




Article

Clustering of LMS Use Strategies with Autoencoders

María J. Verdú ^{1,*},[†] , Luisa M. Regueras ^{1,†}, Juan P. de Castro ^{1,†}  and Elena Verdú ^{2,†} 

¹ Higher Technical School of Telecommunications Engineering (ETSIT), Universidad de Valladolid, 47011 Valladolid, Spain; luireg@tel.uva.es (L.M.R.); jpdecastro@tel.uva.es (J.P.d.C.)

² Escuela Superior de Ingeniería y Tecnología, Universidad Internacional de La Rioja, 26006 Logroño, Spain; elena.verdu@unir.net

* Correspondence: marver@tel.uva.es

† These authors contributed equally to this work.

Abstract: Learning Management Systems provide teachers with many functionalities to offer materials to students, interact with them and manage their courses. Recognizing teachers' instructing styles from their course designs would allow recommendations and best practices to be made. We propose a method that determines teaching style in an unsupervised way from the course structure and use patterns. We define a course classification approach based on deep learning and clustering. We first use an autoencoder to reduce the dimensionality of the input data, while extracting the most important characteristics; thus, we obtain a latent representation of the courses. We then apply clustering techniques to the latent data to group courses based on their use patterns. The results show that this technique improves the clustering performance while avoiding the manual data pre-processing work. Furthermore, the obtained model defines seven course typologies that are clearly related to different use patterns of Learning Management Systems.

Keywords: autoencoders; clustering; deep learning; educational data mining; learning management system; unsupervised learning



Citation: Verdú, M.J.; Regueras, L.M.; de Castro, J.P.; Verdú, E. Clustering of LMS Use Strategies with Autoencoders. *Appl. Sci.* **2023**, *13*, 7334. <https://doi.org/10.3390/app13127334>

Academic Editor: Dimitris Mourtzis

Received: 24 May 2023

Revised: 7 June 2023

Accepted: 19 June 2023

Published: 20 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Learning Management Systems (LMS) are increasingly being integrated into traditional universities as a complement to face-to-face teaching. These systems provide teachers with many resources and utilities to offer materials to students, interact with them, and manage their courses. In the same way that some researchers have aimed to detect students' learning profiles from their behavior to provide appropriate recommendations and improve academic achievement [1,2], detecting teachers' instructing styles from their course designs would allow for recommendations and best practices to be provided. However, a literature review suggests that a research gap exists in the fields of learning analytics and educational data mining from the perspective of teachers [3–5].

With the assistance of an intelligent LMS that analyses teaching style from course structure and LMS use patterns, teachers could improve the design of their courses [5]. This improvement is expected to directly impact the learning process and students' outcomes and satisfaction [6]. On the other hand, many national teacher accreditation systems are based on teaching quality programs that evaluate competence in technology-based learning. Therefore, universities need to certify the use of LMSs by teachers. Currently, this task is carried out manually by experts or with overly basic metrics based on the presence or absence of LMS activity. It would be interesting to be able to automatize this certification process.

We see the process of defining teaching styles from LMS data as an unsupervised problem. Supervised methods require a significant quantity of well-labeled training datasets created by significant manual feature engineering efforts [7]. It is very hard to obtain labeled data in this context. In addition, as there is not a generally accepted model of

online teaching styles [8,9], it is infeasible for experts to assign meaningful labels to online courses [10]. Even unsupervised methods require some important manual data engineering in the phase of data pre-processing and transformation. In our previous work on LMS course clustering [11], we lost information in the processes of reducing the dimensionality and discretizing the input features into only three bins (low, medium and high). We grouped in the same bin values that could have fuzzy limits. In the process of reducing dimensionality, the discarding of some dimensions inevitably led to the loss of information. Thus, the primary problem to be solved is how to obtain a low-dimensional model while keeping the most important characteristics of the original data. Motivated by these limitations, as well as by recent work on deep learning techniques, we propose a course classification approach based on deep learning and clustering.

Deep neural networks outperform linear models in many tasks. For unsupervised learning, one recent approach is the use of autoencoders, which provides an informative representation of the data that can be used, for example, for effective clustering [7,12].

Autoencoders are a type of neural network where the input is the same as the output [13]. They are an unsupervised learning technique. Thus, they can be trained with unlabeled data. The autoencoders have two main components: an encoder and a decoder (see Figure 1). The output of the encoder is a latent (or semantic) representation of the input, usually of a lower dimension. The decoder can reconstruct the input using that lower-dimensional representation. Then, autoencoders allow for the dimensions of the input data to be reduced, while extracting their main characteristics.

