Performance indicators in higher education with learning analytics

Carlos Morais
CIEC-University of Minho
Polytechnic Institute of Bragança
Bragança, Portugal
cmmm@ipb.pt

Paulo Alves
Polytechnic Institute of Bragança
Bragança, Portugal
palves@ipb.pt

Luísa Miranda
Polytechnic Institute of Bragança
Bragança, Portugal
lmiranda@ipb.pt

Abstract* — The regular use of virtual learning environments by educational institutions and the availability of large quantities of data regarding the teaching and learning process both challenge teachers, institutions and researchers to explore such data within the perspective of learning analytics in order to obtain indicators which might contribute to the improvement of the teaching and learning process. Consequently, such improvement may lead to students' better learning outcomes as well as a better relationship with the institution they attend. Taking into account that student retention and dropout from degree courses before their completion have highly negative consequences for a high number of students, the aim of this research is to look into indicators associated with such problems through a mixed methodology, both qualitative and quantitative, and through procedures of analysis and research of data from databases containing academic information provided by a sample of 1,588 undergraduates. The focus of the study lies upon the relation and analysis of the following variables: students' frequency of access per course unit to the virtual environment of the institution they attend; quantity of course units passed; and mean of the marks obtained in the course units passed. Among the results obtained, we highlight the existence of a positive degree of association with moderate correlation between the variables frequency of access per course unit to the virtual environment and the mean of the number of course units passed. Following the division of the sample into four different groups, each group corresponding to one quartile, resulting from the data regarding the frequency of accesses per course unit to the institution's virtual environment, significant differences were found concerning the mean of the course units passed between the groups showing a higher frequency of access to the virtual learning environment and those with a lower access frequency, to the advantage of the groups showing a higher frequency of access to the virtual environment. Therefore, considering the way the groups were constituted, there is proof that the higher the frequency of access to the virtual environment is, the higher the probability of students succeeding in the course units they attend will be.

Keywords - Learning analytics, access to the virtual environment per course unit, learning outcomes, education.

I. INTRODUCTION

Research is one of the best paths to knowledge building. The complexity of the human being makes it a challenge to conduct research into its features. In this case, we looked into the features of higher education students in order to provide strategies which can be accepted by the scientific community and might promote students' good learning outcomes. Therefore, considering the quality and quantity of information technologies and the diversity of data they allow to treat, we used learning analytics to obtain indicators which enable a better understanding of undergraduates' action within the teaching and learning process and to define, or redefine, strategies which may improve their performance within the academic context.

We highlight that the Society for Learning Analytics Research defines learning analytics as the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs [1]. Similarly, Tempelaar, Rienties, and Giesbers [2] refer that Learning analytics looks into improving learning processes through systematic measurements of data related to learning and providing feedback to students and teachers.

Since higher education institutions have, among other purposes, the mission of building and disseminating scientific knowledge, and since the authors of this study are higher education teachers with diverse and complementary specialisations, namely in the areas of engineering, education and technology, the challenge was set to identify the data regarding students' access to a virtual learning environment (VLE) and from there, taking into account the aims and concept of learning analytics, identify relationships and indicators which allow a better understanding of students' action as well as the prediction of situations which are not always identifiable within institutions' daily activity such as retention or dropout.

Higher education institutions are struggling with the need to make use of the large amount of data they have in their databases, especially in their virtual learning environments, for the improvement of their relationship with students and the adaptation of their strategies to their public target.

Considering learning analytics and the possibility of using a large amount of data, including data regarding access to the virtual environment and students' final marks in the various course units, new fields or research are appearing which might contribute to the definition of new educational policies and the prediction of situations of students tending towards retention or dropout before the completion of their degree courses.

As highlighted by Shacklock [3], the main reasons for introducing learning analytics are associated with preventing retention, providing feedback to students, obtaining data concerning students' attendance and improving teaching and learning.

One of the research questions guiding this study is the following: What is the degree of association between the

variables frequency of accesses to the virtual environment per course unit (Frequency_Accesses_VLE_CU) and the number of course units passed (N_CU_Passed)? In addition to this, it is also interesting to test the following research hypothesis: there are significant differences in the number of course units passed between the groups which register higher Frequency_Accesses_VLE_CU and the groups with lower Frequency_Accesses_VLE_CU.

