to evaluate our platform. People working or planning to work in the domain of tourism, especially on a service level, might not be used to the support their training by the means of multi-media in general. So, the availability of a platform that provides them with the possibility to perform their training courses online might have been perceived as very appealing, resulting in these high ratings of user acceptance in terms of learners.

F. Limiting Factor

Performing synchronous evaluations that span multiple days is a resource consuming task. Especially, when the participants have to be recruited from a particular vocational domain, with full-time job schedules and limited time to spend in training that is not directly financed by their employer. Therefore, we limited the comparison between the two learning modes presented in the previous section to the "Tourism & Hospitality" course, and covered only four out of five modules here. We decided to leave out the duration measure, which represents the time spent for a module, due to some technical uncertainties we had with the results they produced. This is definitely something we need to rethink and reimplemented for further evaluations. Currently, our assessment is limited to the usage of multiple choice quizzes. We are aware of the shortcomings of this approach regarding the capabilities of multiple choice questions to assess learning success as a whole. For the future, we are working on integrating the results of the ECVET certified final exams of the courses, done by the ECQA, into our dataset. For the moment, we are not able to consider those here due to some restrictions. During the analysis, we noticed that two learners of the asynchronous phase participated twice in a course, thus we removed both datasets of the participants from the evaluation data. One thing we did find out when perform our analytics was the following. It would have been interesting also to have a cluster of learners that had the same time to finish the course as the synchronous learners had (two weeks), but without providing them the opportunity to participate in the Webinar. Finally, we want to state that our user acceptance questionnaire only aimed at evaluating the platform from a learner's perspective. This results from the fact, that the platform was developed as part of the same project (VIRTUS), where also the VET trainers that provided the learning content via the platform were partners of. Thus, we had no external VET providers that we could have interviewed to find out about their opinions regarding the platform, although we deem this an area worth of further research.

IV. Related Work

Vocational training adapted to the digital, virtual domain not as fast as e-learning conducted in higher education at university level for example. Although its potential was first mentioned already at the end of the last century [8], its adoption is still in its infancy. The result of this is that studies performing learning analytics, to the best of our knowledge, consider either university/college or schools as the educational setting [9]. From a technology point of view, most learning analytics approaches use data generated by Massive Open Online Courses (MOOCs) created by open source course management system, with the most famous one being "Moodle" [10]. Here, using standardized data formats [11] to apply learning analytics on, like the Learning Record Store (LRS) of the xAPI standard [12], is an active research topic. Having a well-defined data (exchange) format allows for an easier handling of analyzing data, but it does not solve the question what data needs to be collected to understand past and predict future student behavior and react accordingly. Here, techniques like Educational Process Mining (EPM) can be used to retrospectively derive meaning of collected data logs. Bogarín et al. [13] have used this technique together with clustering to predict success rates of students taking a course of a Moodle course management system installation. Click-based analysis is also very often applied to learning analytics in virtual learning environments, like for example the authors of [14] used the GUHA (General Unary Hypothesis Automaton) method, a data mining technique, to predict student drop-out rates. In general, we can conclude that most studies consider the prediction and improvement of assessment results, activities and drop-out rates [15] as their main goal. While the first two are of equal importance in asynchronous vocational training, the drop-out rate problem is more bound to the MOOC context, since most vocational training still is done in a teaching-style manner, with often mandatory participation. Without self-regulated learning, drop-out rates in this setting are similar to classroom-like situation, which also corresponds to our study results. This especially lays the focus on the study of self-regulated learning with learning analytics to identify and analyze self-regulated learning strategies and how to improve them [16].

It appears that there is still a lack of research for virtual vocational training, especially in the domain of learning analytics. To the best of our knowledge, there exists no research that compares the effects of asynchronous and synchronous learning virtual vocation training and education.

V. Conclusion & Outlook

We presented an integrated learning analytics approach that allows to compare asynchronous and synchronous learning phases and draw conclusions for certification, interventions and gradual course improvement. It is embedded into a technical platform for the creation, implementation, deployment and performance of virtual training courses. We evaluated our contribution with two vocational training courses, one for social entrepreneurship and one for tourism & hospitality. Based on an English version, we translated both courses in four languages (German, Italian, Greek and Spanish) and conducted them with 114 learners from different European countries. Our main focus for this paper was the interplay of synchronous and asynchronous learning phases to demonstrate the ability of the platform to blend Webinar-style course units with phases of self-regulated learning on the platform. Additionally, we had an eye on important questions like for both virtual training centers offering courses as well as small and medium sized companies sending their employees for vocational training, e.g. drop-out rates.

The interpretation of our learning analytics results shows that the combination of Webinars and self-regulated learning saved additional resources beyond the pure availability of a virtual training center. This is the first evaluation of the learning platform and the evaluation of asynchronous learning phases is still ongoing at the time of the writing. Moreover, we have not differentiated yet the performance of learners in different language versions of the same course. That means that the possibilities for advanced learning analytics are not yet explored and we can expect more and better feedback from the platform and its analytic capabilities. What we can say for now is that the platform scales well both in the number of courses and learners, since the underlying ROLE framework has been in use for self-regulated learning since many years and one installation is capable of serving several thousands of learners. In further research, we will focus also on community learning analytics for understanding the social processes in virtual training centers better from an empirical perspective.

Acknowledgment

This research was funded by the European Commission under the European Union's Erasmus+ project "VIRTUS" (grant no. 562222).

References

[1] M. van Harmelen, "Personal Learning Environments", in Proceedings of the 6th IEEE International Conference on Advanced Learning Technologies, Washington, DC, USA: IEEE Computer Society, 2006, pp. 815–816.

[2] E. Wenger, Communities of Practice: Learning, Meaning, and Identity. Cambridge, UK: Cambridge University Press, 1998.

[3] R. Ferguson, "Learning analytics: drivers, developments and challenges", in International Journal of Technology Enhanced Learning, vol. 4, no. 5-6, pp. 304–317, 2012.

[4] P. de Lange, P. Nicolaescu, and R. Klamma, "VIRTUS Virtual VET Centre (V3C): A Learning Platform for Virtual Vocational Education and Training", in European Conference on Technology Enhanced Learning, 2017, pp. 500–503.

[5] D. Renzel, R. Klamma, M. Kravčík, and A. Nussbaumer, "Tracing Self-Regulated Learning in Responsive Open Learning Environments", in Advances in Web-Based Learning - ICWL 2015: 14th International Conference, Guangzhou, China, November 5-8, 2015, Proceedings, Berlin-Heidelberg: Springer, 2015, pp. 155–164.

[6] RWTH Aachen University, "R2.3: State of the Art Report on Virtual Organizations in Education and Training: Virtus Project Technical Report", 2016.

[7] D. Renzel, R. Klamma, and M. Jarke, "IS Success Awareness in Community-Oriented Design Science Research", in New Horizons in Design Science: Broadening the Research Agenda: 10th International Conference, DESRIST 2015, Dublin, Ireland, May 20-22, 2015, Proceedings, Switzerland: Springer International Publishing, 2015, pp. 413–420.

[8] R. C. Schank, Virtual learning: A revolutionary approach to building a highly skilled workforce. New York: ERIC; McGraw-Hill, 1997.

[9] N. Nistor, M. Derntl, and R. Klamma, "Learning Analytics: Trends and Issues of the Empirical Research of the Years 2011-2014", in Design for Teaching and Learning in a Networked World: 10th European Conference on Technology Enhanced Learning, EC-TEL 2015, Cham: Springer International Publishing, 2015, pp. 453–459.

[10] M. Dougiamas and P. Taylor, "Moodle: Using learning communities to create an open source course management system", in World Conference on Educational Multimedia, Hypermedia and Telecommunications: EDMEDIA'03, 2003.

[11] Á. Del Blanco, Á. Serrano, M. Freire, I. Martinez-Ortiz, and B. Fernández-Manjón, "E-Learning standards and learning analytics. Can data collection be improved by using standard data models?", in Global Engineering Education Conference (EDUCON), 2013 IEEE, 2013, pp. 1255–1261.

[12] J. M. Kevan and P. R. Ryan, "Experience API: Flexible, decentralized and activity-centric data collection", Technology, Knowledge and Learning, vol. 21, no. 1, pp. 143–149, 2016.

[13] A. Bogarín, C. Romero, R. Cerezo, and M. Sánchez-Santillán, "Clustering for improving educational process mining", in Proceedings of Learning Analytics and Knowledge Conference 2014, LAK '14, New York: ACM, 2014, pp. 11–15.

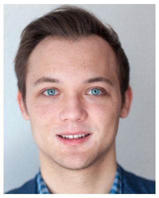
[14] A. Wolff, Z. Zdrahal, A. Nikolov, and M. Pantucek, "Improving retention: predicting at-risk students by analysing clicking behaviour in a virtual learning environment", in Proceedings of the Third International Conference on Learning Analytics and Knowledge: ACM, 2013, pp. 145–149.

[15] J. L. Santos, J. Klerkx, E. Duval, D. Gago, and L. Rodrìguez, "Success, activity and drop-outs in MOOCs an exploratory study on the UNED COMA courses", in Proceedings of Learning Analytics and Knowledge Conference 2014, LAK '14, New York: ACM, 2014, pp. 98–102.

[16] J. Maldonado-Mahauad, M. Pérez-Sanagustín, R. F. Kizilcec, N. Morales, and J. Munoz-Gama, "Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses", Computers in Human Behavior, vol. 8, pp.179-196, 2017.

Peter de Lange

Peter de Lange received his Master's degree in computer science in 2015 from RWTH Aachen University. Peter works as a research assistant and PhD student within the Advanced Community Information Systems group at the Information Systems and Databases chair, RWTH Aachen University, Germany since 2015. He is involved in several EU projects in Technology Enhanced Learning as well as research and teaching. Among his research interests are Web engineering, community information systems, learning analytics and peer to peer infrastructures.



Alexander Tobias Neumann

Alexander Tobias Neumann received his Bachelor's degree in computer science from the RWTH Aachen University in 2016. In his thesis he focused on human-computer interaction with emphasis on tangible user interfaces. Currently, he is working towards his Master's degree and focuses on technology-enhanced learning.

Petru Nicolaescu

Petru Nicolaescu works as a PhD student within the Advanced Community Information Systems group at the Information Systems and Databases chair, RWTH Aachen University, Germany. He is involved in major impact EU projects in technology enhanced learning, research and teaching since 2012. He received his engineering (B.Sc.) degree in computer science in 2009 from Politehnica University Bucharest and his Master degree in Software Systems Engineering in 2012 from RWTH Aachen University. Among his research interests are information systems and infrastructures, collaboration technologies, Web engineering and social network analysis.

Ralf Klamma

Ralf Klamma holds diploma, doctoral and habilitation degrees in computer science from RWTH Aachen University. He leads the research group "Advanced Community Information Systems" (ACIS) at the information systems chair, RWTH Aachen University. He is known for his work in major EU projects for Technology Enhanced Learning (PROLEARN, GALA, ROLE, Learning Layers, TELMAP, Tellnet, CUELC, SAGE, BOOST, VIRTUS and WEKIT). Ralf organized doctoral summer schools & conferences in Technology Enhanced Learning and Social Network Analysis. He published more than 150 scientific papers. He is on the editorial board of Social Network Analysis and Mining (SNAM) and IxD&A. His research interests are community information systems, serious games, augmented reality, web engineering, social network analysis, requirements engineering and technology enhanced learning.

Proposing a Machine Learning Approach to Analyze and Predict Employment and its Factors

Francisco J. García-Peñalvo[1,3,4]*, Juan Cruz-Benito[1,3,4], Martín Martín-González[5], Andrea Vázquez-Ingelmo[1,3,4], José Carlos Sánchez-Prieto[1,4], Roberto Therón[1,2,3]

[1] GRIAL Research Group. University of Salamanca (Spain)
[2] VisUSAL Research Group. University of Salamanca (Spain)
[3] Department of Computer Science. University of Salamanca (Spain)
[4] Research Institute for Educational Sciences. University of Salamanca (Spain)
[5] UNESCO Chair in University Management and Policy. Technical University of Madrid (Spain)

Received 3 November 2017 | Accepted 22 January 2018 | Published 2 February 2018

unir
LA UNIVERSIDAD EN INTERNET

Abstract

This paper presents an original study with the aim of propose and test a machine learning approach to research about employability and employment. To understand how the graduates get employed, researchers propose to build predictive models using machine learning algorithms, extracting after that the most relevant factors that describe the model and employing further analysis techniques like clustering to get deeper insights. To test the proposal, is presented a case study that involves data from the Spanish Observatory for Employability and Employment (OEEU). Using data from this project (information about 3000 students), has been built predictive models that define how these students get a job after finalizing their degrees. The results obtained in this case study are very promising, and encourage authors to refine the process and validate it in further research.

Keywords

Employability, Employment, Artificial Intelligence, Machine Learning, Random Forest, Academic Analytics, OEEU.

DOI: 10.9781/ijimai.2018.02.002

I. Introduction

The concept of employability has steadily gained importance in recent years, becoming one of the pillars of the European educational strategy within the European Higher Education Area (EHEA) framework. However, the empirical research is still insufficient to build a strong theoretical foundation. It is worth noting that applied research on employability has, nowadays, an exploratory approach. This is because this research area presents difficulties regarding to have adequate, reliable and updated data, as well as this early status not only prevents agreement on research results are reached, but also poses many questions as to which methodologies and approaches are most appropriate to address these issues. For these reasons, the area is still growing and need to push the outcomes to further research levels.

Several research projects have been developed in recent years to provide more information on the employability of graduates, many of which have been driven by the OECD and the European Commission. These projects have faced at least two problems: first, the lack of a single, consensual definition of employability and, secondly, the difficulty of obtaining summary indicators to assess it.

Indeed, employability is a theoretical construct whose definition varies according to academic discipline and the perspective used, as well as the socioeconomic context to which it refers. There is no clear consensus on the factors that compose or determine it, nor on the employment outcomes to which it leads. Therefore, evaluating employability is a tough task. In any case, given the complexity of the notion of employability, it would be worth using several variables and indicators that assess different labor, educational and sociodemographic issues, rather than a summary indicator.

Most of the studies on employability that have been developed since the 1990s have focuses on identifying the competencies that graduates will need throughout their career path. Some have gone further, introducing other variables related to the training and education that offer the academic institutions, the sociodemographic context, the institutional and normative framework and the productive structure (“broad” approach as presented in [1]).

This kind of research, despite its initial status, is focused on develop successful strategies and outcomes that could help policymakers and institutions to enhance and promote those detected factors that contribute to get more chances of employment and better employments. For that reasons and its application in the society, it is possible to affirm that the project is in the scope of emergent areas like the Academic Analytics [2-6] or Institutional Intelligence [7, 8].

This paper aims to present a new method to analyze employability factors and to analyze how people gets employed. To achieve that, this paper proposes a machine-learning-based approach that produce predictive models on employment, providing the main factors that affect the predictive model and finding the most relevant ones. This approach contrasts to the previous state of the art in this research area. As will be explained in the Background section, previous approaches are based on basic statistical processes and tries to accomplish the problem of employment and its factors as a whole, instead of weighting the relevance of each factor to build more complex models. To illustrate these considerations, the paper provides a case of study where has been

* Corresponding author.

E-mail address: fgarcia@usal.es

applied the approach and shows some promising results.

The research presented in this paper is developed under the scope of the Spanish Observatory for University Employability and Employment (OEEU in its Spanish acronym) [9]. This observatory gathers data about employment and employability parameters among the Spanish graduates (after they leave the university) to analyze the information they provide and understand what the employment trends and most important employability factors are for this population [10]. To accomplish this mission, the observatory has developed a complex information system [6, 11-13] that collects and analyzes data to present the insights to the researchers [14, 15]. These data collected are used as a dataset to test the machine learning approach that will be presented in the following sections.

This paper has the following structure: second section (Background) presents the state of the art in the case of research applied on employability and employment analysis. Third section (Proposal) describes the machine learning approach and the methods and materials used in this research. Section fourth (Case Study: OEEU) shows the case study developed and the initial results achieved. Fifth section (Discussion) discusses about the implications of the research presented and the results achieved. Sixth section (Conclusions) finalizes the paper with some final remarks and introduces some future work.

II. Background

One of the main competencies studies promoted by the OECD was the "Definition and Selection of Key Competences" (DeSeCo). Since the first editions of the Programme for International Student Assessment (PISA) it had become clear that job success depended on a much greater range of competencies than those considered in the project. The DeSeCo project was created to identify these key competencies, aiming to serve as a framework to guide and complement two international programs to evaluate competencies: the aforementioned PISA and the Adult Literacy and Lifeskills (ALL). The DeSeCo project began in 1997 and ended in 2003, when it published the final report entitled "Key Competencies for a Successful Life and Well-functioning Society" [16]. In this project worked academics and experts from different fields (sociologists, philosophers, psychologists, economists, anthropologists, historians, statisticians, educators, etcetera) and social institutions (political parties, unions, employers, associations, etcetera) to define and figure out key competencies based on previous research. The final list of competencies was discussed in depth in two international symposia until achieve an agreement on the most important. This project is one of the foundational approaches and projects for the employment and employability analytics knowledge area.

One of the most influential projects in the EHEA when defining, identifying and classifying competences was the Tuning Educational Structures in Europe (usually called Tuning project). This project was funded by the European Commission within the Socrates framework. The project was divided into two phases, the first of which was more significant. It was active between 2000 and 2002. Its main objective was to figure out and classify the competences that the graduates require on their career path. Experts from different fields of knowledge (Business Administration, Geology, History, Mathematics, Physics, Education and Chemistry), from several European countries (Germany, Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Iceland, Italy, Norway, Netherlands, Portugal, Spain and Sweden) worked on the project. A total of 5803 graduates, 944 employers, and 998 scholars were surveyed; with the participation of more than 100 educational institutions of the European Union. The research finally divided the competencies into two main groups — specific and generic— and, in turn, the generic ones into three groups: instrumental, interpersonal and systemic.