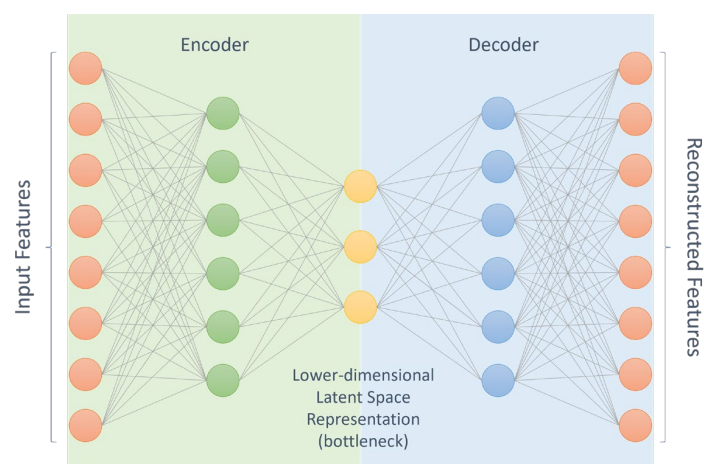


Figure 1. General Autoencoder architecture.

Typical applications of autoencoders include dimensionality reduction [13] and feature extraction [12,14]. Traditional ways to reduce the dimensionality of large datasets are removing variables with a high correlation and/or a high number of missing values, as well as using PCA (Principal Components Analysis). Autoencoders have been shown to be better than PCA as a tool for dimensionality reduction [15–17], although they may involve more computation time and resources than PCA [18]. In fact, autoencoders have also been shown to be better than PCA as a tool for capturing the natural data structure in clustering and obtaining well-defined clustering structures [19]. Other studies have also shown that dimensionality reduction through deep neural networks allows the use of more scalable classification systems while maintaining or improving the accuracy [20,21].

Autoencoders are being used in different fields. For example, Deep Convolutional Autoencoders are used for the unsupervised clustering of seismic data [7], obtaining precisions comparable to those achieved by supervised methods but without the need for labeled data, manual feature engineering and large training sets. In addition, variational autoencoders are used to learn a latent data representation that captures the natural clustering of bank customers' data according to creditworthiness [19]. In addition, autoencoders are successfully used for dimensionality reduction prior to applying k-means clustering to

functional magnetic resonance imaging data [22]. In the literature, there are many other examples of a successful unsupervised approach with autoencoders [23–26].

In the field of learning analytics, neural networks have mainly been used in supervised problems for predicting students' performance, dropout rates and mood [27–33], as well as for classifying students according to a well-known learning style model [1,34]. However, to the best of our knowledge, there is no previous work on latent representations of LMS courses using autoencoders or any other deep learning architecture.

We first use an autoencoder that learns the latent representation of LMS courses (defined as the course structure together with the users' interactions). We then apply the latent data to clustering techniques to group courses based on their use patterns. That is, the autoencoder extracts features automatically from input course data and reduces the features' dimensions. The new dimension-reduced data are used for clustering instead of the original data, with the hypothesis of improving clustering performance compared to that obtained with manual preprocessing [11]. Therefore, in the problem of modelling course typologies according to LMS use, we aim to answer to the following research questions:

- RQ1: Is it possible to avoid the manual data pre-processing work?
- RQ2: Is it possible to improve the clustering performance by reducing dimensionality using deep learning instead of manually transforming data?
- RQ3: Do we obtain a well-defined clustering structure when we start from the latent space?

The remainder of this paper is organized as follows. In the next section, we describe the context and the data mining process implemented in this study, along with the methods used. Next, the results of this study and their analyses are presented. Finally, the last section contains the outcomes and insights regarding future work.

2. Materials and Methods

This section contains a description of the methodology used in this research.

A flow diagram of the proposed methodology is shown in Figure 2. It consists of the following steps:

- Logs acquisition;
- Data preprocessing;
- Dimensionality reduction;
- Classification.

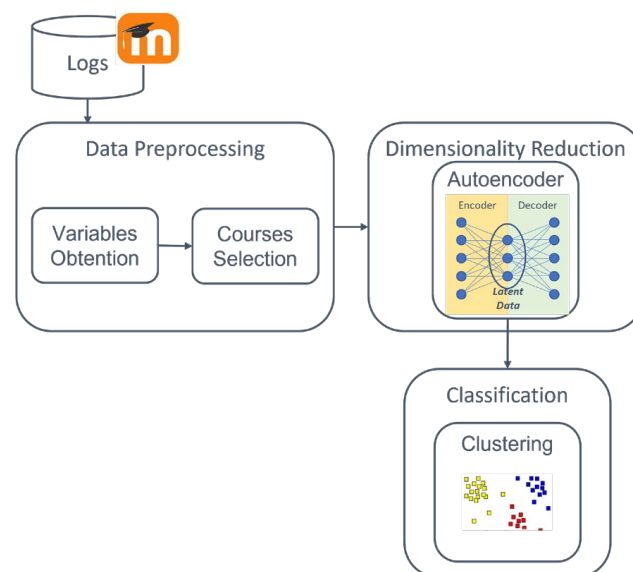


Figure 2. Flow diagram of the proposed methodology.

