Educational Data Mining to Identify the Patterns of Use made by the University Professors of the Moodle Platform

Johan Calderon-Valenzuela, Keisi Payihuanca-Mamani, Norka Bedregal-Alpaca

Universidad Nacional de San Agustín de Arequipa

Arequipa, Perú

Abstract—Due to the events caused by the COVID-19 pandemic and social distancing measures, learning management systems have gained importance, preserving quality standards, they can be used to implement remote education or as support for face-to-face education. Consequently, it is important to know how teachers and students use them. In this work, clustering techniques are used to analyze the use, made by university professors, of the resources and activities of the Moodle platform. The CRISP-DM methodology was applied to implement a data mining process, based on the Simple K-Means algorithm; to identify associated groups of teachers it was necessary to categorize the data obtained from the platform. The Apriori algorithm was applied to identify associations in the use of resources and activities. Performance scales were established in the use of Moodle functionalities, the results show the use made by teachers was very low. Rules were generated to identify the associations between activities and resources. As a result the functionalities that need to be enhanced in the teacher training processes were identified. Having identified the patterns of use of the Moodle platform, it is concluded that it was necessary to use a Likert scale to transform the frequency of use of activities and resources and identify the rules of association that establish profiles of teachers and tools that should be promoted in future training actions.

Keywords—*Clustering; educational data mining; moodle; usage patterns; k-means algorithm; a priori algorithm*

I. INTRODUCTION

The pandemic caused by COVID-19 brought with it a mandatory social confinement that forced societies to change customs and models and to face the need to enhance the use of technological tools to give rise to teleworking and tele education.

The face-to-face educational system necessarily migrated towards virtual teaching-learning environments, which generated new ways of understanding the educational process. Integrating information and communication technologies (ICT) to the educational process implies changes in the forms of communication, in the contents and forms of evaluation, changes in the role of the teacher and students, ICT can be used by teachers as technical-pedagogical support and by students as a tool for autonomous learning [1].

Although some educational institutions were gradually integrating ICTs as a means of enhancing their educational processes, in the context of the pandemic, many of them had to abruptly assume, without considering the context of their educational community, the use of virtual scenarios to carry out their educational activities.

Although some educational institutions were gradually integrating ICT as a means of enhancing their training processes, in the context of the pandemic, many of them had to assume abruptly and without considering the context of their educational community, the use of virtual scenarios to carry out their training activities. In particular, universities had to adapt to this transformation by assuming educational models that use technological tools of support and accompaniment to improve conventional teaching and learning processes. Within these tools are virtual educational platforms or LMS (Learning Management System) that enable new teaching-learning modalities. For [2] in the b-learning modality, which combines face-to-face with non-face-to-face teaching, the LMS favors learning. In the e-learning modality, the use of Internet-based technologies provides a wide range of solutions that promote the acquisition of knowledge and development of skills [3].

Training processes in virtual environments imply changes in the roles of their actors, the student stops being a passive consumer and becomes a producer of information and knowledge; for [4], the students must be autonomous and independent in their information search skills, must create new content in an innovative way and transform it into knowledge and must communicate effectively by decoding messages and transmitting information. For his part, the teacher assumes the role of guide and incorporates technology adequately and effectively, that is, incorporating technological tools with a pedagogical approach [5].

In this information context, in order for students to achieve meaningful learning, the teacher must use the LMS tools to manage resources and activities in a way that enhances the autonomous work of the student and favors the development of competencies and skills such as search and organization of information, teamwork and communication with their peers and with the teacher. That is why this research focuses on improving university teacher performance and consequently student learning.

For the analysis and evaluation of the use of LMS, data mining is a tool that allows determining patterns of behavior in the data obtained from the platform and identifies the factors associated with the success of online learning [6]. In

[7, 8] they found that most works focus on the usability of LMS, limiting themselves to the student's role and neglecting the teacher's work process.

In [9] a statistical analysis is made of how mathematics teachers in a face-to-face model supported by the Moodle platform develop cognitive and action competencies in elementary school students. In [10] educational data mining is applied to the Moodle LMS to identify behavior patterns in students to identify the resources and activities that are best suited to the students' needs; it concludes that there is a correlation between the level of activity and their academic performance.

In [11], the authors investigated the use of Virtual Learning Environments by professors in a higher education institution; they sequentially combined three methods: processing of VLE logs, surveys, and interviews. In [12] data is collected in some Spanish universities to study the uses that university professors make of the virtual campus and the methodology they propose to students. The model of courses carried out concludes that the universities studied make minimal use of the platform and focus on making materials available to the student and that the model is characterized by the presence of basic, complementary, and organizational teaching materials, together with a proposal of individual and group activities.