Driven by the creation of the EHEA, different employability and competency projects have been developed that have enabled comparisons between European countries and universities using a common methodology. One of these projects was the "Higher Education and Graduate Employment in Europe" (usually called CHEERS project — "Careers after Higher Graduation. A European Research Study"). The project was promoted and financed by the European Commission within the "Targeted Socio-Economic Research Programme" (TSER). It began in 1997 and ended in 2000. It was led by The International Centre for Higher Education Research at the University of Kassel (Germany) and included other countries like Germany, Austria, Spain, Finland, France, Iceland, Italy, Japan, Norway, Netherlands, United Kingdom, Czech Republic and Sweden. Between 1998 and 2000 the different research groups sent a standardized questionnaire to the graduates who had completed their studies in the academic year 1994/1995. 37000 subjects were interviewed (about 3000 from each university). The information collected was based on their studies and the career path to analyze the relationship between higher education and employment (job position, mismatch in the labor market, etcetera). An important part of the questionnaire, in which the project had put special emphasis, was the assessment of the level of graduates' competencies and the level that they required by employers.

The CHEERS project was the starting point for the development, in 2006, of "The Flexible Professional in the Knowledge Society: New Demands on Higher Education in Europe" project, usually called REFLEX project. This project aimed to answer, among others, the following questions: what competences do graduates require fulfilling the demands from the modern knowledge society? To what extent has higher education provided these competencies? How can the mismatches between acquired and required competencies be solved?; To what extent are the graduates' expectations met? [17]. It was funded by European Union within the 6th Framework Programme (FP) for Research and Technological Development. It was led by the Research Center for Education and the Labor Market of the University of Maastricht. 14 European countries (Germany, Austria, Spain, Finland, France, Italy, Norway, Netherlands, United Kingdom, Belgium, Czech Republic, Portugal, Switzerland, Estonia) and Japan participated. The methodology and the questionnaire were like that adopted in the CHEERS project. A total of 40787 graduates were surveyed in 1999/2000, 5500 of which corresponded to graduates in Spanish universities (National Agency for Quality Assessment and Accreditation of Spain, 2008). Once again, information about the competencies that the project researchers considered relevant for the promotion of employability was compiled.

In Spain, the study carried out by the Catalan University Quality Assurance Agency (AQU) is one on the most important at the regional level. AQU conducts a telephone survey periodically (every three years) since 2001. The sample is made up of university graduates who finished their studies in any Catalan university three years before the date on which they were surveyed. Among other aspects the outcomes of this study are reports that analyze information related to the quality of employment, job stability, earnings, education-job skills match, job satisfaction, the process of finding a job, mobility, students' satisfaction with their studies, etc. In relation to competences, the questionnaire incorporates a section about skills acquired and their usefulness in the workplace.

Among the latest initiatives to evaluate the competencies of graduates in Spain highlight the *Libro Verde sobre la empleabilidad de los egresados de la Comunidad Valenciana* (Green book on the graduates' employability of Valencian Community). This book was published in 2013 because of a project promoted and funded by the *Generalitat Valenciana* and developed by the Valencian Agency of Assessment and Prospective (AVAP) and the universities of the region. The report presents the results of a study aimed to clarify the employability and employment situation of graduates at the *Comunidad Valenciana (*Valencian Community) and provided a series of conclusions and recommendations to reinforce and improve their employability. In this project was carried out a survey for those who had completed their short and long cycle studies during the years 2008 and 2009, comprising 2099 graduates in the study. This survey gathered information about the level of mastery of 21 generic competences that the graduates should have, as well as the level required by employers and the level acquired in universities.

Regarding to the methodologies applied in these studies, it is worth noting that they mainly use descriptive statistics, basically applying basic measures of central tendency and frequency distribution. They also rely mainly on basic charts to present the information in a simple way, basically using of histograms and bar charts. As an example, it is common to find in this type of studies the percentage distribution of graduates according to the strategy used in the search for employment, the time (months) taken for graduates to find their first job, the average earnings, the percentage of graduates satisfied with their jobs or their studies, the distribution of the graduates according to the education and job match, the average level of competences required by employers and acquired at the universities, etc. The number of variables involved in these studies varies notably depending on the scope of the project and its objectives. Among the largest is the REFLEX project, which almost included 500 variables distributed in 11 categories: study program, other educational and related experiences, transition from study to work, first job after graduation, employment history and current situation, current work, work organization, competencies, evaluation of study program, values and orientations, about yourself, plus the variables of country and study identification.

This kind of studies have attracted the attention from numerous researchers from different disciplines who, given the increasing availability of data, have been able to carry out some empirical papers. Thus, much of the research has focused on identifying the competencies required by employers and evaluating their impact on employment [18-21] as well as assessing the mismatches between acquired and required competencies [22, 23]. In addition, these research works allowed deepening the relationship education-labor market, as well as the mismatches and their effects [23-28]. They have applied all kinds of analytical methodologies. Although there is no common practice, they use more sophisticated techniques and tools than those applied by general studies, linked to econometrics, psychometry and other quantitative and qualitative methods from research in social sciences.

III. Proposal

In the case of this paper, researchers followed this kind of new approaches. Using novel methods and techniques like machine learning could open new possibilities and ways to work in assessing employability and employment. Also, these methods could unleash new ways for manage huge amounts of information as a whole but considering each factor in a weighted way within the predictive models to be built. Previously some projects use basic statistics due the difficulty of handling big datasets in a non-automated mode, but in our way, the same procedure can be used to crunch all data related jointly.

In general, the proposal for the analysis (based on machine learning) follows common principles in data science regarding data structuration, tidy data approaches, etc. [29-31]. Because it is a proposal and it is on its initial stage, the approach should be explainable to assess its appropriateness. For this reason, the machine-learning process has been implemented in a white-box way; thus, the researchers have selected algorithms and methods to make the workflow explainable.

Moreover, these main principles, the different details for the analysis pipeline, and methods used in this research are presented below. All of these details are explained in the following workflow (available at [32]):

1. Retrieve dataset about students from OEEU's information system.
2. Filter the desired fields from the datasets and enclosed them in data frame (a data structure like a table).
3. Data cleaning: remove noise data, remove columns (variables) with too many null (NaN) values, and remove all students who have only partial.
4. Normalize data with the One-hot encoding algorithm for categorical values in columns [33].
5. Considering the data gathered and the kind of variable (labeled) to predict (student gets employed or not), the algorithm to use must be related to supervised learning. This is because this kind of algorithm makes predictions based on a set of examples (that consist of a labeled training data set and the desired output variable). Moreover, regarding the dichotomous (categorical) character of the variable to predict, the supervised learning algorithm to apply must be based on classification (binary classification, as we have a label of finalization equal to true or false). According to the authors' previous experience, the possibility of explaining results and the accuracy desired for the classification, a Random Forest classifier algorithm [34] was selected. In this step, the Random Forest algorithm was executed repeatedly to determine the best setup for the dataset given (obtaining the most adequate parameters for the execution).
6. With the best configuration found, train the random forest algorithm (with 33.33% of the dataset) and obtain the predictive model.
7. Using the predictive model, obtain the most important features for the predictive model.

At this point, researchers have built a model that could predict if a person will get employed or not, showing also what are the factors that affect more the result. After that, researchers could use these factors to filter information, generalize knowledge across the dataset, extract what values on these factors lead to get employed or not, etc. As an example, using these most relevant factors, researchers could clusterize students to gain deeper knowledge about what are the main characteristics between those who get an employment and those who do not.

In general, in this paper will be demonstrated how this kind of approach could be suitable for the goal of modelling employment and its factors. To do that, the algorithms and the code used is available publicly at [32]. Unfortunately, due the privacy restrictions that affect the OEEU will not be revealed some kind of data involving individual graduates, universities or other sensitive information, it will be only displayed aggregated (and anonymized) information. For that reason, regarding the processes related to analyze the factors that define the model, only will be shown some generalist figures that could illustrate the procedure and give some clues about a real implementation.

The programming language used to conduct all the analyses and calculations was Python. The Python software tools and libraries used to code and test the approach were:

- Pandas software library [29, 35], to manage data structures and support analysis tasks.
- Scikit-learn [36] library, to accomplish the machine learning workflow [33].
- Jupyter notebooks [37-39], to develop the Python code used in this research.

IV. Case Study: OEEU

As presented in the previous sections, the projects that investigated employability and employment disciplines varied in number and type of variables available, scope of the project, etc. For this paper, researchers used the information gathered by the Spanish Observatory for Employability and Employment (OEEU). This project keeps information about 182000 students graduated from degree and master studies. It includes about 400-500 variables per each edition of the study (one edition about degree graduates – 134129 students involved – and other edition about master studies –47822 students–).

The data used in this case study correspond to the information available from graduated students that finalized a degree in the course 2009-2010. This is, information about 134129 students, with 493 variables per student [9, 32]. Despite of the dimensions of the dataset, it is worth noting that not all the students have information for all the possible variables. The Observatory gathers the information by using two input methods: the raw records from the Spanish universities and the information provided by the students through fulfilling questionnaires. These two main sources of information have only part of the variables marked as required, for that reason some of them appear with empty values. Also, the fact of ingesting information via web forms (like in this case) make extremely difficult to get all the information, because the graduate can quit the web form in any moment. For that reason, are required methods that clean and wrangle the information like those presented in the previous section.

Apart of cleaning and wrangling the data properly, researchers have excluded some variables included in the OEEU's dataset, since they are related exclusively to some universities (the universities could add some questions to the OEEU questionnaire), and using only the common variables to all students in their test to create the predictive model related to employment. This reduced the dataset from 493 total variables to 383 possible variables per student. These 383 variables per student can be observed in the 8th cell at the provided notebook [32].

Following the workflow outlined in the previous section, after filtering the desired variables, researchers cleaned all those variables which presents to much *NaN* (empty) values. In this case, the threshold used to remove all the weak variables was 10%. This is, all the variables with more than 10% of empty values were discarded for the model construction. This threshold is strict to obtain a stronger model. There are other common procedures to deal with void variables or measures (fill the empty values with others from the dataset, with the mean of the column, etc.), but in this case, researchers preferred to avoid any kind of artificial data that could contaminate the result. After removing all these non-valuable variables, researchers dropped all the students that had any empty value in their information (completing by this way the data-cleaning stage). After this hard-cleaning process, researchers counted with 26 data variables from 9744 graduates. As previously commented, following other conservative methods to deal with empty values would lead to use more variables (columns) and observations (rows), but this is not the focus in this initial test of the approach presented.

After all this work in data preparation, began the machine learning phase. In this case, the third part of the dataset (third part of graduates) was marked as the portion to train the random forest algorithm. Also, researchers selected the variable to predict using the others. In this case, the variable to predict was *'haEstadoDesempleado'* (Have the student been unemployed?) which contains two possible values: 0 (false) value for those students that got a job after finalizing the degree, and 1 (true) for those students that were not employed.

After that, and with this 33% of the observations (3305 students) and the variable to predict, researchers tested programmatically the best setup for the random forest algorithm. This is, the best configuration values for the parameters *randomforestclassifier__max_features* and *randomforestclassifier__min_samples_leaf.* Using the best values found for the parameters, researchers executed (trained) the algorithm to get the corresponding predictive value.

Table I presents the quality metrics [40] of the predictive model built to predict the graduates' employment. As displayed, the precision of the predictive model classifying and predicting the employment or not was of 0.71 (where 0 is the worst precision and 1 the best).

TABLE I. Results of the Predictive Model Built

	Precision[a]	Recall[b]	F1-score[c]	Suport[d]
False	0.73	0.12	0.20	1066
True	0.70	0.98	0.82	2239
Avg / total	0.71	0.70	0.62	3305

[a]The precision is the ratio *tp / (tp + fp)* where *tp* is the number of true positives and *fp* the number of false positives. The precision is intuitively the classifier's ability of not labeling as positive a sample that is negative. This score reaches its best value at 1 and worst score at 0.

[b]The recall is the ratio *tp / (tp + fn)* where *tp* is the number of true positives and *fn* the number of false negatives. The recall is intuitively the ability of the classifier to find all the positive samples. This score reaches its best value at 1 and its worst score at 0.

[c]The F1 score can be interpreted as a weighted average of the precision and recall, where an F1 score reaches its best value at 1 and its worst score at 0. The relative contribution of precision and recall to the F1 score is equal. This score reaches its best value at 1 and its worst score at 0.

[d]The support is the number of occurrences of each class in each predicted label.

On the other hand, the crosstab that expresses the number of good and bad predictions for the predictive model can be found in Table II.

TABLE II. Crosstab for the Predictive Model Built

	False (predictions)	True (predictions)
False (actual)	125	941
True (actual)	46	2193

Following the process, the important factors found for the model are presented in the Table III. Despite the random forest provide a importance score for all the variables involved in the predictive model, researchers stated 0.05 as the minimum value to consider the factor as relevant. This is because 0.05 is a common value to ensure reliable results in several analytical processes.

TABLE III. Most Relevant Factors for the Predictive Model Built

Name of the variable	Explanation	Importance score
universidad_id	The university where studied the graduate	0.437347
otrosCriterios SeleccionarPuesto Trabajo_circunstancias Personal	The graduate's opinion about the relevance of choose a job depending on the conditions related to personal context: family conciliation, etc.	0.150274
sexo_id	Graduate's gender	0.107880
residenciaExtranjero Motivos_cod	Graduate's reasons to live abroad during the degree	0.097754
otrosCriterios SeleccionarPuesto Trabajo_prestigio	The graduate's opinion about the relevance of choose a job depending on the conditions related to the prestige of the employer, the tasks to be done or to the position within the company.	0.075377
titulacion_id	The degree studied by the graduate.	0.054354

The importance score varies between 0–1, where 1 is the best score and 0 the worst one

As previously presented, after obtaining the predictive model, researchers can use the most relevant factors to analyze in deep what are the specific situations (values of factors) that lead students to get or not a job. For example, using the most relevant factors, could be generated clusters that group graduates using their similar characteristics. As an example, Fig. 1 presents the different clusters obtained after applying a hierarchical clustering algorithm to the clean dataset using the factors as variables to group the users. The representation is truncated to show the more related clusters jointly (showing only 12), but in fact, applying the hierarchical clustering were obtained 55 different groups of students.

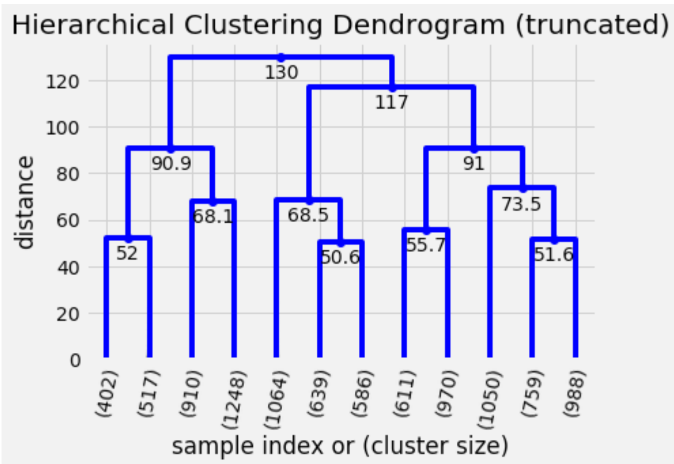


Fig. 1. Dendrogram that represents the clusters of graduates regarding to the most relevant factors detected in the random forest. Each leaf represents a different cluster obtained. The different values that appear near the claves display the Euclidean distance that explains the separation between the different clusters. Finally, the numbers below the leaves (at the bottom of the figure) present the number of users included in the corresponding cluster.

For example, if researchers choose one of the 55 resulting clusters, could observe that (cluster 22):

"229 students from 14 different universities, all of them women who studied one of 20 selected degrees, who do not consider the prestige of employer as a key factor to select a job, who lived abroad during the degree mainly because they work abroad, and do not consider the conditions related to personal context as a key factor to select a job have a chance of get job of 86,90%."

As previously explained, the full information about clusters and results obtained after applying the process is not fully provided in the paper or in an external notebook because of the Observatory's legal and privacy restrictions. Also, this kind of information is out of the scope of this paper, since it is focused on explaining how machine learning methods could be applied to this problem.

V. Discussion

As outlined in the background, the research on employability and employment is a knowledge area in development. The main projects developed in the previous years have pursued to define the main factors that define the employability and employment. The research methods used previously were related to basic statistics and simple analysis, despite some independent researchers went deeper in the methods, applying other related to econometrics, psychometry and other quantitative and qualitative methods of social research. One of the problems observed by the researchers when working with this kind of projects (i.e. in the case of the OEEU), is how to manage to handle large amounts of data and make more specific analysis. The approach presented in this paper regarding the application of machine learning methods, allows to automate some part of the analysis while it allows to gain general knowledge and deal with the problem of analyze how people gets employed from a broader perspective, understanding all the data as a whole.

The machine learning approach has been applied successfully previously by the authors and other researchers in fields like Human-Computer Interaction [41], education [42], etc. For that reason, was tried this approach in a complex research like employment and employability.

Considering the case study presented, the results are quite promising. Despite of the complexity of the data, and the different issues with the dataset, researchers have been able to get a predictive model with a 0.71 precision score (0 the worst score, 1 the best one). This result opens the possibility of keep working in this approach to enhance the results and try deeper analysis. Regarding other scores achieved like, F1, recall, etc., should be outlined that the model built performs poorly in detecting the real "False" values (not employed students), so it could lead to bad predictions regarding students without employment.