Among the aims of this research work, we highlight:

- Identify the degree of association between the variable Frequency_Accesses_VLE_CU and the number of course units passed;
- Obtain indicators from the relationship between the Frequency_Accesses_VLE_CU and the number of course units passed as well as from the relationship between the Frequency_Accesses_VLE_CU and the mean of the marks obtained in the course units passed.

In this paper, we focus on the data regarding a sample of 1,588 students from a public Portuguese higher education institution concerning the 2015/2016 academic year. We present answers associated with questions related to the variables Frequency_Accesses_VLE_CU and learning outcomes in terms of course units which the students passed as well as the mean of the marks they obtained in those course units. The analysis of the relationship between the above variables may help define the profile of students' action within the teaching and learning process and consequently, predict implications to their learning outcomes.

The paper is organised into the following main topics: Learning analytics in undergraduates' performance, Method, Results and Conclusions, ending with the References. Hereinafter is the development of these topics.

II. LEARNING ANALYTICS IN UNDERGRADUATES' PERFORMANCE

Both students' performance regarding learning outcomes associated with the number of course units (CU) they pass and the marks obtained in those units have been the focus of numerous researchers in the field of learning analytics.

Conde & Hernández-García [4] state that learning analytics is an emerging and promising field for educational research and for learning processes supported by information technologies (IT). Agudo-Peregrina et al. [5] define learning analytics as the analysis of data regarding learning supported by IT which enables teachers and degree course coordinators to look for unobserved patterns and underlying information within learning processes, adding that its main goal is to contribute to the improvement of learning outcomes and of the global learning process in e-learning environments.

According to Johnson et al. [6], learning analytics uses web analytics tools applied to education in order to identify students' profile through a process of collecting and analysing individual interactions in the online learning activities. For the same authors, the aim of learning analytics is to build better pedagogies, promote active learning, reach the student population at risk and assess the factors affecting students' academic attainment.

Learning management systems (LMS) enable the record of users' actions, thus providing valuable information on the students' learning process, especially on students at risk of dropout or in need of additional support [7].

Within higher education, there is a high amount of data associated with LMS regarding undergraduates, their learning and the environments they use to study.

As highlighted by Shacklock [3], the relevance is not in the amount of data available but in what is done with it. Data analysis is a process in which data is collected and analysed in order to identify patterns, make predictions and provide support in the decision-making. The analysis of data associated with LMS may enable teachers to identify significant patterns, detect students at risk, provide proactive feedback and adjust teaching strategies, identify changes in students' behaviour and in the way they interact with their peers [8].

A study conducted at Maryland University, United States, showed that undergraduates obtaining low marks use the VLE 40% less than those with higher marks. The same authors also highlight that in another study conducted in California, the use of the VLE explained 25% of students' final classifications variation [9].

Every time a student interacts with any service of their institution, whether it is going to the library, accessing the virtual learning environment or submitting assignments online, a set of digital logs is left behind, which allows the identification of their profile features. Learning analytics is the process of using that data to improve the teaching and learning process [9]. It is a field which focuses on obtaining patterns or trends through the analysis of sets of data related to students or through huge sets of educational data to the development of customised learning systems [10].

The use of learning analytics in education has increased over the last years due to four main reasons: a substantial growth of the amount of data, improved data formats, advances in computing and higher sophistication of the tools available for analysis [11].

Learning analytics is a multidisciplinary field involving learning mediated by computer, artificial intelligence, databases, statistics and visualisation. Also, it contains the research fields regarding educational technology of data mining and adaptive learning systems [12]. This area may provide teachers with various information such as their students' previous classifications and progress, thus enabling them to identify those who are highly likely to drop out and therefore making it possible for teachers or the institution to act through personalised intervention near students at risk by means of advice or other support [9].

Analytic systems enable teachers to identify demotivated students at a much earlier stage, which gives them the possibility to intervene before the demotivation intensifies and therefore improve the possibilities of student attainment [3].

According to Rienties et al. [13], the wider availability of large amounts of data, powerful analysis mechanisms and perspectives of analytic results deftly projected mean that institutions can now more than ever before use past experience to create models of support to students as well as perceptive learning processes.

The importance of learning analytics applications is due to the fact that they give institutions and teachers the possibility to provide customised learning and opportunities to support their students' progress based on the data available about the students, their relationship with the learning process and with the institution they attend.

As an example, we highlight a study carried out by You [8] with data obtained from a sample of 530 undergraduates of an online course, where it is shown that a regular monitoring of the students through learning analytics allows the identification of the frequency of access to the course, to the support materials for the students and the submission of tasks proposed by teachers.