Works such as [13, 14, and 15] have used the CRISP-DM methodology to implement data mining processes aimed at describing students' academic behavior.

In [16], the authors use data mining tools and techniques for academic improvement of the student performance and to prevent drop out. Four classification methods, the J48, PART, Random Forest and Bayes Network Classifiers were used, the data mining tool used was WEKA.

The authors in [17], propose two different guidelines: Learning Analytics focused on descriptive processes, and Educational Data Mining for predictive processes, directing activities adjusted to this environment for obtaining satisfactory results.

In light of the above, the objective of this work was to analyze the use of resources and activities of the Moodle platform, made by teachers at the National University of San Agustin de Arequipa (Peru). The importance of the work carried out is that based on the results obtained, recommendations can be made to enhance the use of this platform and favor the use of interactive and collaborative activities within the framework of a socio-constructivist pedagogical model.

II. THEORETICAL FRAMEWORK

A. LMS Moodle

Modular Object-Oriented Dynamic Learning Environment (Moodle) is a learning management tool or LMS developed to create and manage online training environments. Known by its acronym, Moodle is one of the most widely used content management systems globally. It provides a powerful set of learner-centered tools and collaborative learning environments that empower both teaching and learning.

Moodle fosters active and participatory virtual environments by enabling teachers and students to interact on the platform through chats, forums, videoconferences, among others. Consequently, it facilitates to the teacher the possibility of extending the limits of the classroom to spaces and moments different from the face-to-face class, it gives students autonomy to consult multimedia content and to interact and participate in learning communities [18].

B. Educational Data Mining

Data mining is a set of techniques and technologies that, in an automatic or semi-automatic form, allows the exploration of large databases; the objective is to find repetitive patterns, trends, or rules that explain the behavior of the data in a specific context. Data mining involves the interaction of different techniques and procedures from computer science, statistics, mathematics, and information science. The extraction and analysis of data has become important in multiple areas, because through the application of various techniques, it allows to transform that data into information and knowledge of great utility [19].

In the educational field, there is talk of educational data mining (EDM); which focuses on the development of discovery methods that use data from LMS to understand and improve virtual teaching-learning environments. The EDM builds analytical models that uncover interesting patterns and trends in the use of an LMS [20].

C. CRISP-DM Methodology

CRISP-DM (Cross Industry Standard Process for Data Mining) is a standard and open analytical model of the data mining process [21]. It includes a description of the phases of a data mining project, the tasks required in each phase, and an explanation of the relationships between the tasks. CRISP-DM provides an overview of the life cycle of a data mining process.

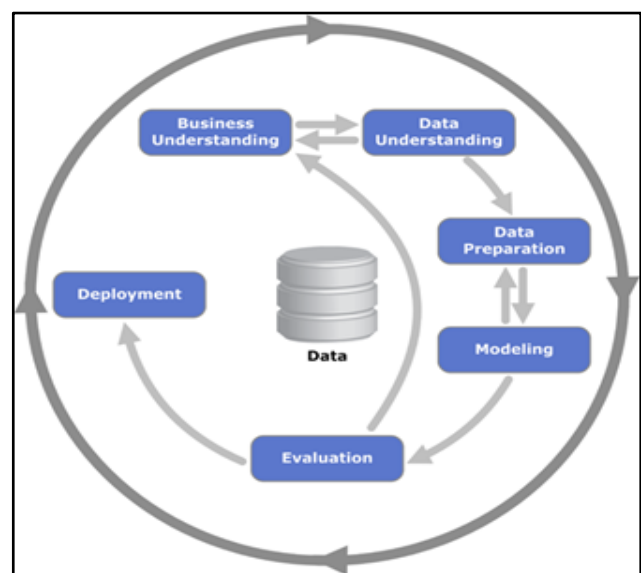CRISP-DM divides the data mining process into six main phases (Fig. 1).

Fig. 1.   CRISP-DM Phases. Source Shearer (2000).

The sequence of phases is dynamic; it is possible to move forward or backward through the phases. The result of each phase determines which phase, or which particular task of a phase, is to be done next. The arrows indicate the most important and frequent dependencies. The advantages of using CRISP-DM include replicability, independence of the application context, and its tool neutrality [22].

III. MATERIALS AND METHODS

A. Context

The research was conducted at the National University of San Agustín (UNSA) in Arequipa, Peru. Traditionally it works under a face-to-face model supported by the use of the Moodle platform. The University Directorate of Information Technologies (DUTIC-UNSA) is the unit in charge of the administration and management of the Moodle platform, in addition, it provides training and technical support services to teachers and students.