In this case, researchers are confident in that managing better the empty values and applying other kind of strategies in data cleaning and wrangling will allow to get better predictive models and outcomes. Following with the results and the case study, it is worth noting that the predictive model involved a support of 3305 observations (graduates), which is a high number in the case of a test like this. It is because researchers want to try a real case to validate the seminal idea of applying machine learning. Despite the model and procedures should be validated in a better way and tested more, the results achieved in a real case like this are considerable.

About the predictive model built, it is not surprising that the some of the most important factors that define if a student gets employment or not after graduating are those related to the university or the degree. Currently there are severe gaps regarding the employment ratio depending on the studies and the students' knowledge area. Also, the predictive model highlights other factor sadly known nowadays: the gender. There are many international studies that deals with the fact that the gender (specially for women) is a handicap to apply for some jobs and positions. The predictive model built found also this variable as a fundamental factor that define how graduates achieve an employment. On the other hand, the predictive model also presents other factors that are not as well-known as the previous ones: the reasons to live abroad during the studies, the importance of employer prestige to choose a job, or the facilities provided by the employer to adjust and balance personal context and life with the work. In general, it is clear that these factors are truly relevant in the model. These 5 factors sum a score of more than 0.8 out of the maximum 1 score that could be achieved by all the 383 factors included in the predictive model. Also, it is possible to think that the least relevant factor of these 5 (the degree obtained by the student) has a very low score (0. 054354), but this score is the weight of a solely factor in a model with 383 different factors, so achieve a 0.05 score out of 1 maximum score it is not low with this amount of different variables observed.

This kind of algorithmic approaches that include all the factors as part of possible models, shed light over some aspects avoiding previous bias and prejudices. In contrast to the background, where the international research community define the factors to study in each project, in this case, authors propose to use all the factors letting the machine learning algorithms to select on their own those truly relevant factors. This switch the *traditional* approach, making the exploratory process to depend only on the dataset available and adapting the focus to the facts and metrics obtained previously.

Also, are provided within the case study some examples on how to obtain deeper information and insights about the concrete metrics that affect the employment. This is, considering the most relevant factors, how to get common characteristics between those students who get an employment and those who do not. There are other many approaches to make this process possible, but the proposed using clustering analysis is evident and easy to understand and apply.

Regarding the implications of this research, authors agree to point out that this kind of approaches could open new possibilities on using data to enhance the education and students' opportunities in labor market. Following with the idea of Academic Analytics, this kind of employment analytics could be included in academic intelligence processes in universities and other higher education institutions. Also, policymakers could follow this kind of data-driven employability and employment analytics to design and propose new ways of preparing students to their professional future.

VI. Conclusion and Future Work

This paper presents a novel study in the field of employability and employment analytics. The main results achieved have been quite promising and encourage authors to continue the labor of improving the generation of predictive models for employability and employment. The nature of this kind of problems is extremely complex and varies on the time, but with this kind of algorithmic and automated processes could address it better than the traditional approaches. Based on the results, the authors are committed to continue developing the approach to get better results and improve the process until it could be applied successfully in further research works.

Acknowledgment

The research leading to these results has received funding from "la Caixa" Foundation. The author Juan Cruz-Benito would like to thank the European Social Fund and the *Consejería de Educación* of the *Junta de Castilla y León* (Spain) for funding his predoctoral fellow contract. This research has been cofounded by the University of Salamanca through the program of financial aid for the predoctoral contract of José Carlos Sánchez-Prieto (*Programa III: Ayudas para contratos Predoctorales*) cofounded by Banco Santander. This work has been partially funded by the Spanish Government Ministry of Economy and Competitiveness throughout the DEFINES project (Ref. TIN2016-80172-R).

References

[1] R. W. McQuaid and C. Lindsay, "The concept of employability," *Urban studies,* vol. 42, no. 2, pp. 197-219, 2005.

[2] P. Baepler and C. J. Murdoch, "Academic analytics and data mining in higher education," *International Journal for the Scholarship of Teaching and Learning,* vol. 4, no. 2, p. 17, 2010.

[3] J. Bichsel, *Analytics in higher education: Benefits, barriers, progress, and recommendations*. EDUCAUSE Center for Applied Research, 2012.

[4] J. P. Campbell, P. B. DeBlois, and D. G. Oblinger, "Academic analytics: A new tool for a new era," *EDUCAUSE review,* vol. 42, no. 4, p. 40, 2007.

[5] D. A. Gómez Aguilar, F. J. García-Peñalvo, and R. Therón, "Analítica Visual en eLearning," *El Profesional de la Información,* vol. 23, no. 3, pp. 233-242, 2014.

[6] F. Michavila, M. Martín-González, J. M. Martínez, F. J. García-Peñalvo, and J. Cruz-Benito, "Analyzing the employability and employment factors of graduate students in Spain: The OEEU Information System," presented at the Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, 2015.

[7] Oficina de Cooperación Universitaria. OCU, "Libro Blanco Inteligencia Institucional en Universidades," ed: Madrid: OCU, 2013.

[8] F. J. García-Peñalvo, "Inteligencia Institucional para la Mejora de los Procesos de Enseñanza-Aprendizaje," *GRIAL Research Group*.

[9] F. Michavila, J. M. Martínez, M. Martín-González, F. J. García-Peñalvo, and J. Cruz-Benito, "Barómetro de Empleabilidad y Empleo de los Universitarios en España, 2015 (Primer informe de resultados)," 2016.

[10] W. Greller and H. Drachsler, "Translating learning into numbers: A generic framework for learning analytics," *Journal of Educational Technology & Society,* vol. 15, no. 3, pp. 42-57, 2012.

[11] A. Vázquez-Ingelmo, J. Cruz-Benito, and F. J. García-Peñalvo, "Improving the OEEU's data-driven technological ecosystem's interoperability with GraphQL," presented at the Fifth International Conference Technological Ecosystems for Enhancing Multiculturality 2017 (TEEM'17), Cádiz, Spain, October 18-20, 2017, 2017.

[12] A. Vázquez-Ingelmo, J. Cruz-Benito, F. J. García-Peñalvo, and M. Martín-González, "Scaffolding the OEEU's Data-Driven Ecosystem to Analyze the Employability of Spanish Graduates," in *Global Implications of Emerging Technology Trends*, F. J. García-Peñalvo, Ed. Hershey, PA: IGI Global, 2018, pp. 236-255.

[13] A. García-Holgado, J. Cruz-Benito, and F. J. García-Peñalvo, "Analysis of Knowledge Management Experiences in Spanish Public Administration," presented at the Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15), Porto, Portugal, 7-9, October, 2015.

[14] J. Alcolea Picazo and S. Pavón de Paula, "Los datos como recurso estratégico Libro Blanco Inteligencia Institucional en Universidades (pp. 17-43)," *Madrid, Spain: OCU (Oficina de Cooperación Universitaria),* 2013.

[15] M. Zeleny, "Management support systems: Towards integrated knowledge," *Human systems management,* vol. 7, pp. 59-70, 1987.

[16] D. S. Rychen and L. H. Salganik, *Key competencies for a successful life and well-functioning society*. Hogrefe Publishing, 2003.

[17] National Agency for Quality Assessment and Accreditation of Spain, *Titulados universitarios y mercado laboral, Proyecto REFLEX*. Madrid: ANECA, 2008.

[18] R. G. Biesma, M. Pavlova, G. Van Merode, and W. Groot, "Using conjoint analysis to estimate employers preferences for key competencies of master level Dutch graduates entering the public health field," *Economics of Education Review,* vol. 26, no. 3, pp. 375-386, 2007.

[19] A. García-Aracil and R. Van der Velden, "Competencies for young European higher education graduates: labor market mismatches and their payoffs," *Higher Education,* vol. 55, no. 2, pp. 219-239, 2008.

[20] H. Heijke, C. Meng, and G. Ramaekers, "An investigation into the role of human capital competences and their pay-off," *International Journal of Manpower,* vol. 24, no. 7, pp. 750-773, 2003.

[21] E. Kelly, P. J. O'Connell, and E. Smyth, "The economic returns to field of study and competencies among higher education graduates in Ireland," *Economics of Education Review,* vol. 29, no. 4, pp. 650-657, 2010.

[22] M. J. Freire Seoane, M. Teijeiro Álvarez, and C. Pais Montes, "La adecuación entre las competencias adquiridas por los graduados y las requeridas por los empresarios," 2013.

[23] P. Kellermann, "Acquired and Required Competencies Of Graduates," in *Careers of university graduates: Views and experiences in comparative perspectives*, vol. 17, U. Teichler, Ed. Dordrecht: Springer Science & Business Media, 2007, pp. 115-131.

[24] P. Kellermann and G. Sagmeister, "Higher education and graduate employment in Austria," *European Journal of education,* vol. 35, no. 2, pp. 157-164, 2000.

[25] H. Schomburg and U. Teichler, *Higher education and graduate employment in Europe: results from graduates surveys from twelve countries*. Springer Science & Business Media, 2007.

[26] U. Teichler, "Research on the relationships between higher education and the world of work: Past achievements, problems and new challenges," *Higher Education,* vol. 38, no. 2, pp. 169-190, 1999.

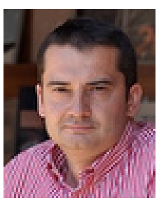
[27] U. Teichler, "Graduate employment and work in selected European countries," *European Journal of Education,* vol. 35, no. 2, pp. 141-156, 2000.

[28] U. Teichler, "New perspectives of the relationships between higher education and employment," *Tertiary Education & Management,* vol. 6, no. 2, pp. 79-92, 2000.

[29] W. McKinney, *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. " O'Reilly Media, Inc.", 2012.

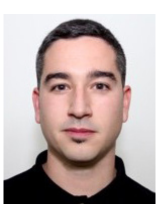
[30] W. McKinney, "Data structures for statistical computing in python," in

Proceedings of the 9th Python in Science Conference, 2010, vol. 445, pp. 51-56: SciPy Austin, TX.
[31] H. Wickham, "Tidy data," *Journal of Statistical Software,* vol. 59, no. 10, pp. 1-23, 2014.
[32] J. Cruz-Benito. (2017). *Jupyter notebook developed to support the research presented in the paper "Proposing a machine learning approach to analyze and predict employment and its factors"* Available: https://github.com/juan-cb/paper-ieeeAccess-2017
[33] S. Raschka, *Python machine learning*. Packt Publishing Ltd, 2015.
[34] L. Breiman, "Random Forests," *Machine Learning,* journal article vol. 45, no. 1, pp. 5-32, October 01 2001.
[35] W. McKinney. (2017). *Pandas, Python Data Analysis Library. 2017*. Available: http://pandas.pydata.org/
[36] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research,* vol. 12, no. Oct, pp. 2825-2830, 2011.
[37] M. Ragan-Kelley *et al.*, "The Jupyter/IPython architecture: a unified view of computational research, from interactive exploration to communication and publication," in *AGU Fall Meeting Abstracts*, 2014.
[38] F. Perez and B. E. Granger, "Project Jupyter: Computational narratives as the engine of collaborative data science," Technical Report. Technical report, Project Jupyter2015.
[39] T. Kluyver *et al.*, "Jupyter Notebooks-a publishing format for reproducible computational workflows," in *ELPUB*, 2016, pp. 87-90.
[40] Scikit-learn. (2017, 4/09/2017). *API Reference - scikit-learn documentation: Metrics*. Available: http://scikit-learn.org/stable/modules/classes.html - module-sklearn.metrics
[41] J. Cruz-Benito, A. Vázquez-Ingelmo, J. C. Sánchez-Prieto, R. Therón, F. J. García-Peñalvo, and M. Martín-González, "Enabling Adaptability in Web Forms Based on User Characteristics Detection Through A/B Testing and Machine Learning," *IEEE Access,* vol. 5, 2017.
[42] A. Zollanvari, R. C. Kizilirmak, Y. H. Kho, and D. Hernández-Torrano, "Predicting Students' GPA and Developing Intervention Strategies Based on Self-Regulatory Learning Behaviors," *IEEE Access,* 2017.



Francisco J. García-Peñalvo

Francisco J. García-Peñalvo received the degrees in computing from the University of Salamanca and the University of Valladolid, and a Ph.D. from the University of Salamanca. He is currently the Head of the Research Group Interaction and eLearning. He has led and participated in over 50 research and innovation projects. He was the Vice Chancellor of Innovation with the University of Salamanca from 2007 to 2009. He has authored over 300 articles in international journals and conferences. His main research interests focus on eLearning, computers and education, adaptive systems, Web engineering, semantic Web, and software reuse. He was a Guest Editor of several special issues of international journals, such as Online Information Review, Computers in Human Behavior, and Interactive Learning Environments. He is also a member of the Program Committee of several international conferences and a Reviewer of several international journals. He is currently the Editor-in- Chief of the International Journal of Information Technology Research and the Education in the Knowledge Society Journal. He is also the Coordinator of the multidisciplinary Ph.D. Programme on Education in the Knowledge Society.



Juan Cruz-Benito

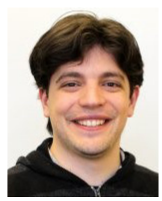
Juan Cruz-Benito received an M.Sc. degree in intelligent systems from the University of Salamanca, Spain, in 2013, where he is currently pursuing a Ph.D. in computer sciences. He is one of the youngest members of the Research Group Interaction and eLearning, where he specializes in software solutions based on technology ecosystems and open source software. He works in Human-Computer Interaction, educational virtual worlds and technologies for educational purposes, disciplines that he has developed in many innovation and research projects. He has participated in many European and national R&D projects, such as TRAILER, VALS, USALSIM Virtual Campus, and the Spanish Observatory for University Employability and Employment (OEEU), where he participated as a software engineer, researcher, and developer.

Martín Martín-González

Martín Martín-González is a researcher at the UNESCO Chair in University Management and Policy of the Technical University of Madrid (UPM). Prior to his work at the UPM he worked as a researcher in the Faculty of Economics in the Autonomous University of Madrid (UAM). He holds a PhD in Economics and a master's degree in Economic Development and Public Policy from the UAM. He majored in economics at the University of La Laguna (ULL). His areas of research are the Economics of Education, the Economics of Higher Education, Employability, the Evaluation of Educational Policies, Public Economics, Applied Economics, Economic Development, Higher Education and Vocational Training.

Andrea Vázquez-Ingelmo

Andrea Vázquez-Ingelmo was born in Salamanca, Castilla y León, Spain in 1994. She received a bachelor's degree in Computer Engineering from the University of Salamanca, Salamanca, in 2016, and she is currently completing her master's degree in computer engineering from the same university. She is a member of the Research Group of Interaction and eLearning (GRIAL). Since 2016, she has been part of the national project "Spanish Observatory for University Employability and Employment" as a developer and researcher.



José Carlos Sánchez-Prieto

José Carlos Sánchez-Prieto received a bachelor's degree in Pedagogy and a master's degree in ICT applied in education from the University of Salamanca (Spain). Currently he is a Researcher in Training at the Faculty of Education of said university, where he conducts his PhD research within the Programme on Education in the Knowledge Society. His area of research is the assessment of attitudes among in-service and pre-service teachers.

Roberto Therón

Roberto Therón studied Computer Science at the University of Salamanca (Diploma) and the University of La Coruña (BA). After joining the Research Group Robotics at the University of Salamanca, he presented his thesis work, "Parallel calculation configuration space for redundant robots," receiving the Extraordinary Doctoral Award. He subsequently obtained a Bachelor in Communication Studies (University of Salamanca) and a Bachelor in Humanities (University of Salamanca). At the same university, he continues to carry out his research work as the manager of the VisUsal group (within the Recognized Research Group GRIAL), which focuses on the combination of approaches from computer science, statistics, Graphic Design and Information Visualization to obtain an adequate understanding of complex data sets. In recent years, he has been dedicated to developing advanced visualization tools for multidimensional data, such as genetics or paleo-climate data. In the field of Visual Analytics, he develops productive collaborations with groups and institutions internationally recognized as the Laboratory of Climate Sciences and the Environment (France) or the Austrian Academy of Sciences (Austria). He is the author of over 100 articles in international journals and conferences.

Exploring the Benefits of Using Gamification and Videogames for Physical Exercise: a Review of State of Art

Carina S. González-González*, Nazaret Gómez del Río, Vicente Navarro-Adelantado

ITED Research Group – University of La Laguna, Tenerife (Spain)

Received 9 January 2018 | Accepted 20 March 2018 | Published 29 March 2018

Abstract

There is a lack of motivation in children and adolescents to do physical exercise and at the same time a worldwide obesity epidemic. Gamification and active videogames can be used to increase the motivation of young people, promoting healthy habits. In this work we explore different studies on active videogames, eSports and gamification applied to physical exercise and health promotion. Main findings include positive effects in a reduction in body weight and in the promotion to continue performing of physical exercise. It also contributes to increase the motivation in children and adolescents to practice exercise. The personalization of user experience and emerging technologies (big data, wearables, smart technologies, etc.) are presented as promising opportunities to keep the engagement in game-based program and gamification of physical exercise.

Keywords

Physical Exercise, Gamification, Active Videogames, Literature Review, State of The Art.

DOI: 10.9781/ijimai.2018.03.005

I. Introduction

In the school or outside, physical exercise (PE) is considered a positive element and widely as a fun, engaging and social activity. Inside schools, there is a curricular formal PE, named Physical Education, which main goals is to develop motor skills, knowledge and healthy behaviors. Outside schools, PE can have many non-formal ways: fitness, sports, dancing, etc. But, many times, the PE lessons in school are the main PE that young people and children have [1]. Physical and active games can attract children and young people to have a regular PE and in this way, to promote healthy habits and wellbeing.