The use of leaning analytics within the exploration of virtual learning environments is also justified by the importance that these environments have had, mainly to teachers and students.

In a study conducted by Morais, Alves and Miranda (2013) [14], which involved a sample of 189 higher education teachers in the assessment of aspects related to their use of a virtual learning environment tools, the aspects most valued by teachers were the availability and access to digital resources, time saving and the possibility to improve communication between teacher and students. The same study also highlights that the most valued features of the digital resources provided in the virtual environment were its accessibility and user-friendliness as well as the provision of general information, resources and supplementary material regarding the course unit.

III. METHOD

Characterising research which is supposed to be innovative is always difficult, considering the perspectives it may follow. Taking this into account, we will characterise this research work mainly as far as its purpose, aims, nature and procedures are concerned.

With respect to its purpose, it follows an applied approach since we expect the results obtained to have a practical application in the improvement of the virtual learning environment of the institution that the students attend, which will lead them to benefit directly from the conclusions drawn from the study.

With regard to its aims, it follows an exploratory approach, since the data was obtained within a real context, but limited to only one higher education institution. For the results to be more generalised, the involvement of students from other institutions would be required. It assumes some of the features of a descriptive study as it describes relations between the variables under study.

As far as its nature is concerned, we consider that it follows a mixed approach, qualitative and quantitative. The qualitative nature is shown by its concern with indicators associated with the quality and support to students with regard to retention and dropout from degree courses before their completion. The quantitative nature is shown by the search for relations between variables, namely the determination of the correlation between pairs of variables and the multiple comparisons to test hypotheses through statistic procedures. The quantitative approach is also present in the objective measurements, structured and systematic data collection, statistics to quantify reality and results which may be conducive to generalisations [15]. Also, the variables assume numerical values and the relation between variables can be tested by means of statistic procedures [16], and the information is targeted towards assessing the amplitude of variation of the variables' values [17]. Furthermore, Coutinho [18] highlights that the researcher's interest in a quantitative study is to assume a scientific, distant and neutral attitude in order to test the hypothesis statistically and contribute to the process-product causal link.

With regard to the procedures, the data was obtained through computing processes related to the search and selection of data in databases associated with the virtual learning environment. We must say that these processes do not fit into the most common data collection techniques as in general, the main data collection techniques are: survey (questionnaire or interview); observation (direct and systematic or participatory) and desktop review. By analogy with the desktop review technique, where the information treated is mainly obtained from documents, we may consider that the main data collection technique used in this study was the analysis of data coming from databases containing academic information.

The data was provided by a sample of 1,588 undergraduates in the areas of Engineering (57.5%) and Business Sciences (42.5%) and refers to the 2015/2016 academic year. Among the students composing the sample, 36.1% are female and 63.9% are male. Their ages range from 19 to 61 years old. They were enrolled in the 1st year (41.5%), in the 2nd year (31.3%) and in the 3rd year (27.2%).

In this study, we assessed the degree of association between the Frequency_Accesses_VLE_CU and the number of course units passed. For comparing the means of the number of course units passed, we considered as an independent variable the Frequency_Accesses_VLE_CU and as dependent variables the learning outcomes associated with the umber of course units passed and the mean of the marks obtained in those course units.

Admitting that the number of course units in which students are enrolled may have an influence in the frequency of accesses to the virtual environment, we determined the degree of association between the variables mentioned above, with results showing a level of significance of 0.01, the coefficient of Pearson of 0.20, showing a rather low, though positive, correlation. Therefore, in order to relate the frequency of accesses to the virtual environment, the sample was divided into four groups considering the frequency of accesses to the virtual environment according to the number of course units in which the students were enrolled. Also, in order to determine the students' profile regarding their attendance to the virtual environment, each undergraduate's number of accesses was divided by the number of course units they were enrolled in. The result is a data sample coming from the reason between (each student's total number of accesses to the virtual environment) and (the number of course units each student is enrolled in), which constitutes a variable named Frequency_Accesses_VLE_CU.

The data sample described above originated four groups of students corresponding to each quartile of the referred sample.

The groups are characterised in Table 1.

TABLE 1: CHARACTERISATION OF THE GROUPS REGARDING THE SIZE AND THE FREQUENCY OF ACCESSES TO THE VIRTUAL ENVIRONMENT PER COURSE UNIT.