As a result of the COVID-19 pandemic, the UNSA migrated towards a mixed model in which face-to-face sessions are carried out through a videoconferencing system and arranged the mandatory use of the Moodle platform to dynamize and enhance the training process.

B. Methodology

Cross-sectional descriptive-exploratory research has been carried out.

To develop the data mining process, the steps of the CRISP-DM methodology have been followed.

C. Data

The study considered a universe composed of 4809 virtual classrooms implemented in the 2020-A academic semester. These classrooms correspond to the three academic areas of the UNSA: engineering, social and biomedical.

IV. DEVELOPMENT OF THE PROPOSAL

A. Phase 1, CRISP-DM Methodology: Understanding the Business

Tasks related to understanding project objectives and requirements were performed to turn them into technical objectives and a project plan.

Objectives: This research focuses on analyzing the use that teachers make of the activities and resources available on the Moodle platform. It is intended to:

- Investigate if there are differences in use between the three academic areas: Natural and Formal Sciences and Engineering, Social Sciences and Biomedical Sciences.

- Identify associated groups of teachers in the use of resources and activities using the Unsupervised Learning Algorithm Simple K-Means.

- Identify associations in the use of resources and activities using the A-priori association algorithm.

Resources:

- Technological: Computer, Google Collaboratory platform in Python language for data processing and model development.

- Technical: Data mining techniques.

- Human: The Researchers.

- Data source: Moodle platform usage logs during the 2020-A Academic Semester.

B. Phase 2, CRISP-DM methodology: Understanding the Data

A first contact with the data was made to familiarize themselves, identify their quality, define the first hypotheses and establish the most obvious relationships.

1) Data collection: The data was obtained from the records of the Moodle platform. For the data to have meaning it was necessary to integrate different tables. The "teacher" table with the correlative code of each teacher and their full name. The "course" table contains the correlative code of each virtual classroom, the short name and the full name of the corresponding subject. Both are related through the "course_teacher" table. The "resource_activities" table contains the number of times that the 13 activities and 7 resources available on the virtual platform have been used. This table relates to "course_teacher" through the course and teacher correlates. Fig. 2 shows the entity-relationship diagram of the integration performed.

2) Verification of data quality: Initially, the database considered the 4809 classrooms configured on the virtual platform. Data was cleaned considering that some classrooms had not been used or had not been assigned a teacher; 2 duplicate teacher records were deleted, 219 classrooms that had the Teaching field blank, and 335 classrooms that had less than or equal to a resource or activity used. As a result, we worked with data from 4255 classrooms.
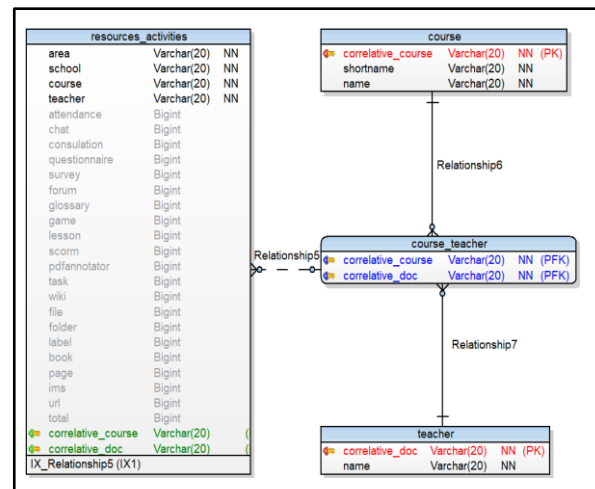

Fig. 2. Entity-relationship Diagram.

3) Data exploration: The use of classrooms was analyzed for each academic area, Fig. 3 shows the percentage of classrooms implemented and the number of teachers who were in charge in each of them. The differences found are explained by the number of professional schools that are integrated into each academic area.

At UNSA, the academic areas are divided into Schools and these into Professional Schools. Table I shows the number of virtual classrooms implemented in the different schools, as well as the number of professors administered them.

Tables II and III show the frequency of the use of Moodle resources and activities in each academic area and the amount of total usage.
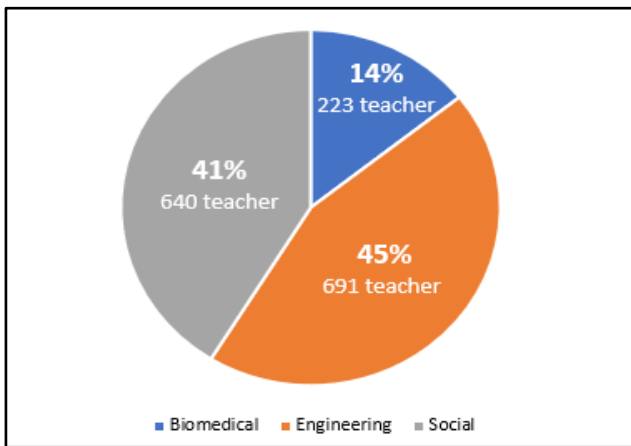

Fig. 3. Classrooms Percentage and Number of Professors in each Academic Area.