PE can be gamified or transformed into an active game if we consider that any process that satisfies the following premises can be gamified: "the activity can be learned; the user actions can be measured, and the feedbacks are timely delivered to the user" [2].

In the creation of active games, as in other types, designers must take into account fundamental elements. These game elements can be classified in many diverse ways, for example, as:

- Mechanic, story, aesthetics, and technology [3].
- Interfaces, rules, entity manipulations and goals [4].
- Mechanics, dynamics and aesthetics [5].
- Mechanics and dynamics [6].
- Dynamics, mechanics and components [7], organized in a pyramid structure, according if the element is conceptual or tactical.

Although, the most common elements associated to gamification are points, badges and leader-boards (PBL), there are diverse frameworks to design or gamify systems with a variety of elements related to the intrinsic and extrinsic motivation of user [8, 9]. These elements allow the design of personalized gamified experiences according to the user preferences.

* Corresponding author.

E-mail address: cjgonza@ull.edu.es

In this paper we perform a literature review to extract the elements of games and gamification used for PE and their effectiveness.

II. Studies About Games and Gamification for PE

In the literature we can find several studies about games and PE related with the energy expenditure in educational programs carried out in schools combined with children's leisure time [10]. Main findings include that these kind of physical educative programs in schools increase the motivation of children and adolescents to continue performing physical exercise [11]. But, not all physical activity impacts in the reduction of body weight, because it depends on the frequency, duration and intensity of the activity [12]. For example, in the study carried out in [13] related to the dancing game, it was found that the cardiorespiratory answer was comparable to an aerobic dance of medium to high intensity [13].

Regarding the motivation to PE, some studies conclude that is better to encourage children participating in team rather than in individual sports [14]. Also, in other study on a group of overweight children and adolescents using an exergame (Dance Dance Revolution of Konami) as a routine physical activity, found that the game was not enough to motivate them to participate. Then, to increase participation researchers encouraged cooperative play, increased the musical variety and included a competitive mode in the activities [15]. Besides, a dancing videogame used in children´s homes was evaluated while they were playing weekly in group. Results showed that the motivation and participation increased due to the group sessions [16].

Moreover, the use of active videogames has positive effects promoting an active lifestyle. A study compared the energy expenditure (EE) required by a sedentary game and two active videogames, finding that the non-active game increased 22% the EE, in contrast to active games, that increased 108% in the case of torso's movements and 172% in the case of dancing. The same study found that non-obese children had lower EE when playing the dancing video game than

TABLE I. Gamification, Games And Physical Exercise

Reference	Research	Major Findings
[10],[11], [12]	Physical activity and energy expenditure through educational intervention programs	Positive effects in a reduction in body weight and in the promotion of physical exercise.
[14], [15], [16]	Motivation	Increasing the motivation in children and adolescents to practice exercise.
[12], [13], [14], [17], [18], [19], [20],[21], [22]	Energy expenditure with active games	Energy expenditure of an active game is significantly higher than that derived from other sedentary activities or non-active videogames. Children burned the same amount of calories when they walked moderately and three times more than while resting (moderate intensity).
[23],[24], [25],[26], [27],[28]	Interactive technologies to promote healthy habits in children	Gaming platforms, playgrounds and technologies, with gestural and body movement as an interactive element, promotes physical activity. Also, mobile devices and wearables promotes outdoor physical activity.
[29],[30], [31],[32]	Design of active games (collaboration, social aspects, structural elements, etc.)	Patterns for design collaborative games and structural framework based on the fundamentals of motor play to guide the design and evaluation of active videogames.

obese children [17]. Other studies about EE with active games found that the EE was significantly higher in these kind of games than with conventional videogame [18], [19]. Similar findings of other study showed an increase over the base line of between 120% and 140% in the EE and the energy consumed playing active games, similar when the participant did other PE (walk, jogging, swimming) [20]. Besides, in other study the calories consumed at rest, while watching television and while walking was measured in children and compared with the calories burned when playing active videogames. The results revealed that children burned the same amount of calories when they walked moderately and three times more than while resting [21].

Although positive effects were found in several studies on the EE with active videogames, PE and sports cannot be replaced. As we mentioned above, EE depends on the intensity, duration and frequency of the activity and only a few active videogames allow performing PE with moderate intensity [22].

Several studies describe the effective use of technologies with children in the promotion of healthy habits [23]. Some gaming platforms includes body movement as an interactive element that can be used for PE [24]. Also, outdoor PE are favored by smartphones supporting different sensors tracking biometrics and allowing augmented reality activities [25]. Other technology is the Playware based on the use of sensors, actuators, hardware and software for playgrounds [26]. In Playgrounds, some game elements like social interaction, simplicity, challenge, goals, and feedback should be considered [27]. Playgrounds can be used to spatial cognitive development, considering multiple perspectives, zooming in and out, distances, experiencing movement or finding visual cues [28].

Other game elements that should be considered in the design of educative videogames in general and for active videogames in particular are collaboration and social aspects [29]. A set of theoretical elements for the design of educational videogames were proposed [30] and for monitoring and evaluating educational videogames [31]. A framework based on the fundamentals of motor play to guide the design and evaluation of active videogames have been developed [29]. Other study, analyzed the effectiveness of commercial active platform (Nintendo Wii) and an active platform designed following principles of educational, collaborative and active videogames [32]. In Table I is summarized the major findings of the literature review analyzed regarding gamification, games and physical exercise.

III. eSports and Physical Exercise

Electronic sports (eSports), videogames competitions (digital sports, exergaming, cybersport, etc.) are gaining popularity around the world [33]. Perhaps, the name eSports is based on the transference of the classic classification of "sports" for electronic games that were based on a sport. Despite the eSports are officially accepted as sport in about 60 countries [34], still there is no consensus in a common definition of eSports [33]. Some authors define "eSport" as "an umbrella term used to describe organized, sanctioned video game competitions, most often in the context of video game tournaments" [34]. Other authors define" eSports" as an "area of sport activities in which people develop and train mental or physical abilities in the use of information and communication technologies" [35]. Thus, the term "eSports" sometimes is used as a direct synonymous of digital "sport", but they are not the same. Maybe this confusion is because there are concepts, like 'sport gaming,' 'virtual sports,' and 'exergaming', that are being used to describe the digitalization of playful activities in different ways [35].

The eSports are considered as a "sedentary activity" (or with a low level of physical activity) to be considered an "sport" [36]. And, in the philosophy of sport literature there seems to be 'a solid point of agreement in that the physical skill is a necessary component of all sports [37]. But among the eSports there are games which require physical skill and games which do not [38]. Some authors claims that eSports "require the learning and performance of motor skills and that embodiment within a virtual environment may be considered playful or even athletic" [39]. So, some eSports can be utilized for the development of motor skills, but maybe or not be implies physical exercise [40], as we can observe in examples of eSports with official competition leagues like Counter Strike, WOW (World of Warcraft), League of Legends or FIFA [40].

Although, certain types of eSports cannot be left out of the category of sports, there is not a common criterion about to including eSports into educational contexts [41]. The authors argue that although some eSports can be considered games because it implies the learning of motor skills [39], however, do not concluded that these skills should be taught in physical education programs. They argue that the embodied interaction and visibility of movement behavior can useful to learn about the movement and physical education [39].

IV. Related Projects for Physical Exercise

There are several research projects related to gamification, education and PE. In Table II we related some of the last relevant European Projects. Among the Spanish related projects, we note the project "Play the Game: gamification and healthy habits in physical education". Hernando et al (2015) [42] have studied the impact of the gamification as learning strategy in PE subject at school. The study has been designed as a didactic unit named "Play the game" where students of three secondary schools has to achieve a healthy cardiac frequency in their physical activity through different challenges, levels, points, leaderboards and badges. Therefore, Play the game has innovative elements, such as personalization, cooperation, emotions, technologies and a combination of formal and informal contexts. The results of Play

TABLE II. European Projects for PE

	Description
PEGASO	Uses wearable technologies and games to encourage healthy lifestyles amongst teenagers [20] shows a case study on the use of Gamification strategies and wearable lifestyle technologies for personal health management, in the framework of the European Project PEGASO. This author describes the results of two-year project exploring the potential of various lifestyle tracking and health monitoring equipment and the impact on the health parameters and well-being. The data captured and visualized by the mobile applications linked to these lifestyle technologies illustrates how gamification and enabling technologies have evolved in support of pervasive personal health management. URL: http://www.pegaso4f4.eu/
DOREMI	The main goal of DOREMI project is the "active ageing" based on three keys: healthy eating, active lifestyle and social interaction. The decline of cognitive ability is strongly related to lifestyle, as well as social engagement, cognitive stimulation, nutrition and physical activity. DOREMI is focused on social inclusion with the help of cognitive games and the development of a social and gamified environment. URL: http://www.doremi-fp7.eu/project
BEACONING	The project is named "Breaking Educational Barriers with Contextualized, Pervasive and Gameful Learnin". The 'anytime anywhere' learning concept is developing through pervasive, context-aware and gamified techniques and technologies, framed under the Problem-Based Learning approach. The project will create a platform as ubiquitous solution based on context aware systems, procedural content generation, pedagogy-driven gamification, learning analytics and cloud technology. URL: http://beaconing.eu/
ProsocialLearn	The project "Gamification of Prosocial Learning for Increased Youth Inclusion and Academic Achievement" is founded on the hypothesis that children at risk of social exclusion, lacking empathy and showing high levels of aggressive or anti-social behaviors, should benefit from digital games tailored to teach prosocial skills (the ability to identify the benefits of cooperation, recognize the emotions and needs of others and express trustworthiness) that can help them achieve academically, appreciate team work and recognize the value of understanding other people's needs. The gamification of prosocial learning will be driven by a set of well-defined prosocial learning objectives that are designed for the development of specific prosocial skills, in terms of prosocial theory, gameplay and game mechanics. Moreover, ProsocialLearn will create a new ecosystem through a new market for digital games, designed to support learning and development of prosocial skills, and offering games developers scientifically proven prosocial game elements for development digital games. An application programming interface (API), ProsocialAPI, will allow developers to integrate functions into games including visual sensing, identification of prosocial signals from in-game actions, personalized adaptation of game elements, player profiles, game mechanics and expressive virtual characters, and support for data collection with protection of personal data. URL: http://prosociallearn.eu/
LEGEND	The main goal of this project is to create an ecosystem of tools and apps to key agents of the sport sector. As Google does with its platform but with the sport as the soul, the core of LEGEND platform is Search but then it provides with dozens of tools. With this tool, public (local government, Universities or Schools) and private (sport facilities, individuals, federations, associations, etc.) managers from all around the world can send real time alerts, manage the classifications in an easy way, generate automatic calendars, publish the rules and news, define roles as referees or team managers and many more features to come. Also they have developed mobile apps for following the league but also the users can organize their friendly matches, comment on them before and after the match, check the results of their friends and to share them in other social networks. So, to motivate people to report the results, the project added an algorithm to calculate their skill level of every sport and also added gamification technics to the app so users can get badges depending on their results. Finally, users can track and check all their stats and historical results. URL: https://ec.europa.eu/easme/en/sme/5462/legend-sport-goes-viral

the game show the potential of gamification as an emergent learning strategy in PE because increase the motivation and promote the healthy habits in the students.

Other relevant project on gamification in PE is "ExpandeEF" or Expanded Physical Education, work developed by Lucía Quintero (2017) [43]. ExpandeEF was applied during one academic course with students of the second year in a high school of Tenerife (Canary Islands, Spain). Quintero (2017) used not only gamification strategies to design the experience, but emergent didactic methodologies such as mobile learning, flipped classroom or service-learning was applied (Fig. 1).

Fig. 1. ExpandeEF project. Source: http://bit.ly/2FNItIx

Also, several related projects have been developed by our research group, such as:

- VIDEM (Developing healthy habits and physical education through active educational games for hospitalized children and adolescents) funded by the Ministry of Science and Innovation, Ref. EDU 2010-10010, had the main goal of developing healthy habits through motor games and active video games in hospital classrooms. Among the objectives of the project, there are: a) Designing a model of educational intervention through physical exercise and ICT. The exercise is the transversal educational strategy, related attitudes and communicative values for integration of minors; b) Evaluating the influence of physical activity with learning games and motor play in learning healthy habits. Besides, training interventions and effectiveness of game models and tools applied are valued. In this context, it has been made and validated an integrated educational program formed by motor games and active videogames for the development of healthy lifestyles at a primary school [32]. URL: http://videm.es/
- SALUD-in (Platform for Interactive Virtual Rehabilitation with Physical Social Games for Health and Techniques of Natural Interaction) Ref PROID20100218, funded by the Canarian Agency for Research, Innovation and Information Society. This interactive platform, aimed at hospitalized children, allow the virtual-based rehabilitation based on physical social games for health and natural interaction techniques. It is based on multiplayer games, with games designed for physical and cognitive rehabilitation, a motion

capture system and biomedical data based on a low cost system (Kinect sensor and wearables devices). In this project TANGO:H (Tangible Goals platform) was created (Fig. 2). TANGO:H is a platform for hospitalized children with functional diversity developed by the Technological and Renewable Energy Institute (ITER) and the Interaction, Technology and Education Research Group (i-TED)) of the University of La Laguna. Further it comprises a clinical management system and remote monitoring of rehabilitation exercises and medical records of patients. URL: http://saludin.es/

Fig. 2. TANGO:H allow the design of personalized physical games in individual and collaborative modes. URL: http://tangoh.iter.es

- PROVITAO (Active videogames program for Outpatient Treatment of Obesity). The PROVITAO Ref OBE05 project, funded by the CajaCanarias Foundation (2014-2017), aims to support the treatment of obesity at early ages, contributing to improving the state health patients and preventing future disorders in adulthood [44]. It has a model of educational intervention designed for education in healthy habits, with an exercise program, motor games and commercial and own active video games, created in the research group, such as TANGO:H. The whole program is «gamified», in order to motivate and to achieve the engagement of children during the intervention in schools and home (one school year) (Fig. 3). URL: http://provitao.webs.ull.es

Fig. 3. Children using TANGO:H in the school under the PROVITAO project.

Finally, some example of the last gamified platforms or active games to develop healthy habits and can be used in PE are:

- SuperBetter: This platform increases resilience - the ability to stay strong, motivated and optimistic even in the face of difficult obstacles, making more capable of getting through any tough situation. URL: https://www.superbetter.com/
- Zombies, Run!: is a mobile game for running in which players has to run away from zombies. Then, while players run, stories are narrated, random sprinting to avoid zombies are launched. Players can collect items and be punctuated in personal music playlists. URL: https://zombiesrungame.com/
- Fitocracy: In this app, users work on exercise goals in group/community, and, also, can have a personal coach to keep motivated in fitness and nutrition. URL: https://www.fitocracy.com/
- VirZoom: this technology combines Virtual Reality (FR) and Fitness Technology. VirZOOM is a static bicycle connected to VR games. For example, the user can live different virtual experiences: power a horse in a race, a tank in battle, a fire-breathing dragon through mountains. URL: http://virzoom.com/

V. Trends in Gamification and Games For PE

A. Personalization

People have different ways to get fun. So, the research has identified different player types and motivations to play. Bartle (1996) [45] identified four player types: killer, achiever, socializer, and explorer. Regarding the motivations, Lazzaro (2004) [46] detected four motivational factors for playing games: hard fun, easy fun, altered state and people factor, and Yee (2006) [47] identified three main motivation components: achievement, social and immersion. So, the student model must represent the way people play, and the types of players. The personalization of game elements in the system [48] should take into account the forms of adaptation proposed by Kobsa et al. (1999) [49]: to user data, to usage data and to environment data. Besides, a typology of engaged behaviors to determine if a player is engaged or not has been proposed by Bouvier et al. (2013) [50]. Some research can help to understand the influence of environment data. For example, Cheng (2011) [51] tried to find the good moments to play at work.

Therefore, in gamification it is important to know how to motivate a particular and different person at the right moment using different types of motivations [52]. Thus, it is possible uses gamification strategies based on intrinsic motivation (inherent in the person, taken for its own sake or interest, for example, status, power, access to certain skills, or to contribute to a common good) or in extrinsic motivation (outside the person, made for reward or feedback). Social strategies can be used to, for example to compete, to collaborate or to compare achievements. In social games, there are collective mechanical equipment (projects, group scores, etc.) and other mechanical applied to the individual (motivation, positive reinforcement, etc.) [53].

Adaptation and personalization are concepts closely related and similar, which have a mutual goal: to offer a closer user experience by offering content close to the user, personalized to your interests and looking for increasing fidelity and satisfaction [53]. To perform this adaptation / personalization, the basic elements are: to define the user profile, to define the content and functionality that you want to adapt, and to define the interface elements that allow this adaptation / personalization. Personalization allows the adaptation of system through different techniques, such as content filtering or rule-based filtering, to infer the user's needs and preferences [54]. For personalization / adaptation of a gamified system, we must think about what are the features that make the system fun and if the system can work with or without these gamified features. We must also think about how these features relate to gamified different user profiles. Moreover, we must also consider whether the system can work independently to gamification without affecting the core functionality, which in our case is learning. For example, a leaderboard can be activated for the most competitive users, while not for others like introspective or special needs users. For the adaptation / personalization experience, the gamification engine must decide when and how specific and general features will be activated, taking into account: a) the student model (consisting of the user profile or static information and user history or dynamic information, and b) contextual information.