Groups	n	Frequency of accesses per CU	Mean	Median
G1	397	0 – 6.3	2.3	1.9
G2	397	6.4 – 16.3	11.0	10.9
G3	397	16.4 – 28.4	22.1	22.0
G4	397	28.4 – 337.0	45.6	38.7

We can see that all groups have the same size and that the figures of each statistic grow in the sense of the group number. Thus, resulting from the definition of the groups, the number of accesses to the virtual environment per course unit of the subjects in group 1 is lower than that of those in group 2, the number of accesses in group 2 is lower than that in group 3, and the number of accesses in group 3 is lower than that in group 4. An identical relationship was found regarding the mean and median of the distribution of the data concerning access to the virtual environment per course unit.

The characterisation of the groups' ages and gender is presented in Table 2.

As we can see from observing Table 2, the minimum age is the same in all groups, the maximum age and the mean of ages are higher in group 1. The majority of the subjects are male in all the four groups.

The processes used in the data collection and treatment lie upon the research field of Analytics, which according to Bichsel [19], focuses on the use of data, statistical analysis and exploratory and predictive models in order to obtain predictions

and act in complex issues.

This research work focuses on the learning analytics context, which in Mah's view [20], may support reflection on the interventions or activities targeted at perfecting the teaching and learning process, predicting and shaping activities to students, defending its use in early intervention to prevent dropout or adapt services and curricula within institutions.

Learning analytics is highly connected to each institution's virtual environment since this environment is where undergraduates have increasingly more information available, thus needing to access it to check timetables, assessments and course-related information, access learning materials, interact with others through forums and submit assignments proposed by teachers. While the student is learning, the virtual learning environment may generate big volumes of available data and identify useful measures related to students' behaviour and attainment [9].

TABLE 2: CHARACTERISATION OF THE GROUPS REGARDING AGE AND GENDER.

Group	Age					Gender	
	Min	Max	Mean	Med	STD EV	F (%)	M (%)
G1	19	61	27.1	25	6.4	28.5	71.5
G2	19	49	24.1	23	4.3	35.3	64.7
G3	19	52	22.3	22	4.2	42.3	57.7
G4	19	55	23.7	23	4.7	38.3	61.7

The main variables under study are Frequency_Accesses_VLE_CU, CU_Passed and Mean_CU. The variable Frequency_Accesses_VLE_CU corresponds to the quotient between the number of students' logs of accesses to the institution's virtual environment and the number of course units they were enrolled in. The variable CU_Passed represents the number of course units in which each student had a passing mark, meaning a mark higher or equal to 10. The variable Mean_CU represents the mean of the marks obtained in the course units each student passed.

IV. RESULTS

The results refer to data concerning 1,588 undergraduates during the 2015/2016 academic year, obtained from databases associated with the virtual learning environment used in the higher education institution they attended.

In order to determine the degree of association between the Frequency_Accesses_VLE_CU and the number of course units the undergraduates passed, we started by determining Pearson's correlation coefficient [20].

The correlation coefficient varies between -1 and +1 and the degree of association between the variables is all the stronger as the correlation coefficient is closer to -1 or +1 and all the weaker as the correlation coefficient is closer to zero. By using the Statistical Package for the Social Sciences (SPSS) between the variables Frequency_Accesses_VLE_CU and N_CU_Passed, we obtained a Pearson's correlation coefficient of 0.402 with a level of significance of 0.01. Therefore, considering the research question initially formulated, the results show that the two variables have a positive and moderate degree of association. Consequently, we can admit that the Frequency_Accesses_VLE_CU represents an indicator of the students' profile in the teaching and learning process as far as the course units they pass is concerned.

After analysing the degree of association between the variables, we moved on to exploring that relationship in order to assess the influence of the Frequency_Accesses_VLE_CU in the mean of the number of course units passed.

Considering that the variable Frequency_Accesses_VLE_CU assumes great variability of values, from 0 to 337, we made up four different groups according to this variable by dividing the distribution of data into quartiles, naming them group 1 (G1), group 2 (G2), group 3 (G3) and group 4 (G4) according to the students associated with the 1st, 2nd, 3rd or 4th quartile of the distribution, respectively.

In Table 3 we present the average number of course units passed by each group.