TABLE I. NUMBER OF CLASSROOMS BY SCHOOLS

School	No. of classrooms	No. of professors
School of Administration	269	115
School of Agronomy	93	48
School of Architecture	159	76
School of Biological Sciences	194	102
School of Accounting and Financial Sciences	172	80
School of Historical and Social Sciences	317	130
School of Natural and Formal Sciences	151	93
School of Law	189	77
School of Economics	130	49
School of Education	263	122
School of Nursing	76	47
School of Philosophy and Humanities	271	96
School of Geology, Geophysics and Mining	174	85
School of Civil Engineering	203	84
School of Process Engineering	403	210
School of Production and Services Engineering	750	313
School of Medicine	83	44
School of Psychology RRII Cs. Of Communication	358	146

TABLE II. FREQUENCY OF USE EACH RESOURCE BY AREA

Resources / Area	Biomedical	Engineering	Social	Total
File	13373	43299	46283	102955
Folder	675	1764	1512	3951
Label	2253	13042	9173	24468
Book	21	102	68	191
Page	149	224	238	611
Ims	0	2	0	2
Url	9476	33376	37343	80195
Total	16471	58433	57274	

TABLE III. FREQUENCY OF USE EACH ACTIVITY BY AREA

Resources / Area	Biomedical	Engineering	Social	Total
Attendance	1481	5421	5335	12237
Chat	275	1800	1614	3689
Consultation	10	81	106	197
Questionnaire	2771	6590	6697	16058
Survey	5	2	16	23
Forum	1469	4761	6962	13192
Glossary	71	87	138	296
Game	20	85	114	219
Lesson	19	287	176	482
Scorm	0	5	4	9
Pdf Annotator	152	169	117	438
Tasks	6808	21066	18617	46591
Wiki	14	29	39	82
Total	13081	40354	39896	

C. Phase 3, CRISP-DM Methodology: Data Preparation

In this phase, the necessary activities are carried out to build the data set that will serve for the modeling.

Each classroom was evaluated according to the frequency of use of each activity and resource. Finally, only the activities or resources used in at least 97% of the classrooms in each area were considered (Tables IV and V).

Consequently, in the academic area of Social Sciences ten tools were considered: Attendance, Chat, Questionnaire, Forum, Glossary, Task, File, Folder, Label and URL.

In the Engineering area, eleven tools were included: Assistance, Chat, Questionnaire, Forum, Glossary, Task, File, Folder, Label, Page and URL.

In the area of Biomedical Sciences, twelve tools were considered for the study: Attendance, Chat, Questionnaire, Forum, Glossary, Pdf Annotator, Task, File, Folder, Label, Page and URL.

In order to complement the information, Tables IV and V show the number of classrooms that have never used the activities or resources indicated by area of knowledge.

TABLE IV.   NUMBER OF CLASSROOMS THAT HAVE NEVER USED EACH RESOURCE BY AREA

Resources / Area	Biomedical	Engineering	Social
File	10	47	34
Folder	337	1422	1545
Label	284	1198	1481
Book	437	1801	1932
Page	419	1774	1916
Ims	446	1838	1969
URL	46	327	265

TABLE V.   NUMBER OF CLASSROOMS THAT HAVE NEVER USED EACH ACTIVITY BY AREA

Resources / Area	Biomedical	Engineering	Social
Attendance	173	1705	983
Chat	353	1781	1538
Consultation	436	1792	1916
Questionnaire	117	734	634
Survey	443	1838	1958
Forum	10	62	66
Glossary	416	1783	1881
Game	434	1803	1935
Lesson	433	1810	1925
Scorm	446	1835	1966
Pdf Annotator	429	1811	1918
Task	70	226	255
Wiki	437	1827	1948

Calculated the frequencies of use of the Moodle functionalities, it was found that the values were very dispersed so it was necessary to reduce them. Considering that scale transformations are efficient instruments for reducing a data set, it was decided to transform the values from numeric to nominal.