The static part of the student model or profile contains data such as age, gender, administrative information, learning style, type of player and preferences. Identifying the type and player preferences will increase the student motivation. The dynamics of the model student or history contains information of student interaction with the learning system and the state of their learning. However, a gamified system must also incorporate the trace of student interaction with the system for activation or deactivation of the functionality of gamification to increase the degree of engagement. Moreover, contextual information is essential in a gamification engine. The students can perform the activities from school, work or in their free time, in the classroom with their peers and with the teacher, or remotely. Student can also do the activities from a tablet, a mobile device, a laptop or desktop computer. All these contextual characteristics affect the gamified experience and the gamification engine must be able to adapt the features to different contexts. For example, if the activity is carried out in the classroom with teacher assistance, the chat cannot be very useful.

According to Gadiyar (2014) [55] "many gamification initiatives use points, badges and leaderboards as a way to motivate and incent participants to alter their behavior and use analytics to measure and monitor users' actions and social components to increase the user motivation", but, most of them fail to keep the user involved over the long term. The proposed solution for this problem is the personalization of the entire gamification process.

Gamification techniques should try to understand users, their personality, feelings, behaviors and actions. Big data [56], behavioral insights and elements of psychology can be used in gamification to provide a better end-user experience. Thus, in a gamification experience, every feedback, message or response should relate to user characteristics and situation properly. Typical gamification approaches, includes PBL, Levels, Feedback, Reward and Recognition techniques. The social gamification includes social media, communities, Web 2.0 elements, and big data analytics. Next generation of gamification systems, includes the elements for a personalized and contextual experience, such as: behavior-based frameworks, mental models, neuroscience and big data analytics.

B. Technology for Gamification and Games for PE

Technology plays a central role in the lives of today's children and young people. So, the use of new technologies, apps and devices into schools, could offer more engaging physical activities and healthier lives to students. Technology should be the core of engagement strategies in PE.

Lister et al (2014) [57] establish that apps "represent a promising opportunity for getting people active and have received considerable attention but this has been at the expense of in-depth analysis of effectiveness" and if the applications are not developed properly, "they will end up in a common technology cycle of hype with the users' feelings of failure and frustration on technology". To promote positive and active user experience in apps, many apps uses gamification, but they only use the most convenient game elements and did not use the full potential of gamification to create a success gaming experience. So, Lister et al (2014) [57] have conducted an analysis of 132 most popular apps in markets and seems to agree with this criticism. Moreover, the authors studied if the apps addressed correctly the motivational components to produce a behavior. They found the apps ignore the individual ability to perform the behavior, being this the main issue to achieve long-term behavioral change. So, digital rewards like badges or points may not produce a long-term behavioral change. We believe that apps have potential in physical education, but the design of these apps needs put attention in to achieve a sustained change in behavior.

According to a report on active life style in young people, it is found that "today's children and adolescents live sedentary lives full of computers, video games and television" [58]. But, certain technology (i.e. wearables, smartphones) that can encourage children to perform outdoor exercise. In this sense, some organizations like Youth Sports Trust, highlighting "the need to include wearable technology and gamification in physical education classes in schools" [59]. Furthermore, they said "in order to get children active from a young age, a more holistic approach to PE is needed, one which integrates technology and the delivery of a seamless, intuitive and digitally enhanced form of physical activity" [58]. Therefore, smart textiles can be introduced in the schools, as PE uniforms, and can take an important role in PE. For example, there are biosignal-monitoring underlayers produced by Athos, which read muscle effort, heart and breathing rates, analyses this information and push recommendations through mobile devices [58].

But tracked data on the activities and statistical information cannot be enough to transform this data into knowledge, skills, attitudes and behaviors [59], motivating and enhancing physical activity. Coaching support in physical activity can be an important key of success in PE. Moreover, school PE teachers will need skills and resources to offer a diverse set of PE activities focused on health, fitness and emotional wellbeing, and supported by technologies. Digital literacy should start at early ages and PE teachers must be put attention at this area. To use effectively technologies teachers requires a specific knowledge, but most primary school teachers tend to be generalists. And, to integrate the technology with the PE in a transparent and intuitive way it is needed to have a holistic approach.

VI. Conclusions

In this paper we presented a review about different studies on games and gamification applied to physical exercise, specially focused on the promotion healthy habits. We also present a review about the concept of "eSport" and its relationship with physical exercise and some considerations to be include in physical education programs. We found that many studies on active games or gamified physical exercise has been focused on energy expenditure and motivation.

Although, we found several studies and related project about how to use interactive technologies to promote healthy habits, most of games and gamified programs fail to keep the user involved over the long term. So, we believe that providing a more personalized experience can solve the problem of the engagement in long term. Personalization and emerging technologies (big data, wearables, smart technologies, etc.) based games and gamification for physical activity promising opportunity for getting people active.

Finally, we note that active videogames and gamification can be used in educational programs to increase the motivation of children and adolescents in physical exercise. In this sense, the nursing profession can play a fundamental role in health education. So, the educational programs to promote physical exercise and healthy habits should be designed and developed from the Primary Health Attention Centers and, also, in schools. And, the introduction of gamification into these educational programs (i.e. PROVITAO), can improve the quality of life of children who suffer from childhood obesity thanks to the acquisition of healthy habits.

Acknowledgment

This work has been supported partially by Fundación Caja Canarias, Project PROVITAO OBE05.

References

[1] Pearce, M. et al. (2011), "Who children spend time with after school: associations with objectively recorded indoor and outdoor physical activity", international journal of behavioral nutrition and Physical Activity, 11:45 (2011).

[2] Cook, W. (2013) Training Today: 5 Gamification Pitfalls. Training Magazine. Retrieved from: http://www.trainingmag.com/content/training-today-5-gamification-pitfall

[3] Schell J. (2008). The Art of Game Design: A book of lenses. Taylor & Francis US.

[4] Zagal, J., Mateas, M., Fernandez-Vara, C., Hochhalter, B. & Lichti, N. (2005) «Towards an Ontological Language for Game Analysis». In Proceedings of the Digital Interactive Games Research Association Conference (DiGRA 2005), Vancouver B.C.

[5] Zichermann G. & Cunningham C. (2011). Gamification by Design: Implementing game mechanics in web and mobile apps. O'Reilly Media, Inc.

[6] Simões J., Díaz Redondo R. & Fernández Vilas A. (2012). A social gamification framework for a K-6 learning platform, Computers in Human Behavior.

[7] Werbach, K. (2012). For the Win: How Game Thinking Can Revolutionize Your Business. Wharton: Wharton Digital Press.

[8] Nacke, L. E., & Deterding, S. (2017). The maturing of gamification research. Computers in Human Behavior, 71, 450-454.

[9] Mora, A., Riera, D., Gonzalez, C., & Arnedo-Moreno, J. (2015, September). A literature review of gamification design frameworks. In Games and virtual worlds for serious applications (VS-Games), 2015 7th international conference on (pp. 1-8). IEEE

[10] Papastergiou, M. (2009). Exploring the potential of computer and video games for health and physical education: A literature review. Computers & Education, 53(3), 603-622.

[11] Sánchez López AM, Aguilar Cordero MJ, González Jiménez E, Padilla López CA, Álvarez Ferre J & Ocete Hita E. (2011). La Obesidad como factor pronóstico de la falta de motivación en el niño y en el adolescente. X Congreso Nacional de la SEEDO. Revista Española de la Obesidad. 9(2), 99.

[12] Tan B, Aziz AR, Chua K, & Teh KC. (2002). Aerobic demands of the dance simulation game. Int J Sports Med. 23(2), 125–129.

[13] Cantallops Ramón J, Ponseti Verdaguer FJ, Vidal Conti J, Borràs Rotger PA & Palou Sampol P. (2012). Adolescencia, sedentarismo y sobrepeso: análisis en función de variables sociopersonales de los padres y del tipo de deporte practicado por los hijos. Retos. Nuevas tendencias en Educación Física, Deporte y Recreación. 21, 5-8

[14] Borràs Rotger PA, Vidal Conti J & Ponseti Verdaguer X. (2008). Escuelas promotoras de salud; intervenciones y estrategias para incrementar el nivel de actividad física: revisión y recomendaciones. Libro de ponencias de la IV Congreso Internacional y XXV Nacional de Educación Física. Córdoba: Universidad de Córdoba.

[15] Maddison R, Mhurchu CN, Jull A, Jiang Y, Prapavessis H & Rodgers A. (2007). Energy expended playing video console games: an opportunity to increase children's physical activity?. Pediatr Exerc Sci. 19(3), 334–343.

[16] Chin A Paw MJ, Jacobs WM, Vaessen EP, Titze S, & van Mechelen W. (2008). The motivation of children to play an active video game. J Sci Med Spor. 11(2), 163-166.

[17] Lanningham-Foster L, Jensen TB, Foster RC, Redmond AB, Walker BA & Heinz D. (2006). Energy expenditure of sedentary screen time compared with active screen time for children. Pediatr.. 118(6), 1831-1835.

[18] Lanningham-Foster L , Foster RC , McCrady SK , Jensen TB , Mitre N & Levine JA. (2009) Activity-promoting videogames and increased energy expenditure. J Pediatr. 154(6), 819-823.

[19] Mellecker RR & McManus AM. (2008). Energy expenditure and cardiovascular responses to seated and active gaming in children. Arch Pediatr Adolesc Med. 162(9), 886-891.

[20] Maddison R, Mhurchu CN, Jull A, Jiang Y, Prapavessis H & Rodgers A. (2007). Energy expended playing video console games: an opportunity to increase children's physical activity?. Pediatr Exerc Sci. 19(3), 334–343.

[21] Wetzsleon RJ, Swanson KS & Pickett K. (2008). Energy expenditure and ground reaction forces of an active video game, Dance Dance Revolution, in healthy weight and overweight children. Med Sci Sports Exerc. 40(5), S255.

[22] Wortley D. (2015). Gamification and Lifestyle Technologies for Personal Health Management. Europea Conference on Game Based Learning (ECGBL 2015). Steinkjer, Norway 8-9 October 2015.

[23] Hourcade, J.P (2015). Child Computer Interaction. First Edition. Retrieved on 01 July from: http://homepage.cs.uiowa.edu/~hourcade/book/child-computer-interaction-first-edition.pdf

[24] ExergameFitness.com Retrieved on March 15, 2017 from: http://www.exergamefitness.com/

[25] Magielse, R. and Markopoulos, P. (2009). HeartBeat: an outdoor pervasive game for children. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). ACM, New York, NY, USA, 2181-2184.

[26] Lund, H. H., Klitbo, T., & Jessen, C. (2005). Playware technology for physically activating play. Artificial life and Robotics, 9(4), 165-174.

[27] Sturm, J., Bekker, T., Groenendaal, B., Wesselink, R. and Eggen, B. (2008). Key issues for the successful design of an intelligent, interactive playground. In Proceedings of the 7th international conference on Interaction design and children (IDC '08). ACM, New York, NY, USA, 258-265.

[28] Seitinger, S. (2009). Designing for spatial competence. In Proceedings of the 8th International Conference on Interaction Design and Children (IDC '09). ACM, New York, NY, USA, 123-130.

[29] González C.S. & Navarro V. (2015). A Structural Theoretical Framework Based on Motor Play to Categorize and Analyze Active Video Games. Games and Culture. 1555412015576613, first published on March 26, 2015 as doi:10.1177/1555412015576613

[30] Padilla N., Collazos C. A., Gutierrez F., Medina N. (2012). Videojuegos educativos: teorías y propuestas para el aprendizaje en grupo. Ciencia e Ingeniería Neogranadina, VOL 22-1, PP 139 - 150, Bogotá.

[31] Collazos C. A., Gonález C.S., Gutierrez F. (2014). Patterns for Monitoring and Evaluating Collaborative Learning Processes in Videogames. LAP Lambert Academic Publishing.

[32] González C.S., Gómez N., Navarro V., Cairós M.,Quirce C., Toledo P., Marrero-Gordillo N. (2016). Learning healthy lifestyles through active videogames, motor games and the gamification of educational activities, Computers in Human Behavior, Volume 55, Part A, February 2016, Pages 529-551, ISSN 0747-5632, http://dx.doi.org/10.1016/j.chb.2015.08.052

[33] Rosell Llorens, M. (2017). eSport Gaming: The Rise of a New Sports Practice. Sport, Ethics and Philosophy, 11(4), 464-476.

[34] Witkowski, E. 2009. Probing the sportiness of eSports. In eSports yearbook 2009, edited by J. Christophers and T. Scholz. Norderstedt: Books on Demand GmbH: 53–56.

[35] Witkowski, E. 2012. On the digital playing field: How we "Do Sport" with networked computer games. Games and Culture 7 (5): 349–374.10.1177/1555412012454222

[36] Whalen, S.J. 2013. Cyberathletes' lived experience of video game tournaments. PhD diss., University of Tennessee, MIT Press.

[37] Tamboer, J.W.I. 1992. Sport and motor actions. Journal of the Philosophy of Sport 19: 31–45.10.1080/00948705.1992.9714493

[38] IESF (2018). URL: https://www.ie-sf.org/

[39] Hilvoorde, I. V., & Pot, N. (2016). Embodiment and fundamental motor skills in eSports. Sport, Ethics and Philosophy, 10(1), 14-27.

[40] Hutchins, B. 2008. Signs of meta-change in second modernity: The growth of e-sport and the World Cyber Games. New Media & Society 10 (6): 851–869.10.1177/1461444808096248

[41] Wagner, M.G. 2006. On the scientific relevance of eSport. In Proceedings of the 2006 International Conference on Internet Computing and Conference on Computer Game Development, 26–29 June 2006. Las Vegas, NV: CSREA Press: 437–440.

[42] Monguillot, M., González, C., Zurita, C., Almirall, Ll., and Guitert, M. (2015). Play the Game: gamification and healthy habits in physical education. Apunts. Educación Física y Deportes 2015, n.º 119, 1.er trimestre (enero-marzo), pp. 71-79. ISSN-1577-4015.

[43] Quintero González, L. E. (2017). La gamificación estática versus dinámica: una experiencia de aula a través de una pedagogía lúdica. ExpandEF. In CIVE'17 (Congreso Internacional de Videojuegos y Educación). Retrieved from: http://bit.ly/2FNItIx

[44] Gómez del Río, N., Martín González, R., González González, C. S., Navarro Adelantado, V., Toledo Delgado, P. A., Navarro Campillo, L., ...

& Quirce González, C. (2017). PROVITAO: a research program based on active games for help the abulatory treatment of childhood obesity. In 1st Workshop on Gamification and Games for Learning (GamiLearn'17). Universidad de La Laguna.

[45] Bartle, R. (1996). Hearts, Clubs, Diamonds, Spades: Players Who Suit MUDs. Journal of MUD Research 1, 1.

[46] Lazzaro, N. (2004). Why we play games: 4 keys to more emotion. Retrieved from http://www.xeodesign.com/xeodesign_whyweplaygames.pdf.

[47] Yee, N. (2006). Motivations for play in online games. CyberPsychology & Behavior 9, 772–775.

[48] Monterrat B., Lavoué E & George S. (2013). Toward Personalised Gamification for Learning Environments. 4th Workshop on Motivational and Affective Aspects in Technology Enhanced Learning (MATEL 2013).

[49] Kobsa, A., Koenemann, J. & Pohl, W. (1999). Personalized hypermedia presentation techniques for improving online customer relationships. Technical report No. 66 GMD, German National Research Center for Information Technology, St. Augustin, Germany.

[50] Bouvier, P., Lavoué, É., Sehaba, K. & George, S. (2013). Identifying Learner's Engagement in Learning Games: a Qualitative Approach based on Learner's Traces of Interaction. 5th International Conference on Computer Supported Education, Aachen, Germany.

[51] Cheng, L., Shami, S., Dugan C., Muller, M., DiMicco, J., Patterson, J., Rohall S., Sempere A. & Geyer W. (2011). Finding moments of play at work, in: CHI 2011 Workshop (2–5). Vancouver: Canada.

[52] Tondello G., Wehbe R., Nacke L. (2015). Towards a Personalized Playful Digital Wellness Assistant. In proceedings Workshop on Personalization in Serious and Persuasive Games and Gamified Interactions, London, UK, 2015.

[53] González C.S., Toledo P., Muñoz V. (2015). Enhancing the Engagement of Intelligent Tutorial Systems through Personalization of Gamification. International Journal of Engineering Education (IJEE). Vol. 32-1. Pp. 532-541.

[54] Uchyigit G. & Ma M. (2008). Personalization Techniques and Recommender Systems. Series in Machine Perception and Artificial Intelligence. Word Scientific Publishing. Singapure.

[55] Gadiyar A. R. (2014). Gamification 3.0: The Power of Personalization. White paper. Cognizant's Global Technology.

[56] Secades, V. A., and O. Arranz. (2016). Big Data & eLearning: A Binomial to the Future of the Knowledge Society, International Journal of Interactive Multimedia and Artificial Intelligence 3(6): 29-33. doi: 10.9781/ijimai.2016.364

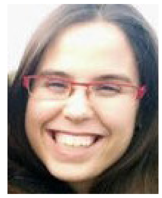
[57] Lister C, West JH, Cannon B, Sax T, Brodegard D (2014). Just a Fad? Gamification in Health and Fitness Apps. JMIR Serious Games 2014;2(2):e9. URL: http://games.jmir.org/2014/2/e9. DOI: 10.2196/games.3413

[58] Future Foundation: Special reports. The Class of 2035. Promoting a brighter and more active future for the youth of tomorrow (2015). URL: http://futurefoundation.net/wp-content/uploads/2015/06/the_class_of_2035_report.pdf

[59] Bhattacharya, S., and S. Nath. (2016). Intelligent e-Learning Systems: An Educational Paradigm Shift, International Journal of Interactive Multimedia and Artificial Intelligence 4(2): 83-88. doi:10.9781/ijimai.2016.4212

Carina S. González González

Dr. Carina Gonzalez is an Associate Professor at the Department of Computer Engineering and Systems at the University of La Laguna (ULL). PhD in Computer Science, specialized in AI and HCI techniques. Also, she is the head of Interaction, Technology and Education (ITED) research group at Department of Computer Engineering and Systems (ULL). For over 20 years she has focused her expertise in the field of the technologies applied to education and human computer interaction (intelligent tutoring systems, intelligent interfaces, human centered design, UX, serious games, gamification, eLearning, digital culture). She has published widely on these topics.