TABLE 3: MEANS OF THE COURSE UNITS PASSED PER GROUP

Groups (*)	Mean of course units passed	Standard Deviation
G1	1.0	2.2
G2	4.3	3.4
G3	6.6	3.3
G4	7.0	3.3

(*) each group has 397 subjects

The data in Table 3 shows that there are differences between the means of each pair of groups. Therefore, in order to answer the research question initially formulated, there is a need to assess whether such differences are significant, in other words, whether the differences detected between the number of course units passed vary significantly depending on a higher or lower frequency of accesses to the virtual learning environment per course unit.

In order to test the research hypothesis that "there are significant differences in the means of the number of course units passed between the groups with a higher Frequency_Accesses_VLE_CU and the groups with a lower Frequency_Accesses_VLE_CU", we assumed as a null hypothesis that there are no significant differences between the means of the course units passed by the students in each pair of groups.

Due to the impossibility of ensuring the distributions' normality and homogeneity through the statistical tests performed, we ran a less demanding test as far as normality and homogeneity are concerned. Considering that we intended to compare means between 4 groups, two by two, we conducted a multiple comparison of means by running Tukey's test, which is considered one of the most robust tests to deviations from normality and homogeneity of variances for large samples [21].

From the application of Tukey's test resulted the figures presented in Table 4.

TABLE 4: MULTIPLE COMPARISONS OF THE MEANS OF COURSE UNITS PASSED BY USING TUKEY'S HSD TEST.

Groups (I)	Groups (J)	Difference between means (I-J)
1	2	-3.224*
	3	-5.537*
	4	-6.010*
2	3	-2.312*
	4	-2.786*
3	4	-0.774

*. The difference between means is significant at level 0.05.

By observing Table 4 we find that except for the pair of groups 3 and 4, there are significant differences between the means of the number of course units passed in each pair of

groups, with a level of significance lower than 0.05. Considering the results presented in Tables 3 and 4 as well as the groups constitution process, we can state in response to the research question that there are significant differences between the mean of the number of course units passed in the following pairs of groups: group 1and group 2; group 1 and group 3; group 1 and group 4; group 2 and group 3; and group 2 and group 4.

Therefore, the results show that the subjects with a higher frequency of accesses to the virtual environment per course unit obtain higher results concerning the number of course units passed than those in the groups with a lower frequency of accesses to the virtual environment per course unit.

In order to assess the relationship between the number of CU each group passed and the mean of the marks obtained in the CU passed, we considered in each group the subjects who passed at least one CU and we excluded those who did not pass any CU, as shown in Table 5.

TABLE 5: DATA REGARDING LEARNING OUTCOMES IN TERMS OF COURSE UNITS PASSED

Groups (*)	With no CU passed (%)	With at least one CU passed (%)
G1	67.3	32.7
G2	11.6	88.4
G3	3.0	97.0
G4	1.5	98.5

CU – Course unit

(*) each group has 397 subjects

The data in Table 5 shows that it is in group 1, which has the lowest mean of frequency of accesses to the virtual environment per course unit, that we find the highest percentage of students who did not pass any curse unit (67.3%), followed by group 2 (11.6%), group 3 (3.0%) and group 4 with only 1.5% of students not passing any course unit. From these results we conclude that the lower the frequency of accesses to the virtual environment per course unit is, the lower the number of course units students pass will be.

In order to determine the mean of the marks obtained in the course units that students passed, we excluded from each group the subjects who did not pass any course unit, which gave birth to the following groups: group 1 turned into Ga1, group 2 into group Ga2, group 3 into Ga3 and group 4 into Ga4. In Table 6, we present the means of the number of course units passed, the mean of the marks obtained in those course units and the number of subjects composing each group.

TABLE 6: DATA REGARDING ACADEMIC RESULTS OF THE STUDENTS WHO PASSED AT LEAST ONE COURSE UNIT.

Groups	Mean of CU passed	Mean of the marks of the CU passed	N. Subjects
Ga1	3.2	11.8	130
Ga2	4.8	11.5	351
Ga3	6.8	12.1	385
Ga4	7.2	12.3	391

The observation of Table 6 allows the conclusion that the mean of the number of course units passed varies in an ascending trend from group Ga1 (3.2) to group Ga4 (7.2), which according to the constitution of the groups, shows that the higher the frequency of accesses to the virtual environment per course unit is, the higher the number of course units passed is too. With regard to the mean of the marks obtained, we find that except for group Ga1, there is a slight upward trend as the frequency of accesses to the virtual environment per course unit increases, since in group Ga2 the mean is 11.5, in group Ga3 it is 12.2 and in group Ga4 it is 12.3.