On the cleaned data, the maximum and minimum values of the frequency of use of each activity and resource were identified. With these values the range was calculated and divided into five parts of equal length, obtaining the values to construct a five-level Likert scale. The scale values and weight of each level assigned were:

0 = Very low
1 = Low
2 = Medium
3 = High
4 = Very high

D. Phase 4, CRISP-DM Methodology: Modeling

In this phase, the most appropriate modeling techniques are selected for the specific data mining project.

The Simple K-Means classification algorithm was applied, using the elbow method. It was looked the part of the graph

where the line changes abruptly which forms an "elbow"; that number of clusters will help when classifying the data. The appropriate was 30 iterations to determine the optimal number of clusters, obtaining k = 5 (Fig. 4).

The objective was to group similar observations and discovers patterns in the use of Moodle resources and activities in the transformed data based on the Likert scale.

It was worked with 4255 classrooms; each classroom has 30 fields including: area, faculty, school, course, teacher, activities and resources. Each classroom has the amount of resources and activities used numerically, therefore the Likert scale was applied to have a standard score from 0 to 4 with respect to usability; then, we eliminate columns that are not going to be trained, the data set used consists of 20 fields between activities and resources, which will be grouped using the K means algorithm, resulting in 5 groupings, cluster 1, 2 and 4 have less usability, cluster 5 has a medium usability followed by cluster 3 (Fig. 5) and Table VII.

To identify associations in the use of resources and activities, the A-priori association algorithm was applied in each academic area.

First, we converted the data into 0's (activities and resources with score 0) and 1's (activities and resources with score 1,2,3 and 4) to work the model.
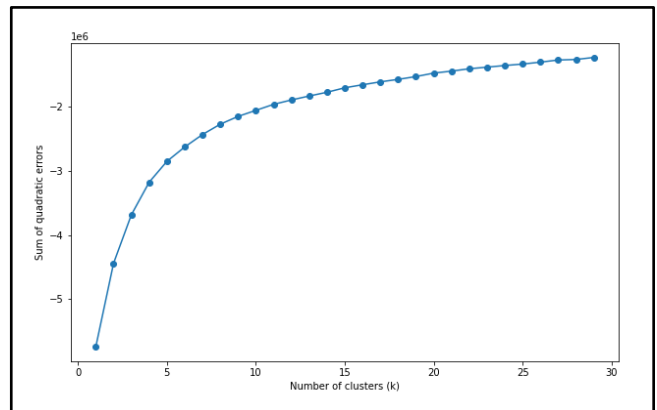

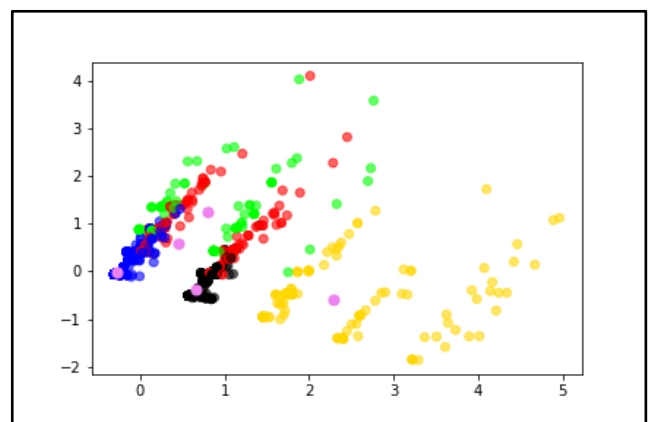Fig. 4.   Elbow Method to Determine the Optimal value of k.

Fig. 5.   Clusters Determined by Simple K-Means.

In the social area, associations were obtained with the tools file, URL, and task. In the biomedical area, associations were obtained with the questionnaire, homework, URL, and attendance. However, in the engineering area no association was found; therefore. Due to these differences, it was decided to perform the process again with the total data set.

Considering the definition of support given by [23], the association rules were constructed with a minimum support of 0.8; increasing the number of associations, where the following is true: antecedent = is used => consequent = is used, the probability of each association is observed in the "Confidence" column, the Lift column shows the increase in the probability of use when considering both tools rather than only the antecedent tool. (Table VI shows the results).