Nazaret Gómez del Río

Nazaret Gómez del Río is a PhD student in the doctoral Education Program at the University of La Laguna. She is Nurse and has the Master's Degree in Research, Management and Quality in Health Care, with specialization in education. Her main research line is the design and development of gamified educational programs using technologies and active videogames to prevent obesity in children.

Vicente Navarro Adelantado

Dr. Vicente Navarro Adelantado is an Associate Professor at the Department of Specific Didactics, Plastic and Corporal Expression, University of La Laguna, La Laguna, Spain. PhD in Physical Education, specialized in Teaching and learning in physical education and sports, motor play and active games. He has published widely on these topics

Planning and Allocation of Digital Learning Objects with Augmented Reality to Higher Education Students According to the VARK Model

Antonia Mireles Medina*, Francisco Javier Carrillo García, José Angel Montes Olguín

Instituto Tecnológico Superior Zacatecas Norte (México)

Received 3 November 2017 | Accepted 22 January 2018 | Published 9 February 2018

Abstract

In the present research, the authors propose the planning, assignment and use of digital learning objects with augmented reality according to the learning style of students in higher education, according to the VARK Model. It is found that students with treatment have had better results in their final grades than students who have not undergone the treatment of having used digital learning objects with augmented reality. The digital objects of learning (DLO's) with augmented reality designed according to the learning style of the students are an attractive and adequate option so that the teachers who are the main responsible for the didactic planning can spread the knowledge in the students. So that traditional forms of education are put aside and as a result of taking advantage of Information and Communication Technologies that have come to break with the paradigms that have prevailed for years in the teaching - learning process. On the other hand, education based on e-learning platforms facilitates the training of students at a distance allowing them to build and self-manage learning, as well as facilitate the dissemination of digital learning objects with augmented reality according to the learning style according to the VARK Model.

Keywords

Learning Digital Objects, VARK Model, Teaching-learning Process, Augmented Reality, E-learning.

DOI: 10.9781/ijimai.2018.02.005

I. Introduction

EDUCATION and the teaching-learning process evolved throughout history to allow the development of different models, techniques and tools that the teacher uses to transmit knowledge to their students. Students can acquire the necessary skills to function in the workplace, however, there are some factors that influence for a student fails a course, these factors may be related to: a) The student. Diversity in learning styles, IQ (Intelligenz-Quotient), interpersonal relationships, ability to work as a team, b) The teacher. Teaching style, didactic strategy, c) The external context. Family, social and economic situations, space physical conditions, among others. The present research studies the need to consider the student learning style as a factor to pass a subject.

In the teaching-learning process the student learning style must agree with the teacher's teaching style, on this topic the Instituto Tecnológico Superior Zacatecas Norte has a particular interest for improving their teaching methods and techniques.

In this study participated 18 students, the course is "Taller de Investigación II", they are in the seventh semester, this course belongs to the 2009-2010 study plan on the "Ingeniería en Administration" career. This course supports the student's certification process of the "Tecnológico Nacional de México", and allows the writing, presentation and defense of the research project [1].

To assign digital learning objects with augmented reality to support the teaching-learning process at the higher education students taking into account the student learning style, according to the VARK Model, is the general objective for this research. [2] mention, "The teacher is no longer the main information source, but rather a person who promotes the learning of cognitive and metacognitive strategies so that the student can convert for himself all that information that he receives in knowledge and knows how to apply it to different areas and situations ". They also comment that "work with digital learning objects facilitates and favors attention to diversity." Digital learning objects (DLO) have to be commensurate with their learning style, DLO's are designed based on the needs of each student.

DLO's according to [3] are characterized by presenting a series of audiovisual, interactive and dynamic activities that favor the development of student skills such as visualizing, guessing, analyzing, among others, these objects are designed based on the needs of each student. DLO's are small pedagogical elements that allow a great flexibility of access through the web [4]; they can also be accessed through local networks, desktop computers, mobile devices such as Tablet, Smartphone, among others.

The reason for proposing the use of DLO's is because an object of learning is a unit that has a meaning by itself and as such can be integrated into different contents or contexts of teaching and learning so sustain it [2]. The authors of this research agree with the authors that the objects of learning have value in themselves and can be used in the educational field with all freedom of time by the student, ie not necessarily in the classroom if not in your study space. Another reason is that learning objects as they say [5] may consider "interoperability that refers to the ability of two or more systems to exchange information,

* Corresponding author.

E-mail address: mirelesmed_7@hotmail.com

and then reuse that information" advantage that the present authors article emphasize, because ODAS can be reused as needed or the opportunity presented, this means that they can be overexploited in the sense that they can not only be used once.

For the purposes of this article, based on the learning styles proposed by the VARK model, it is a simple and appropriate instrument to identify teaching - learning styles; this instrument was developed by Neil Fleming in conjunction with Collen Mills in 1992 according to [6]. The model is called VARK (Visual, Aural, Read / Write, Kinesthetic) of the sensory modalities that they identified as Visual, Auditive, Reading and Kinesthetic or Kinesthetic, as it affirms [7]. Similarly, they say that people select information according to their interests or preferences to learn. According to González, [8], "each individual can present from one to the four modalities with all their combinations". However, in order to identify such modalities, the corresponding questionnaire must be applied to the following website: http://vark-learn.com/el-cuestionario-vark/. The questionnaire consists of 16 multiple choice questions, with four possible answers, which should only select one.

According to [6], the instrument for the identification of the learning style initially consisted of 13 questions with three and four possible answers, however for the year 2006 it was modified as already described in the previous paragraph. Below are described each of the sensorial modal preferences considered in the VARK model according to the [9]:

a) Visual (visual): preference for graphical and symbolic ways of representing the information.
b) Reading / writing: preference for printed information in the form of words
c) Auditive (aural): preference for listening the information.
d) Kinesthetic: perceptual preference related to the use of experience and practice, whether real or simulated.

It is crucial and vital that teachers at higher level institutions know the learning styles of their students, so that they can provide them with the right tools and thus achieve the acquisition of learning in a simple way. This is stated by [10]: "In educational processes it is important to take into account the characteristics associated with how students can learn, that is, their predominant learning style, their skills and participate in meaningful learning". Identifying learning styles in the student should become a daily task in those who are dedicated to facilitating knowledge. [11] refer to 38 instruments to identify learning styles. For the purposes of this research, the VARK model is used.

The teaching-learning process is considered as "the movement of the cognitive activity of the students under the direction of the teacher, towards the mastery of the knowledge, the skills, the habits and the formation of a scientific conception of the world" [12]. In this respect there is the constructivist theory, where Piaget, Ausubel and Vygotsky stand out; the first makes two contributions from the epistemological theory: knowledge as construction of schemes and levels of cognitive development. On the other hand, the contributions of the second author deal with significant learning and previous knowledge. Vygotsky is awarded as a key contribution to school education as a development context, the area of proximal development and the teacher as mediator, according to Betoret Teaching Strategies for Learning Styles [13]. This research rescues significant learning, previous knowledge and the teacher as mediator of knowledge.

Another concept to highlight in this article is augmented reality, which combines the real and virtual world so that it includes synthetic information to perceived real world images. Augmented reality must be interactive, real-time and 3D alignment, since virtual world information must be three-dimensional [14]. The same authors argue that augmented reality has its applications in medicine, manufacturing, entertainment and advertising, nevertheless it can be applied in the educational field. In the case of this research, it is also important to resume e-learning, which can reach a large number of students, can tailor e-learning activities tailored to individuals, and can provide help by instructors among colleagues [15].

II. Methodology

A. Description of the Context and Objects of Study

The research was carried out in a group of 18 students of seventh semester of Engineering in Administration of the "Instituto Tecnológico superior Zacatecas Norte" that study the subject "Taller de Investigación II", taking into account the topic, "desarrollo de la metodología del proyecto de investigación". Such topic contains the following subtopics: "2.1 Aplicación de los instrumentos y métodos experimentales seleccionados"; "2.2 Desarrollo de la metodología"; "2.3 Recolección y tratamiento de datos"; "2.4 Análisis de resultados"; "2.5 Propuesta de ajustes de parámetros de la investigación y/o del prototipo" [1].

B. Itinerary of the Actions Carried Out

The following is how the investigation of the use of digital learning objects with augmented reality according to the learning style of the students has been carried out.

1. Students have been asked to respond to the VARK questionnaire to identify their learning style.
2. At random, two subgroups of nine students were formed, each named group A and group B.
3. Following the sub-topics, a diagnostic test was developed that was applied to both groups to identify prior knowledge.
4. The two groups were given classes considering the sub-topics already mentioned. Group A has also been provided digital learning objects with augmented reality from each of the subtopics (it is specified that they have made use of them by means of Tablet and Smartphone according to their learning style). Group B was provided with DLO´s with augmented reality, in addition, the teacher of the subject has not taken into account their style of learning for the delivery of sub-topics.
5. Students have been asked to answer a final questionnaire and perform exercises to identify if they actually acquired knowledge and improved on the basis of the diagnostic questionnaire.
6. Results were tabulated in the diagnostic test, exercises and final questionnaire, to identify if there was an improvement in their learning process when using LDO´s with augmented reality.

C. Student Learning Styles

Fig. 1 shows the results of the learning styles of the seventh semester group. Fig. 2 shows the learning style of students who used LDO´s.

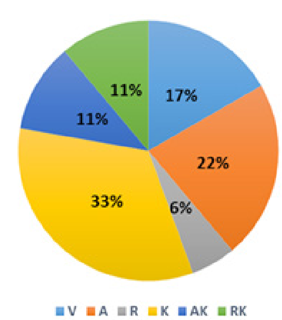


Fig. 1. Percentage of learning styles of the 18 students.

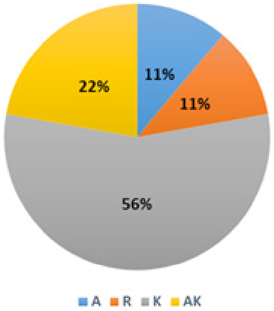


Fig. 2. Percentage of student learning styles that used LDO´s.

As can be seen in Fig. 1, 33% (6 students) are kinesthetic (K), while 22% (4) are auditory (A). The 17% (3) are visual (V). According to [8], these students are "unimodal". It can also be seen that 11% of students who are bimodal (AK) and (RK) coincide, that is, they stand out with two learning styles.

In Fig. 2, it can be seen that 56% have a kinesthetic learning style (K). The 22% is bimodal, since they have the preference to learn auditory (A)-kynesthetic (K). Finally, they coincide in an 11% reader / writer (R) and auditory (A). These students have been provided with LDO´s with augmented reality. In Fig. 3, it can be seen that 34% have a visual learning style (V). 33% are auditory, 22% are reader (R) - kynesthetic (K), that is, the latter corresponds to a bimodal learning style.

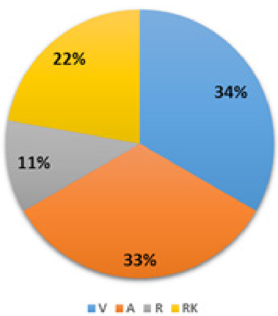


Fig. 3. Percentage of learning styles of students who did not use LDO´s.

It is necessary to consider the recommendations of [16], which are shown in Table I. Based on these recommendations, students were assigned LDO´s with augmented reality. These learning objects have consisted of videos, audio, booklets in digital and printed on each of the subtopics.

TABLE I. Teaching Strategies [16]

Teaching Strategies for VARK Learning Styles			
Teaching strategies for visual style Use of: • Written Instructions • Conceptual maps • Diagrams, models, synoptic tables • Computer animations • Videos, transparencies, photographs and illustrations.	Teaching strategies for auditory style Use of: • Verbal Instructions • Repeat similar sounds • Audio • Debates, discussions and confrontations • Brainstorming • Read the same text with different reflection • Guided and commented reading.	Teaching strategies for reader/writer style Use of: • Write a minute • Literary compositions, journals, blogs and reports, • Elaboration summaries, reviews and synthesis of texts, • Review of peer texts.	Teaching strategies for kynesthetic style Use of: • Role-play and dramatization • Group dynamics that require sitting and standing • The chalkboard to solve problems • Manipulation of objects to explain phenomena • Gestures to accompany oral instructions.

D. Evaluation Results

The results of the evaluation are shown below. It is important to mention that the following criteria have been considered in the final grade of each student (CF): Written exam, product or practices, tasks, performance and attitude; these criteria are those proposed by the Instituto Tecnológico Superior Zacatecas Norte, on a scale of 0 to 100. The diagnostic test (DI) has been applied to students to identify prior knowledge. This test has been applied to both students who used DLO´s and did not use it.

TABLE II. Student Results Who Have Used LDO´s (Group A)

n	DI	CF
Student 1	60	100
Student 2	60	100
Student 3	80	100
Student 4	80	100
Student 5	30	90
Student 6	60	100
Student 7	20	80
Student 8	50	90
Student 9	40	100

TABLE III. Student Results Who Have Not Used LDO´s (Group B)

N	DI	CF
Student 11	30	100
Student 12	20	90
Student 13	100	100
Student 14	70	100
Student 15	50	85
Student 16	50	85
Student 17	10	57.5
Student 17	60	58.5
Student 18	50	80

Obtained the results of Tables II and III have been replaced in the corresponding equations by means of a logical operation to make the comparison between the students who used LDO´s and those who did not use LDO´s and in this way to check if their academic performance has been improved with the use of LDO´s.

III. Data Analysis

Once the data corresponding to the diagnostic examination (previous knowledge), exercises and final exam or questionnaire were tabulated, a comparison of the results between Group A and Group B was made to determine if there was actually an improvement in their learning process. For this the equation (1) has been designed:

$$S_{ODAS} = \frac{\sum_{i=0}^{n} CF_i}{n} - \frac{\sum_{i=0}^{n} DI_i}{n} \quad (1)$$

Where

S_{ODAS} refers to students who used LDO´s (Group A) and

CF is the final grade of each student

DI is the diagnostic test grade

n is the total number of students

$$S_{ODAS} = \frac{860}{9} - \frac{480}{9}$$

$$S_{ODAS} = 42.22$$

Likewise, for students who have been given the class without LDO´s (Group B) we have:

$$S_{sin\_ODAS} = \frac{\sum_{i=0}^{n} CF_i}{N} - \frac{\sum_{i=0}^{n} DI_i}{N} \quad (2)$$

Where S_{sin_ODAS} refers to students who did not use LDO´s and

CF is Final grade of each student

DI is Diagnostic test grade

N is Total number of students

$$S_{\sin\_ODAS} = \frac{756}{9} - \frac{440}{9}$$

$$S_{\sin\_ODAS} = 35.11$$

In order to determine if the use of LDO´s effectively supports and improves the academic performance of the student, the following logical operation is considered: $S_{ODAS} > S_{sin\_ODAS}$ S where: 42.22> 35.11, it is higher, therefore, students who made use of the LDO´s have a better performance or academic achievement.

IV. Hypothesis Testing

Likewise, for a better accuracy in the obtained results a hypothesis test is applied that is shown in Table IV.

TABLE IV. Hypothesis Testing

	100	100
Arithmetic average	95	82
Variance	57.1428571	269.5
Observations	8	8
Pearson correlation coefficient	0.80006709	
Hypothetical difference of means	0	
Degrees of freedom	7	
Statistic t	3.24909357	
P(T<=t) a tail	0.00703746	
Critical value of t (a tail)	1.89457861	
P(T<=t) two tails	0.01407491	
Critical value of t (two tails)	2.36462425	

The hypotheses are proposed, the null hypothesis against the alternative hypothesis: $H_0: U_1 = U_2$ e $H_1: U_1 > U_2$

Thus, t_c=3.25 is in the rejection zone of H_0, so it is accepted H_1, as shown in Fig. 4.

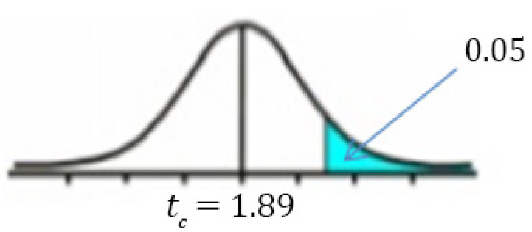


Fig. 4. Normal distribution curve.

V. Summary of Results

In this research work has been achieved the allocation and use of digital learning objects with augmented reality, to support the teaching-learning process of seventh-semester students of the "Ingeniería en Administration" career who have taken the subject of "Taller de Investigación II", considering the learning style of the student according to the VARK Model.

The authors recommend that the teacher be a mediator in the teaching-learning process and that their way of teaching is adapted to the student's way of learning, since this way the educational task can be facilitated.

Likewise, they propose the use of LDO´s in the different educational levels and in the different subjects, since these can be adapted according to the nature of each topic or theme. Likewise, they recommend that LDO´s be assigned to students in all topics of each subject.

Finally, it is important to mention that it is recommended to assign LDO´s according to the learning style according to the VARK Model. Because the VARK model has a structure for collecting data faster, this is due to the number of queries it contains. Without a doubt, the use of LDO´s makes the learning process more attractive and dynamic for students.