In the light of the results obtained, we find that there is a positive moderate correlation between the variables Frequency_Accesses_VLE_CU and the number of course units passed. Also, we find significant differences between the means regarding the number of course units passed between the students revealing a higher frequency of accesses and those revealing a lower frequency of accesses per course unit, to the advantage of the groups with a higher frequency of accesses to the virtual environment per course unit.

In a nutshell, the results show that the low frequency of accesses to the institution's virtual environment is an indicator to be taken into account as far as learning outcomes are concerned, since the groups revealing a lower frequency of accesses per course unit showed to have passed a lower number of course units than those who reveal a higher frequency of accesses to the virtual environment.

V. CONCLUSIONS

This study involved the analysis of data concerning 1,588 undergraduates related to the frequency of access to a virtual learning environment of a Portuguese higher education institution, the course units which the sample subjects passed and the marks obtained in the course units they passed.

Among the results obtained, we highlight that the degree of association between the variables Frequency_Accesses_VLE_CU and the mean of the number of course units which students passed is positive and moderate; Considering the division of the sample into quartiles, G1, G2, G3 and G4, respectively 1st, 2nd, 3rd and 4th according to the number of accesses to the virtual environment per course unit, the size of each group is of 397 subjects and we find that the frequency of accesses to the virtual environment may be a good indicator of the student's profile, since the vast majority (67.3%) of students who did not pass any course unit was found to be part of group 1 (1st quartile of the distribution of accesses to the virtual environment), followed by group 2 (11.6%), group 3 (3.0%) and group 4 with 1.5% of the students not passing any course unit; Except for group 3 and group 4, significant differences with a level of significance below 0.05 were found between each pair of groups as far as the mean of the number of course units passed is concerned, and the groups with a higher frequency of accesses to the virtual environment per course units are the ones who show the highest mean of the number of course units passed. In the distribution of the sample into quartiles according to the number of accesses to the virtual environment per course unit, the subjects from the 2nd quartile passed more course units than those from the 1st quartile, those from the 3rd quartile passed more CU than those from the 2nd and 1st quartiles, and those from the 4th quartile passed more CU than those from the 3rd, 2nd and 1st quartiles, which means that the higher the frequency of accesses to the virtual environment per course unit is, the higher the number of course units which students pass will be.

These results show the importance of research studies in the field of learning analytics as they provide indicators on students' action within the teaching and learning process from concrete data recorded in virtual learning environments, thus enabling the prediction of some situations which might constrain students' future within the institution they attend, especially regarding retention and dropout from the degree courses before their completion.

This study shows that knowing the frequency of accesses to an institution's virtual environment enables stakeholders to

predict quite reasonably their students' attainment or retention in terms of course units they will pass as well as their likeliness to drop out from the institution and consequently, from the teaching and learning process in which they should be regularly involved. Therefore, it is in the highest interest that the institutions' virtual learning environments provide students, institutions and teachers with the statistics regarding students' access to the virtual learning environment. By comparing the means of accesses to the course units in which a student is enrolled, it is possible to anticipate support to the student and consequently, increase the probability of decreasing the risk of retention or dropout from the institution they attend.

The need for institutions, teachers and students to take notice of students' frequency of accesses to the virtual environment of the institution they attend is reinforced by the results obtained in a study conducted by Alves, Miranda and Morais [22], which involved data regarding the academic years of 2009/2010 and 2013/2014, from a sample of approximately 7,000 undergraduates per year. That study showed that virtual learning environments represent spaces for sharing contents between teachers and students, although the months of higher teachers' provision of contents does not correspond to the months of higher students' access to the virtual environment, since the latter access contents more frequently during the months they undergo final assessments than during the rest of the year.

Therefore, virtual environments must increasingly constitute not only support tools for both teachers and students but also monitoring tools which may help prevent undesirable situations related to students' academic attainment.

REFERENCES

[1] Siemens, G., Gasevic, D. "Guest Editorial - Learning and Knowledge Analytics". Educational Technology & Society, 15 (3), 1–2. 2012.

[2] Tempelaar, D. , Rienties, B., Giesbers, B. "In search for the most informative data for feedback generation: Learning Analytics in a data-rich context". Computers in Human Behavior, 47, 157-167. 2015.