TABLE VI.   ASSOCIATIONS FOUND WITH THE APRIORI ALGORITHM

Antecedent	Consequent	Support	Confidence	Lift
URL	file	0.840	0.988	1.009
forum. URL	file	0.810	0.987	1.009
task	file	0.859	0.987	1.009
task, forum	file	0.830	0.987	1.008
task	file, forum	0.830	0.954	1.008
URL	file, forum	0.810	0.953	1.008
File	task	0.859	0.878	1.009
file, forum	task	0.830	0.877	1.008
file	URL	0.840	0.858	1.009
file, forum	URL	0.810	0.856	1.008
file	file, forum	0.830	0.848	1.008
file	forum, URL	0.810	0.828	1.009

V.   RESULT AND DISCUSSION

When analyzing Tables II and III, it can be affirmed that the resource "File" has been used primarily, followed by the resource "URL"; it follows then, that the virtual classroom is being used primarily as a repository of content and links to web content. Likewise, the activity "Homework" has been mostly used, followed by the activity "Questionnaire"; which means that the virtual classroom is being used to collect assignments and make assessments through questionnaires.

Table VIII (at the end of the paper) shows the results obtained by applying the Simple K-Means algorithm to the processed data. The nominal values of the Likert scale were considered for each of the tools in the clusters, since the distribution of the frequency of use is asymmetric, the value of the model was taken as a representative measure. Table VIII also shows that the teachers' use of Moodle activities and resources was classified with very low performance. Although this situation differs slightly in the areas of Social Sciences and Biomedical Sciences, it was found that in Engineering there was not enough support to find association rules.

There are no reasons to assume that teachers do not know how to use the different functionalities of the virtual platform; however, these results suggest the need to reinforce teacher training in the use of the activities and resources provided by Moodle. The training process should focus on the use of the virtual platform as a support to the synchronous sessions conducted by the teacher so that learning outcomes are enhanced through the implementation of active teaching-learning methodologies that can be easily implemented with the functionalities provided by Moodle [24,25,26].

Table VII shows the association rules obtained as a result of applying the Apriori algorithm to the cleaned data corresponding to the three academic areas. The association rules relate to the use of Moodle activities and resources.

TABLE VII.   ASSOCIATION RULES OBTAINED FROM THE APRIORI ALGORITHM

#	Rules of association	Interpretation in Natural language
1	URL = is used=> file= is used	If the URL tool is used in the classroom, then the file tool will be used.
2	forum, URL = is used => file = is used	If both the forum and URL tools are used in the classroom, then the file tool will be used.
3	task = is used => file = is used	If in the classroom the task tool is used, then the file tool will be used.
4	task, forum = is used => file = is used	If both task and forum, are used in the classroom, then the file tool will be used
5	task = is used=> file, forum = is used	If in the classroom the task tool is used, then the file and forum tools will be used.
6	URL = is used => file, forum = is used	If the URL tool is used in the classroom, then the file and forum tools will be used.
7	file = is used => task = is used	If the file tool is used in the classroom, then the task tool will be used.
8	file, forum = is used => task = is used	If in the classroom the tools file and forum are used, then the task tool will be used.
9	file = is used => URL = is used	If the file tool is used in the classroom, then the URL tool is used
10	file, forum = is used => URL = is used	If both the file and forum tools are used in the classroom, then the URL tool will be used.
11	file = is used => task, forum = is used	If the file tool is used in the classroom, then the task and forum tools will be used.
12	file = is used => forum, URL = is used	If in the classroom the file tool is used, then the forum and URL tools will be used.

Association rules have been generated and they have allowed the identification of associations between activities and resources, rules that give evidence of the activities and resources that need to be enhanced in the teacher training processes. For example, from rules 5, 6, and 11 it can be inferred that if the teacher uses the task activity and the resources file and forum, then it is very likely that he/she will implement discussion forums among the activities to be implemented in the virtual classroom; therefore, he/she will not only be imparting knowledge but also supporting the development of communication skills and critical thinking.

TABLE VIII.   RESULTS OF THE SIMPLE K-MEANS CLUSTERING ALGORITHM

Variable	Group 1 (3291 classrooms)		Group 2 (72 classrooms)		Group 3 (89 classrooms)		Group 4 (621 classrooms)		Group 5 (182 classrooms)	
Attendance	3231	Very low	65	Very low	66	Medium	595	Very low	160	Very low
Chat	3265	Very low	64	Very low	88	Very low	594	Very low	178	Very low
Consultation	3291	Very low	72	Very low	89	Very low	621	Very low	182	Very low
Questionnaire	3254	Very low	55	Very low	79	Very low	583	Very low	122	Very low
Survey	3291	Very low	72	Very low	89	Very low	621	Very low	182	Very low
Forum	3232	Very low	60	Very low	88	Very low	603	Very low	163	Very low
Glossary	3280	Very low	67	Very low	87	Very low	610	Very low	176	Very low
Game	3291	Very low	72	Very low	89	Very low	621	Very low	182	Very low
Lesson	3291	Very low	72	Very low	89	Very low	621	Very low	182	Very low
Scorm	3291	Very low	72	Very low	89	Very low	621	Very low	182	Very low
Pdf Annotator	3285	Very low	69	Very low	89	Very low	621	Very low	182	Very low
Task	3291	Very low	26	Very low	45	Very low	549	Very low	118	Very low
Wiki	3291	Very low	72	Very low	89	Very low	621	Very low	182	Very low
File	3140	Very low	38	Very low	66	Very low	537	Very low	109	Low
Folder	3274	Very low	62	Very low	82	Very low	604	Very low	177	Very low
Label	3230	Very low	62	Very low	82	Very low	582	Very low	158	Very low
Book	3291	Very low	72	Very low	89	Very low	621	Very low	182	Very low
Page	3291	Very low	71	Very low	89	Very low	616	Very low	182	Very low
Ims	3291	Very low	72	Very low	89	Very low	621	Very low	182	Very low
url	3158	Very low	31	Very low	59	Very low	553	Very low	103	Low