VI. Conclusion

In conclusion, the group with treatment that corresponds to Group A has better results in its final scores than Group B that did not undergo any treatment. The hypothesis of H_0, is rejected, whereas the hypothesis of H_1, is accepted, therefore, the use of LDO´s with augmented reality supports and as a consequence improves the learning process of students. The objective of the research has been achieved. Likewise, it is considered essential that teachers know the learning style of their students, so that they look for the way to provide them with the necessary tools in their daily tasks in the classroom.

Taking advantage of new technologies in education is a task that should not be overlooked in the educational field, since technologies are useful tools to support the teaching-learning process. The use of digital learning objects with augmented reality is a feasible option that can be taken up so that not only in institutions of higher level are implemented, but in all educational levels.

Planning and assigning digital learning objects is a task properly focused on the teacher who teaches the subjects, this depends on success for the student to achieve an acceptable academic performance, that is, sufficient for the acquisition of passing grades, skills and competencies.

The didactic planning, the assignment and the use of digital learning objects imply the organization of a set of ideas and activities that allow to develop an educational process with meaning, meaning and continuity [17], the authors of this article consider it of the utmost importance that in the process - teaching and learning the teacher as the main responsible for the didactic planning provides students with digital objects with augmented reality to achieve sufficient academic performance. On the other hand, digital learning objects with augmented reality have reusability as a characteristic [18], this allows them to be used whenever the student requires it and in the same way they can facilitate their learning process and that the LDO´s are assigned according to their learning style.

Finally, as stated by [19] the educational technology, gives high importance to the teaching work in its planning phase, so that based on the objectives, the contents are broken down, the learning activities and teaching resources are designed. They will lead to the expected behavior change and the evaluation is defined to corroborate the learning; the teacher must not forget his objective as facilitator of knowledge, so he must always keep in mind that a correct planning will lead him to achieve a better educational quality.

Acknowledgment

The authors of this article thank the "Instituto Tecnológico superior Zacatecas Norte" for the unconditional support to carry out this research.

References

[1] Tecnológico Nacional de México, «Asignatura Taller de investigación II, Clave ACA-0910, » México, 2016.

[2] I. C. Gordillo y I. Fernández Antelo, «Diseño y gestión educativa de objetos digitales de aprendizaje,» eUniverSALearning'08, 26-28 November 2008.

[3] W. d. Santos y P. Vedovatti, «El aula más allá de las paredes. Introducción a la tecnopedagogía mediante la integración de objetos digitales de aprendizaje, » 2015.

[4] M. C. López, R. Mora García, L. Rodríguez Valenzuela, E. Toledo Marhuenda, A. Jimenez Delgado y M. R. Pacheco Mateo, «Investigación en la mejora docente mediante Learning Objects,» pp. 3096-3111, January 2012.

[5] L. C. García, G. López Morteo, R. Morales Gamboa, J. Muñoz Arteaga, J. P. Contreras Paredes, J. J. Contreras Castillo, C. M. Curlango Rosas y R. Villa Angulo, «Diagnóstico de la capacidad para interoperar,» Norma Mexicana para la Interoperabilidad entre Entornos de Objetos de Aprendizaje, vol. 2, 2014.

[6] J. R. G. Nájera, «El modelo VARK: Instrumento diseñado para identificar estilos de enseñanza-aprendizaje,» INED, pp. 86-90, 2007.

[7] F. d. J. N. Cardenas, R. Hernández Palacios, V. T. Tomás Mariano y F. R. Felipe Redondo, «Identificación de Estilos de Aprendizaje en Alumnos Universitarios de Computación de la Huasteca Hidalguense mediante Técnicas de Minería de Datos, » Huejutla de Reyes, Hidalgo.

[8] B. González, C. Alonso y R. Rangel, «El modelo VARK y el diseño de cursos en línea,» Revista mexicana de bachillerato a distancia, pp. 96-103, 2011.

[9] M. Pedraza, «Los estilos de aprendizaje VARK UIS-Seminario de Orientación».

[10] E. L. O. Cañón, A. L. Sánchez Aradilla y A. Lozano Rodríguez, «REA y estilos de aprendizaje según Vark en el aprendizaje de matemáticas,» Revista Internacional Magisterio: Educación y Pedagogía, 64, pp. 91-93, 2013.

[11] J. L. G. Cué, J. A. Santizo Rincón y C. M. Alonso García, «Instrumentos de medición de estilos de aprendizaje,» Revista Estilos de Aprendizaje, pp. 2-23, 2009.

[12] K. H. Ortiz, Plataforma para el control del uso de software educativos, Cuba, 2009.

[13] F. D. Betoret, «La enseñanza y el aprendizaje situation educativa,» Aprendizaje y Desarrollo de la personalidad (SAP001), 2012.

[14] C. G. Morcillo, D. Vallejo Fernández, J. A. Albusac Jiménez y J. J. Castro Sánchez, Realidad Aumentada, Un Enfoque Práctico con ARToolKit y Blender, España: Bubok Publishing S.L., 2012.

[15] J. C. Almenara, «Bases pedagógicas del e-learning,» Revista de Universidad y Sociedad del Conocimiento (RUSC), Vol. 3, nº 1, April 2006.

[16] Lozano, 2001. [On line]. Available: http://www.cca.org.mx/profesores/cursos/cep21/modulo_2/modelo_vark.htm. [Last access: 03 Noviembre 2016].

[17] C. A. Peralta, «Adecuación de la Planeación Didáctica como Herramienta Docente en un Modelo Universitario Orientado al Aprendizaje,» Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, vol. 14, nº 3, pp. 109-130, 2016.

[18] L. Zorita, A. López Medina, M. Latorre, M. Blázquez, E. San Cristobal, S. Martín, G. Díaz y M. Castro, «Creación de objetos digitales de aprendizaje y su inclusión en el repositorio institucional espacio -UNED,» AIESAD, vol. 17, nº 1, pp. 149-177, 2014.

[19] A. Tejeda y M. E, «La planeación didáctica, » 2009. [On line]. Available: http://webcache.googleusercontent.com/search?q=cache:0ybCpUp1yrcJ:uiap.dgenp.unam.mx/apoyo_pedagogico/proforni/antologias/LA%2520PLANEACION%2520DIDACTICA.pdf+&cd=1&hl=es&ct=clnk&gl=mx. [Last access: 19 December 2017].

Antonia Mireles Medina

She was born in "Río Grande, Zac. México", on March 15, 1977. She is of Mexican Nationality.She graduated in 2000 with a degree in Informatics with a specialization in Information Systems at the "Instituto Tecnológico Superior Zacatecas Norte". She completed her studies corresponding to the Masters in Administration at the "Universidad Autónoma de Zacatecas". She has successfully completed her Bachelor of Laws studies in the "Universidad Autónoma de Durango, Zacatecas Campus". She has been teaching at the "Instituto Tecnológico Superior Zacatecas Norte" for 15 years. "Areas of interest are: Informatics, artificial intelligence, robotics, education and law.

Francisco Javier Carrillo García

He was born in "Río Grande, Zacatecas, México", on September 10, 1957. He is 59 years old. Gender Male, Mexican Nationality. He graduated as a Mechanical and Electrical Engineer at the Autonomous University of Mexico, where he worked as a Professor for 12 years. Since September 1994 he works as a Professor at "Instituto Tecnológico Superior Zacatecas Norte". In September 2013, he obtained a Master's Degree in Science in Educational Mathematics from the "Universidad Autónoma de Zacatecas". His areas of interest are Mathematics and Electromechanical Engineering and Mechatronics.

José Angel Montes Olguín

He was born in "Rio Grande, Zacatecas, México", on November 19, 1977. Mexican nationality. He graduated as a Bachelor in Computer Science at the "Instituto Tecnológico Superior Zacatecas Norte", where he worked as a Professor since 2005. His areas of interest are: Artificial Intelligence, Augmented Reality and Data Science, applied mainly to the areas of education and disability.

UC@MOOC's Effectiveness by Producing Open Educational Resources

Abdellah Idrissi J., Sofia Margoum, Rachid Bendaoud, Khalid Berrada*

Trans ERIE – Faculty of Sciences Semlalia – Cadi Ayyad University, Marrakech (Morocco)

Received 3 November 2017 | Accepted 22 January 2018 | Published 15 February 2018

Abstract

Open education is one of the most important settings in every society. It grants everyone the right to learn freely. Today, technology is helping to make learning even more open by providing an environment of online education, which plays a remarkable role in shortening distances and encouraging students to learn. At Cadi Ayyad University (UCA) the new-enrolled students are facing linguistics barriers as well as overcrowding in classrooms, in particular for those in open access institutions. Subsequently, they cannot have an easy access to their face-to-face courses. To help students to overcome these problems, the university has decided to design an online environment for all courses and programmes. The most innovative project adopted at UCA to face massification was inspired from the massive open online courses and was designed as an open Educational platform entitled UC@MOOC. More than 120 scripted courses have been posted online so far. In this paper we will describe and discuss an analytics research on geometrical optics course designed for around 2000 students at UCA. Through out this research we will explain how this initiative has been considered as a source of producing open educational resources.

Keywords

Open Education, E-Learning, MOOC, Blended Learning, Educational Platform.

DOI: 10.9781/ijimai.2018.02.007

I. Introduction

NOWADAYS, e-learning has become an unbroken part of our daily use. A vague definition of e-learning would be using the electronic components that an individual own or may have direct access to in order to access a knowledge resource remotely. These resources are widely spread over Internet forming a huge number of valleys which makes the access to them even more fast, easy and simple unlike searching in a physical library for example. Educational institutions are the most effective providers of organized and efficient online trainings since the most capable educators in the society form their human resources.

A few years ago, Cadi Ayyad University (UCA) in Morocco was offering a traditional learning to its learners in face-to-face mode, relying on direct exchange of information until 2013 when was the birth of UC@MOOC [1] project to respond to some serious big challenges.

In December 2016, Morocco declaration on Open Educational Resources was addressed to the Moroccan Government, education agencies, schools, middle schools, high schools, universities, the third sector, and all organizations and individuals involved in teaching and learning including galleries, libraries, archives and museums. This declaration was addressed in the frame of the OpenMed project, a focused project on development of OER, open frameworks for technology enhanced learning and massive open online courses (MOOCs), recognition of prior learning, adoption of Open Educational Practices [2]. As this declaration was a result of many initiatives of Cadi Ayyad University, it is carrying an educational responsibility to achieve some important goals [3] that are resulted in:

- A top-down and bottom-up implementation of OER.
- Supporting staff in using and integrating Open Practices and Open Resources.
- Collaborative creation in communities of practice, enhancing the quality of student learning.
- Licensing of OER content.

The open access establishments of UCA are having an increase of students surpassing the actual physical places each year [4] as shown in Fig. 1.

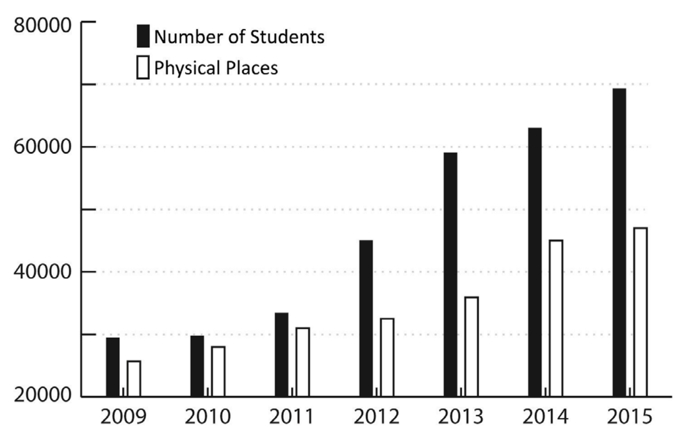


Fig. 1. Number of students and the provided places at UCA.

Providing places has become one of our university's top occupying issues that requires a satisfying resolution as well as for most of

* Corresponding author.

E-mail addresses: berrada@uca.ma

Moroccan universities. Because of massification (Total number of students in Morocco increase from 350.000 students in 2012 to around 850.000 in 2017), an interesting number of students are not able to be a part of the face-to-face learning position due to the insufficient places in classrooms. Even more, those who managed to have a place and sits far from the professor, they face the most common difficulties of receiving the information. We will try to summarize some of these face-to-face learning defects as follows:

- The disturbing noise generated by students that makes hearing the professor more difficult.
- Unclear and blurred sighting of the course projection on the wall and the explaining table.
- The interrupting questions to the professor during the course.

Fig. 2 is showing that for the last 3 academic years there was an increase of 15-20% of the entire enrolled students each year in the faculty of sciences Semlalia (FSSM) only. Moreover, there are 9 other similar open access establishments at Cadi Ayyad University.

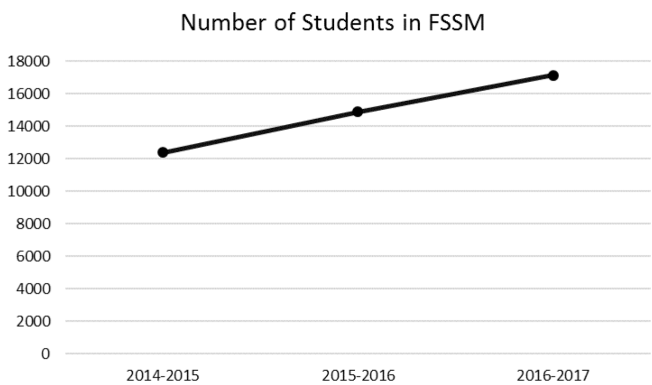


Fig. 2. Number of enrolled students at Faculty of Sciences Semlalia.

Along with massification, the new students aren't prepared to enter Higher education. Earlier in their studies everything has been teached by Arabic language except for foreign language materials (English, French, Spanish etc.). In one year, a sudden change from Arabic to French without any preparation that only increases the challenges that complicates learning process.

As a result, learners are not able of assimilating knowledge and which leads them to double the year or in the worst cases, abandon their studies. The percentage of students who abandoned their studies in the year 2015-2016 was calculated around 25% [5].

UC@MOOC is an original initiative of Cadi Ayyad University. Its main goal is to help reducing the effects of the massification and the other challenges by taking advantage of the revolutionary MOOC in the e-learning history. A simple emulation consists of producing the actual courses to open educational resources with a similar shape of MOOC. Basically, a professor presents his course in a digital studio by the same form respecting the given statement of work. The output is a scripted video put online with free open access to everyone.

Simply by transforming the face-to-face to an online environment using open educational resources (OER), students are now having a helpful escort to have online access to the OER presented by the same professors in face-to-face. This form of MOOC has some of MOOC's characteristics defined by Glance et al. [6]. This hybrid mode of education has proven its efficiency by combining the present and physical with the digital worlds. Moreover, a big number of viewing rates and viewing durations goes with an intensive presence of students in classrooms seems to complements one another [7].

These courses offered as an open educational resources are essentially meant to serve the students of the Cadi Ayyad University, but not only. However, the fact of being open they have attracted learners from other Moroccan universities and learners from the outside of Morocco. As part of Trans ERIE research team on innovation we have conducted this study to find out the added value of these online resources and their effects on UCA students.

II. Methodology

The access the courses of UC@MOOC whose main targeted learners are the students of the university can be done via 3 different tunnels as shown in Fig. 3.

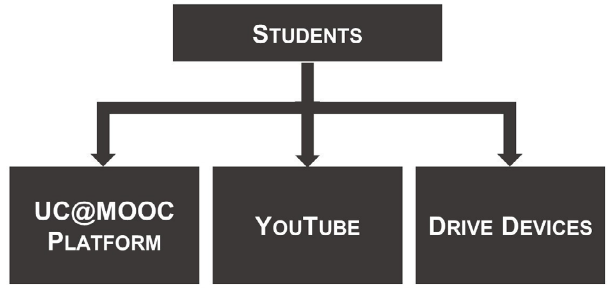


Fig. 3. Accessibility of OER UC@MOOC's.

The first one is through the official platform of UC@MOOC project [1] where all the scripted courses are not uploaded but being called and gathered from YouTube in one place organizing the resources' accessibility by providing many ways of filtering courses, search by establishment, disciplinary fields etc. [8]. It also offers to students some interactive learning activities. However, 10.6% of learners access this platform to learn from UC@MOOC's resources and the other 89.4% watches them from YouTube platform.

The second way of access is via YouTube knowing that it's the platform used to upload these resources, which are open so anyone can have direct access to these videos. YouTube also has another content organizing methods. Playlists are a good example of that where some of learners organize all the videos of a course in one link. YouTube can be described as a less professional platform for education because it is not meant for that and some of the irrelevant suggestions may redirect learners to some unrelated floods.

Inside of YouTube there are many routes to reach the videos of courses. Research inside YouTube is a less taken route to arrive at the courses, suggestions and playlists are the top traffic-driving areas in YouTube then comes the "others" category that combines the traffic from the platform and the shares in other places like the social media. This data is resumed in Fig. 4.

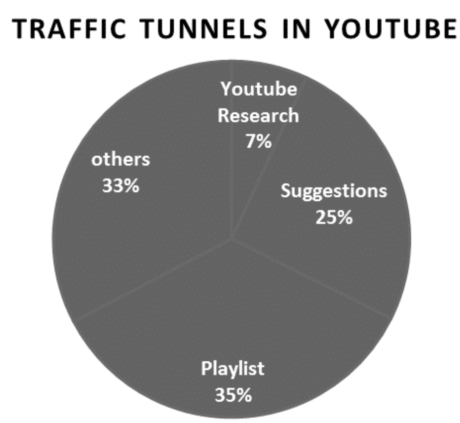


Fig. 4. Trafic Tunnels in Youtube.