[3] X. Shacklock, From Bricks to Clicks: The Potential of Data and Analytics in Higher Education. Higher Education Commission, 2016.

[4] M. Conde and Á. Hernández-García, "Learning analytics for educational decision making" Computers in Human Behavior, vol. 47, pp. 1-3, 2015

[5] Á. Agudo-Peregrina, S. Iglesias-Pradas, M. Conde-González and Á. Hernández-García, "Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning", Computers in human behavior, vol.31, pp. 542-550, 2014.

[6] Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Hall, C., NMC horizon report: 2016 higher education edition. Austin, TX: New Media Consortium, 2016.

[7] Long, P.D., & Siemens, G., Penetrating the fog: Analytics in learning and education. Educause Review Online. Retrieved from http://www.educause.edu/ero/article/penetrating-fog-analyticslearning-and-education. 2011.

[8] J. You, "Identifying significant indicators using LMS data to predict course achievement in online learning", The Internet and Higher Education, pp. 23-30, 2016

[9] N. Sclater, A. Peasgood and J. Mullan, "Learning analytics in higher education. A review of UK and international practice". Full report. JISC. Publiziert unter der Lizenz CC BY, vol.4, 2016

[10] L. Johnson, S. Adams Becker, V. Estrada, and A. Freeman, "NMC Horizon Report: 2014" Higher Education Edition. Austin, Texas: The New Media Consortium, 2014.

[11] G. Siemens and R. Baker, "Learning analytics and educational data mining: towards communication and collaboration" Proceedings of the 2nd international conference on learning analytics and knowledge. ACM, 2012.

[12] M. Vahdat, L., Oneto, D. Anguita, M. Funk, and M. Rauterberg, "Can machine learning explain human learning?", Neurocomputing, Vol. 192, pp. 14-28, 2016.

[13] B. Rienties, A. Boroowa, S. Cross, C. Kubiak, K. Mayles, and S. Murphy, "Analytics Action Evaluation Framework: A Review of Evidence-Based Learning Analytics Interventions at the Open University UK", Journal of Interactive Media in Education, vol. 2, pp. 1–11, 2016.

[14] Morais, C., Alves, P., & Miranda, L. Valorização dos ambientes virtuais de aprendizagem por professores do ensino superior. In A. Rocha, L. Reis, M. Cota, M. Painho, M. Neto (eds.). Sistemas e Tecnologias de Informação, Atas da 8ª Conferência Ibérica de Sistemas e Tecnologias de Informação. Vol. I, pp. 289-294. Lisboa: Associação Ibérica de Sistemas e Tecnologias de Informação. ISBN: 978-989-96247-9-5. 2013

[15] M. D'Ancona, "Metodología cuantitativa: estrategias y técnicas de investigación social", Madrid: Editorial Síntesis, 1998.

[16] J. Creswell, "Research design: Qualitative, quantitative, and mixed methods approaches" (4th ed.). London: SAGE Publication Ltd, 2014.

[17] R. Kumar, "Research methodology: A step-by-step guide for beginners", London: SAGE Publication Ltd, 2011.

[18] C. Coutinho, Metodologia de investigação em ciências sociais e humanas: teoria e prática. Coimbra: Almedina, 2013.

[19] J. Bichsel, "Analytics in Higher Education, Benefits, Barriers, Progress and Recommendations", EDUCAUSE Center for Applied Research, 2012

[20] Mah, D. "Learning Analytics and Digital Badges: Potential Impact on Student Retention in Higher Education". Tech Know Learn, 21, 285–305. 2016.

[21] J. Maroco, "Análise estatística: Com utilização do SPSS (3. ª Ed.)", Lisboa: Edições Sílabo, 2010

[22] Alves, P., Miranda, L.,& Morais, C. Publicação e Acesso a Conteúdos no Ambiente Virtual de Aprendizagem de uma Instituição de Ensino Superior. In A. Rocha, A. Martins, G. Dias, L. Reis, & M. Cota (eds.). Sistemas e Tecnologias de Informação, Atas da 10ª Conferência Ibérica de Sistemas e Tecnologias de Informação, CISTI 2015 (Vol. I, Tomo 2, pp. 881-886). Braga: APPACDM-Associação Portuguesa de Pais e Amigos do Cidadão Deficiente Mental, ISBN: 978-898-98434-4-8, DOI: 10.1109/CISTI.2015.7170468. 2015.