VI. CONCLUSION

With the development of this work, it was possible to identify the behavior patterns of university professors in the use of the activities and resources offered by the Moodle platform.

The use of a Likert scale to transform the frequency of use of activities and resources allowed to reduce the spectrum of values and to be able to find associations when applying the K-Means algorithm; using the elbow method with 30 it was determined that the optimal was to work with 5 clusters.

The activity of the teachers was characterized and it was found that the activities: chat, wiki, lesson, workshop, questionnaire, games and survey are not used despite their great potential as didactic material that can enhance the results of the teaching and learning processes.

The results obtained in this work will serve to implement teacher training processes focused on the proper use of Moodle activities and resources that allow the development of virtual or blended courses based on constructionist and social constructivist approaches.

Applying data mining techniques to the large amount of information generated by the Moodle platform can contribute to the creation of dynamic profiles in the development of a virtual course. In addition to improvements in the teacher's use of resources and activities, students' behavior patterns could be considered in order to adapt courses to their level of learning.

REFERENCES

[1] I. Martí Castro, "Aprendizaje-Virtual". En Diccionario Enciclopédico de Educación. Grupo Editorial Ceac S. A. (LEXUS), 2003.

[2] A. Bartolomé, "Blended Learning. Conceptos básicos. Píxel-Bit". En Medios y Educación, 23. Pp. 7-20. [en línea], 2004.

[3] D. Alemany Martínez, "Blended learning: modelo virtual – presencial de aprendizaje y aplicación en entornos educativos". España, Alicante: Universidad de Alicante, 2010.

[4] J. Cabero, "La formación en la sociedad del conocimiento. INDIVISA - Boletín de Estudios E Investigación - Monografía X: Las TICs en los contextos de formación universitaria". 13-48, 2008.

[5] F. Pedró, "Tecnología y escuela: lo que funciona y por qué: Documento Básico". Madrid: Fundación Santillana. Obtenido de http://www.fundacionsantillana.com/upload/ficheros/noticias/201111/documento_bsico.pdf, 2011.

[6] R. Herrera, Myriam; Ruiz, Susana; Romagnano, María; Ganga, Leonel; Lund, María Inés y Torres, Estela. "Aplicando métodos y técnicas de la ciencia de los datos a datos universitarios". XXI Workshop de Investigadores en Ciencias de la Computación, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de San Juan, Argentina, Abril 2019.

[7] A.Van Leeuwen; J. Janssen; G. Erkens; M. Brekelmans, "Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL". Computers & Education, vol. 79, 28-39, 2014.

[8] J. M. Dodero, F. J. García-Peñalvo, "Development of E-Learning Solutions: Different Approaches, a Common Mission". Revista Iberoamericana de Tecnologías del Aprendizaje, IEEE-RITA vol. 9(2), 72-80 et al. 2014.

[9] G. Hernández, "Análisis del uso y manejo de la plataforma Moodle en docentes de matemáticas, para el desarrollo de competencias integrales en estudiantes de primaria". Revista Q, 10 (19). http://dx.doi.org/10.18566/revistaq.v10n19.a01, 2015.

[10] B. Hidalgo Cajo, "Minería de datos en los Sistemas de gestión de Aprendizaje en la Educación Universitaria". Campus Virtuales, ISSN-e 2255-1514, Vol. 7, Nº. 2, 2018.

[11] G. Samaniego, L. Marqués, y M. Gisbert, "El profesorado universitario y el uso de Entornos virtuales de aprendizaje". Campus Virtuales, Vol. IV, Num. 2, pp. 50-58, 2015.

[12] Salinas, "Modelos didácticos en los campus virtuales universitarios: Perfiles metodológicos de los profesores en procesos de enseñanza-aprendizaje en entornos virtuales". Reporte de investigación, Enero, Conferencia Virtual Educa, 2008.