The last resource provided by the university is DVDs. All these courses were burned on DVDs and put at their disposal inside the campus. As a result, students are able to share these courses between themselves offline using the different storage supports they have.

Our model for this study is the Geometrical Optics course. This course was the first recorded. It was scripted and put online in April 2013. It is divided in 6 chapters each one was encapsulated in one video except for chapter 4, which was divided in 2 videos for being long as presented in Table I. These videos were uploaded on YouTube on Nov-Oct 2013.

TABLE I.
Online & Face-to-face Duration of Geometrical Optics Course

	Chap 1	Chap 2	Chap 3	Chap 4.1 & 4.2		Chap 5	Chap 6
Duration	32:28	11:13	29:05	25:28	28:32	10:54	35:13
Total Online duration	2 hours 52 minutes 53 seconds						
Face-to-face duration	40 Hours						

As shown in Table I it is evident the face-to-face duration is much longer than the same course online duration. The professor in class gets interrupted frequently by questions from students then explaining one thought takes more time than the expected. When the professor presents the course in studio he takes enough and necessary time in order to mount the course properly, thus the recording time is far too small in comparison with presenting it before students. Another advantage has been shown, many professors found themselves practicing more pedagogy when using scripted videos of their courses. Some others asked to re-make new scripted courses after putting online their courses.

The other massification factors such as noise goes also against the regular course rate. Online, if the learner missed an idea in the video he can simply go back and replay the less-understood part.

Geometrical Optics course is being teached around the year during two semesters. In the first one it serves one curricula while in the second one it serves a different curricula. There is a slice difference between the programs of these trainings so the course has been produced online twice with adapted modification in every case. Online, there is a version of the same course for each training/semester.

In our study we will process the version of the course presented in the first semester where the students are new-graduates and they face this course right after their entry to higher education.

III. UC@MOOC Serving UCA Students

A. Providing an International Needs of OER

During the online life of this course until the end of October 2017 it recorded a significant number of views calculated from our UC@ MOOC YouTube Channel only.

The number of views reached 244898 and counting every second. Fig. 5 shows the variation of views each year. It is obvious that since the creation of the course in October 2013 until October 2017 the number of views has been increasing each year where 5858 views was counted in 2013 and 76679 views in 2015 alone. But in 2016 this number dropped to 67748 views. The reason behind this fall is some new channels on YouTube have reuploaded the course into their channels so the views of the course were separated between our official channel and the other channels.

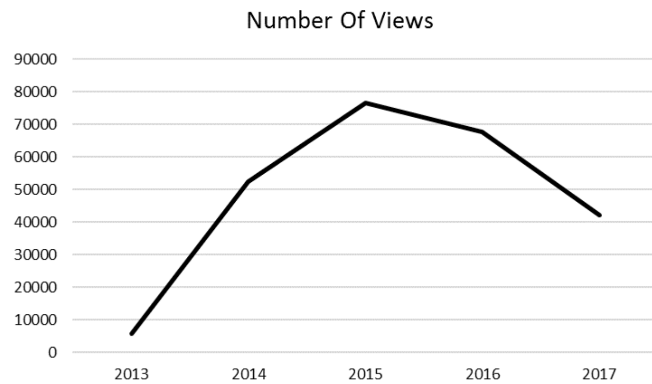


Fig. 5. Number of views in each year from Nov 2013 - Oct 2017.

Let's take a deeper look at the course's views variation in months and see what happens to the course monthly (see Fig. 6).

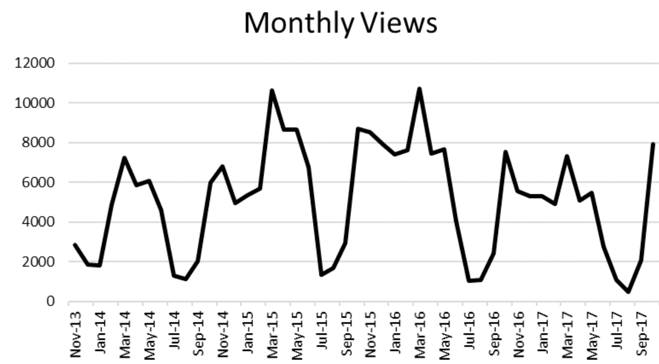


Fig. 6. Number of views from Oct. 2013 to Feb. 2017.

This figure shows that there is a resemblance of the shape of graph twice every year. Well, each year there is the start of two semesters where students try to learn online in conjunction with the faculty. A semester started along with the course after October 2015 where the entire views of that month were 1667. Suddenly this number grows up to 8698. Then when the exams week approach the numbers grows-up again in March to reach their higher limits (10735) in comparison with the rest of the months.

The only reasonable explanation that may accord to any logic is that students find something they need and they are interested in. To strengthen this idea, we'll prove that the number of views is also related to the viewing duration of the course.

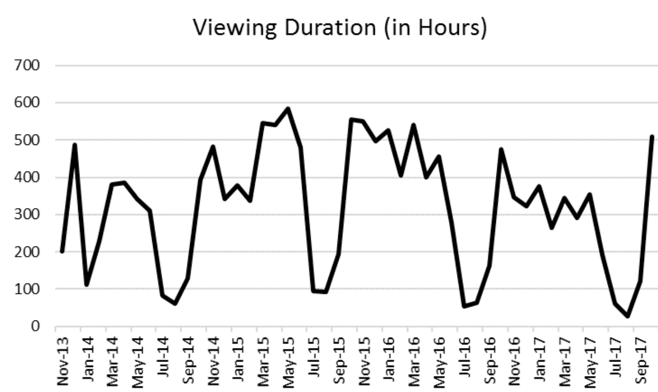


Fig. 7. Duration of views from Oct. 2013 to Feb. 2017.

The course's watching duration graph (Fig. 7) is a lot similar to the number of views graph concerning the semester's start and exams' week period. 2015 is the year that harvest the most viewing durations.

In May, the viewing duration in hours was too close to hit 600. In fact, 583 hours of views equals 24 days and a third. The duration of the entire course is about 2 hours and 53 minutes. Then we found that the whole course has been watched continuously 18 hours a day for 31 days straight.

The viewers of this course are not all Moroccans as shown in Fig. 8.

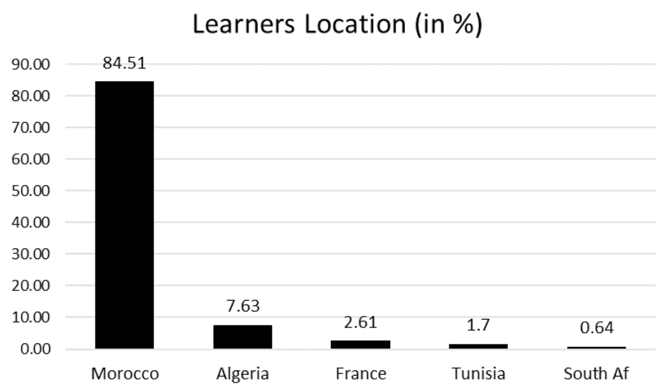


Fig. 8. The origin countries of viewers.

According to Fig. 8, viewers in Morocco contribute with 84.51% of the views and the 15.49% is coming from viewers from the outside of Morocco. This course's targeted learners are the students of the Cadi Ayyad University who, unwillingly, are not able to learn in the recommended and adequate face-to-face situations as mentioned before. These countries are Algeria, France, Tunisia and other south-African countries. Egypt for example is excluded from the equation knowing that they do have English language instead of French as their first foreign language. French is the first foreign language used in Morocco and in this course, which was the key factor that led the learners from these francophone countries to consider and learn using this course even if it wasn't meant for them [9,10].

These videos, being watched on YouTube until October 31st, 2017, have also sparked 823 likes, 98 comments, 1556 shares to social networks and triggered 585 new and unique subscribers to the channel. All of this results to the significant importance of this course as an effective online educational resource by students. Then, did it bring some benefits to student's curriculum? and how?

B. The Effectiveness of UC@MOOC

1) Successful Rate:

After this course was put online to serve the face-to-face training's students at the first place by helping them surpassing the major difficulties they encounter in campus. A positive effect has been projected in the successful rate of students during the years where this online course was put online.

The results that follows were gathered from the archive of the faculty where a lot of data is recorded. The judgement point of our study and how the online course helped is the results obtained by the students in the exams. And the course under our radar is still Geometrical Optics.

Firstly, to prove that the big part of viewers are students we'll present the data that shows viewing duration in comparison with age ranges as provided in Fig. 9 during the month of October 2017 only.

Surly young people are the ones whose viewing duration exceeds all the other ranges combined together. The new-graduates who studies in the first semester have their ages between 17-24 years. Only a small number of students are 16 years old. At this point, for sure students are watching the course [11].

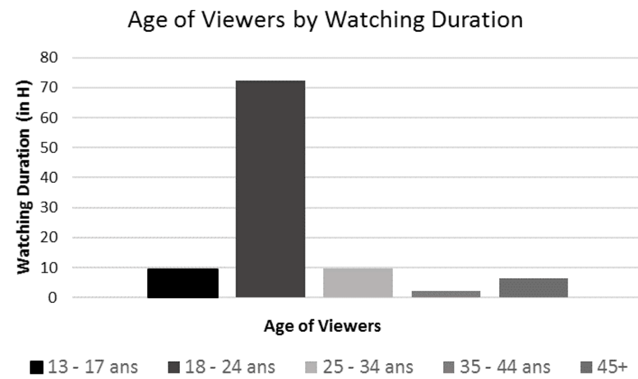


Fig. 9. The age of viewers related with viewing duration.

For the first semester of the academic year 2014-2015 a total of 2514 students were enrolled in the face-to-face course. This is the second year after the course was put online. Only 14.75% of them succeeded in this course. In the next academic year 2015-2016, the number of enlistees went up to 3023 students and success rate was 17.5%. As for the next and last academic year 2016-2017, the success percentage hit 17.8% knowing that the number of enrolled students (2155) was less than the previous year. All of this data is shown in Fig. 10.

From 14.5% to 17.8% of success rate gives a clear vision on the effectiveness of UC@MOOC initiative. Based on the data provided before, a bigger success rate is estimated to be observed in the current semester (First semester of academic year 2017-2018).

Now we have proven that there is a noticeable improvement on success rates of students. In fact, if this open online course helps at least 10 students to overcome defection and support them to succeed each year in the face-to-face course then the efforts provided for the purpose of creating this helpful course are worth investing time, competences and money [12].

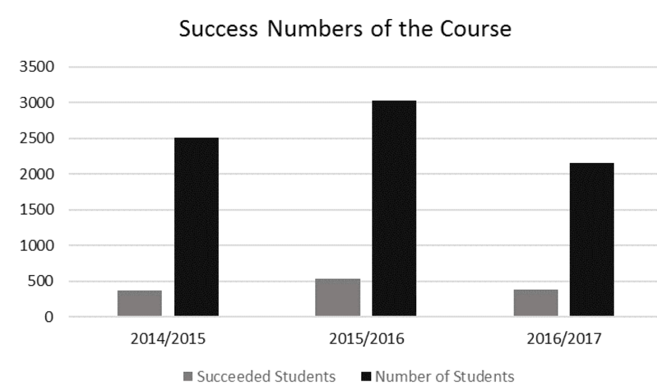


Fig. 10. Number of enrolled and succeeded students on Geometrical Optics Course.

2) Successful Rate of Non-Supported Online Courses:

Geometrical Optics is without a doubt can be called an online supported course. Other courses in the faculty starts parallelly with this course. Not all of them have an online version. We will process now the Mathematics course which is being teached until now in face-to-face only and there is no online version of it. The professors of this course haven't recorded it and it will be an interesting point to focus-on and its success rates [13].

Mathematics course has the same enrolled students as Geometrical Optics. They are being teached in the same time-frame and the exams are also taken in the same period and they serve the same trainings.

In the academic year 2015-2016, Mathematics course recorded a success rate of 14.35% whereas Geometrical Optics had 17.5%. The next year, the courses simultaneously recorded 16.8% and 17.8%. The fact that Geometrical Optics has higher success rates results to the fact that there is a result of providing OER and supporting students online. The graph in Fig. 11 presents the numbers.

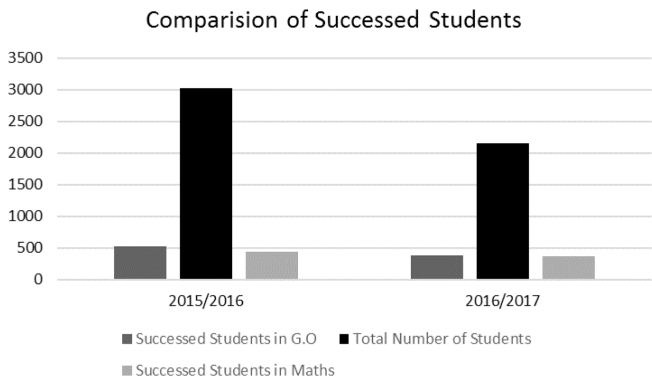


Fig. 11. The difference between the two courses in success numbers.

IV. Conclusion

The massification may have made learning in face-to-face difficult for some students but it may be a diverging point of creating a good or perhaps some better alternatives. As a result of massification we created UC@MOOC. This initiative proved to be a helpful and supportive online educational environment that is not serving the students of the university only, but keeps on providing online learning and knowledge to international learners in their proper context of requesting information. A genius idea, low cost material, a small number of staff and pedagogical researchers and professors are the major elements of building such a wonderful project.

The pedagogical innovation was the start of the route of creating UC@MOOC project. Research on education will be an unstoppable science as long as societies continue on development and invention. However, it is still only a supporting system to students and not an ultimate solution of massification.

The work on providing more open educational resources within the innovative initiative of UC@MOOC is still ongoing. The form of producing these OER will definitely change and thrive to higher levels of efficiency where online learning breaks the current barriers.

Acknowledgment

Authors would like to thank all the people involved in UC@MOOC platform for their engagement and for making this project happen at Cadi Ayyad University. Special thanks also to the OpenMed Erasmus plus consortium for considering this initiative as an innovative practice of Open education in the MENA region.

References

[1] Website of Cadi Ayyad University, Marrakech 2013, http://mooc.uca.ma

[2] OER Morocco Declaration https://www.change.org/p/universit%C3%A9s-d%C3%A9claration-du-maroc-sur-les-ressources-educatives-libres-oer-morocco-declaration?utm_source=embedded_petition_view

[3] Compendium Of Case Studies Of Open Education In The Mediterranean, retrieved from: http://openmedproject.eu/results/compendium/

[4] K. Berrada, A. Miraoui, La capacité d'accueil de l'université Cadi Ayyad face au problème de massification des étudiants, MENA Higher Education Conference, Beirut, Lebanon (May 2015). http://cmimarseille.org/sites/default/files/newsite/library/files/fr/Khalid%20Berrada.pdf

[5] D. Glance, M. Forsey, M. Riley M., "The pedagogical foundation of massive online courses", 2012, First Monday Journal, retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/4350/3673

[6] A. Idrissi. J, S. Machwate, R. Bendaoud, K. Berrada, A. Miraoui, UC@MOOC: Pedagogical innovation to challenges of massification at university level in Africa, ICCMIT, Poland, 2017.

[7] L. Guàrdia, M. Maina, A. Sangrá A, "MOOC Design Principles. A Pedagogical Approach from the Learner's Perspective", May 2013, retrieved from http://elearningeuropa.info/sites/default/files/asset/Indepth_33_4_0.pdf

[8] N. Mirriahi, S. Dawson S., "The Pairing of Lecture Recording Data with Assessment Scores: A Method of Discovering Pedagogical Impact", 2013, doi: 10.1145/2460296.2460331

[9] P. Hill, "Emerging Student Patterns in MOOCs: A (Revised) Graphical View", 2013, online blog post retrieved from http://mfeldstein.com/emerging-student-patterns-in-moocs-a-revisedgraphical-view/

[10] R. Kizilcec, C. Piech. E.Schneider, "Deconstructing Disengagement: Analyzing Learner Subpopulations in Massive Open Online Courses", 2013, retrieved from http://www.stanford.edu/~cpiech/bio/papers/deconstructingDisengagement.pdf

[11] N. Schroeder, O. Adesope, "A Case for the use of Pedagogical Agents in Online Learning Environments", 2012, retrieved from http://jotlt.indiana.edu/article/view/2126/3045

[12] M.P. Papazoglou, "Agent-oriented technology in support of ebusiness", 2001, Communications of the ACM, Vol. 44 No. 4, pp. 71-77.

[13] Commonwealth of Learning (COL), "Perspectives on Open and Distance Learning: Open Educational Resources: Innovation, Research and Practice", May 2013, retrieved from http://www.col.org/resources/publications/Pages/detail.aspx?PID=446

Abdellah Idrissi Jouicha

Abdellah Idrissi J. is a PhD student at Cadi Ayyad University. He is conducting research on pedagogical innovation, MOOC & OER at Trans ERIE group of research of Cadi Ayyad University.

Sofia Margoum

Sofia Magroum is a PhD student at Cadi Ayyad University. She is developing research on Micro-computer Based-Laboratory and OER at Trans ERIE group of research of Cadi Ayyad University.

Rachid Bendaoud

Rachid Bendaoud is professor of physics in charge of e-Learning at Cadi Ayyad University. He is working on blended learning and one of developers of Uc@MOOC initiative at Cadi Ayyad University.

Khalid Berrada

Khalid Berrada is professor of physics at Cadi Ayyad University. He is actively involved in basic research and education development for many years. He is in charge of pedagogical innovation at Cadi Ayyad University and UNESCO Chairholder on "Teaching physics by doing" since 2010. Currently, he is coordinating the Uc@mooc project and Director of the Centre for pedagogical innovation at Cadi Ayyad University.