[13] A. I. Oviedo Carrascal, J. Jiménez Giraldo, "Minería de datos educativos: Análisis del desempeño de estudiantes de ingeniería en las pruebas" SABER-PRO. Revista Politécnica, 15(29), 128-140. https://doi.org/10.33571/rpolitec.v15n29a10, 2019.

[14] N. Bedregal-Alpaca, V. Cornejo-Aparicio, J. Zárate-Valderrama, P. Yanque-Churo, "Classification models for determining types of academic risk and predicting dropout in university students". International Journal of Advanced Computer Science and Applications (IJACSA), Volume 11 Issue 1, 2020.

[15] D. Buenaño-Fernández, "Uso de la metodología CRISP-DM para guiar el proceso de minería de datos en LMS". In book: Tecnología, innovación e investigación en los procesos de enseñanza aprendizaje Rosabel Roig-Vila (Ed.) Chapter: Uso de la metodología CRISP-DM para guiar el proceso de minería de datos en LMS. Editors: Ediciones Octaedro, 2016.

[16] S. Hussain, N. Abdulaziz, F. Ba-Alwi & N. Ribata. "Educational Data Mining and Analysis of Students' Academic Performance Using WEKA", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 9, No. 2, February 2018, pp. 447~459, ISSN: 2502-4752, DOI: 10.11591/ijeecs.v9.i2.pp447-459.

[17] J. Salazar-Cardona & J. Triviño-Arbeláez, "Aplicación de learning analytics y educational data mining en una institución de educación superior en Colombia", Revista Ingenierías Universidad de Medellín, 19(36) • Enero-Junio de 2020 • pp. 71-89 • ISSN (en línea): 2248-4094. DOI 10.22395/rium.v19n36a4.

[18] I. Peña, C. Córcoles, & C. Casado, "El profesor 2.0: docencia e investigación desde la red". Revista sobre la sociedad del conocimiento, 3, 1-9. Obtenido de http://ullviu.blog.cat/gallery/12161/12161-68417.pdf, 2006.

[19] B. Martín Galán, & D. Rodríguez Mateos, "La evaluación de la formación universitaria semipresencial y en línea en el contexto del EEES mediante el uso de los informes de actividad de la plataforma Moodle". RIED. Revista Iberoamericana De Educación a Distancia, 15(1), 159–178. https://doi.org/10.5944/ried.1.15.782, 2012.

[20] R. Cristóbal, S. Ventura, E. García," Data mining in course management systems: Moodle case study and tutorial". Computers and Education, 51(1), 368-384. doI: https://doi.org/10.1016/j.compedu.2007.05.016, 2008.

[21] C. Shearer, "el modelo CRISP-DM: el nuevo plan para la minería de datos, almacenamiento de los datos". J 5:13-22, 2000.

[22] J.M. Moine, A.S. Haedo, & S. Gordillo, "Estudio comparativo de metodologías para minería de datos". CACIC 2011 - XVII CONGRESO ARGENTINO DE CIENCIAS DE LA COMPUTACIÓN, 2011.

[23] K. C. Mondal, B. D. Nandy, & A. Baidya, "A Factual Analysis of Improved Python Implementation of Apriori Algorithm". Methodologies and Application Issues of Contemporary Computing Framework, 139–151, 2018.

[24] N. Bedregal-Alpaca, D. Tupacyupanqui-Jaén, V. Cornejo-Aparicio. "Video and Cooperative Work as Didactic Strategies to Enrich Learning and Development of Generic Competences in numerical Methods". XIII Latin American Conference on Learning Technologies (LACLO). 2018. DOI: 10.1109/laclo.2018.00038.

[25] N. Bedregal-Alpaca, D. Tupacyupanqui-Jaen, M. Rodriguez-Quiroz, L. Delgado-Barra, K. Guevara-Puente and O. Sharhorodoska, "Problem-Based Learning with ICT Support: An experience in teaching-learning the concept of derivative," 2019 38th International Conference of the Chilean Computer Science Society (SCCC), Concepcion, Chile, 2019, pp. 1-7. DOI: 10.1109/SCCC49216.2019.8966396.

[26] N. Bedregal-Alpaca, V. Cornejo-Aparicio, A. Padron-Alvarez, E. Castañeda-Huaman, "Design of cooperative activities in teaching-learning university subjects: Elaboration of a proposal", International Journal of Advanced Computer Science and Applications Volume 11, Issue 42020 Article number 0445. DOI 10.14569/IJACSA.2020.0110445.