We used a MySQL database engine for data collection from the educational environment. In addition, we chose R for data processing and as a data mining tool, since we have used it in our previous tasks, and it also offers a wide variety of packages for deep learning [35]. We built the autoencoder using the Keras API in R [36].

2.1. Logs Acquisition

A preliminary phase in the data mining process is data collection. Data were collected from the Moodle LMS corresponding to the virtual campus of a face-to-face university for the 2015–2016 academic year. These data included approximately 2 million records of log-data. For each course, we registered actions and resources about teaching and learning activities of all participants (teachers and students). We used SQL scripts to create a summary table of courses with aggregated information about users and activity indicators.

2.2. Data Preprocessing

We collected data about all official face-to-face courses that have corresponding virtual courses in Moodle. An LMS is used as optional support in face-to-face classes. Each teacher decides how to use the virtual campus, resulting in a variety of uses.

In the first preprocessing stage, courses without students and empty courses were eliminated, leaving a total of 3303 pre-selected courses. Then, we identified the cases with interesting information to maximize efficiency and validity [37]. We selected all cases that met a predetermined criterion of importance, including courses with at least five students enrolled (since this is the minimum number of students required for an optional course to be taught, according to the Academic Management Regulation of the target university). Finally, 3046 courses were selected. We decided to leave in courses with very low use of the LMS to check if the analysis itself could detect them, which is different from other studies [38,39]. The described preprocessing is common to our previous study [11].

In addition to course filtering, we transformed and selected the variables to conduct a sound analysis of Moodle usage patterns. A Moodle course can integrate and configure resources (such as files, links, labels), activities (such as forums, assignments, quizzes, glossaries, workshops, wikis) as well as management tools (such as the event calendar and the gradebook) [40]. From these data, 17 numerical variables were selected to count resources, activities and actions, as indicated in Table 1. All variables about students' interactions were normalized to the number of students enrolled in the course. The role of the actor for the recorded data is also shown. For example, resources and activities are uploaded and created by teachers, while students are responsible for submissions and views. We grouped together data on resources and some activities of very limited use, as detailed in our previous study [11].

Table 1. Description of numerical variables.

Variable Name	Counted Data	Role
Resources	Resources (html, pdf documents)	Teacher
ResourceViews	Resource views or downloads	Student
Forums	Discussion forums	Teacher
ForumNews	Teachers' forum posts	Teacher
ForumInteractions	Students' forum views and posts	Student
Assigns	Assignments	Teacher
AssignSubmissions	Assignment submissions	Student
Quizzes	Quizzes	Teacher
QuizSubmissions	Quiz submissions	Student
AdvActivities	Advanced activities	Teacher
AdvActivitySubmissions	Advanced activity submissions	Student
GradeItems	Gradebook items	Teacher
GradeFeedbacks	Feedbacks of gradebooks	Teacher
GradeAdvanced	Manual or calculated gradebook items	Teacher
BasicInteractions	Entries (glossary, database, chat)	Student
Feedbacks	Feedback activities (surveys)	Teacher
CalendarEvents	Manual calendar events	Teacher

2.3. Dimensionality Reduction

From the course summary table with the 17 variables that describe the activity of the courses, we used autoencoders to reduce the dimensionality and facilitate the analysis. This is where we tried to improve the methods used in our previous study, where we had to pre-process data by (1) pre-selecting the features of interest, avoiding correlations and low-variance features across samples, and (2) discretizing data using K-means clustering [11]. These techniques reduce the complexity of the data and make the analysis easier to understand; however, they suffer from a significant loss of information.

With autoencoders, when we select a number of bottleneck layer nodes that is smaller than the number of original input nodes, we obtain a compressed representation of the input. Therefore, we obtain the desired effect of dimensionality reduction [41]. Thus, the dimension of the latent representation (bottleneck layer, see Figure 1) is one of the main parameters to be set for autoencoders in dimensionality reduction. For clustering, lower dimensionality can lead to a lower clustering accuracy due to a higher reconstruction error. There is not a rule for the choice, so empirical approaches are typically used [7].

To build the autoencoder, other parameters that must be defined are the number of layers or depth for the encoder and decoder, the number of nodes per layer and the loss function (or distance between the compressed representation of the data and the decompressed representation). We must choose these parameters carefully so that the model does not become overfitted; that is, the dimension of the latent representation should be low enough to let the autoencoder learn useful and meaningful characteristics of the input data.

We implemented a fully connected symmetric autoencoder with 4 dense layers (see Figure 3). The number of nodes in each layer, as well as the activation function, were chosen to minimize reconstruction errors. We found the best option to be 11 nodes for the first hidden layer and 6 nodes for the low-dimensional layer (latent space). Moreover, the encoder and decoder used Rectified Linear Units (ReLU) as the activation function and were trained using the Mean Squared Error as the loss function.

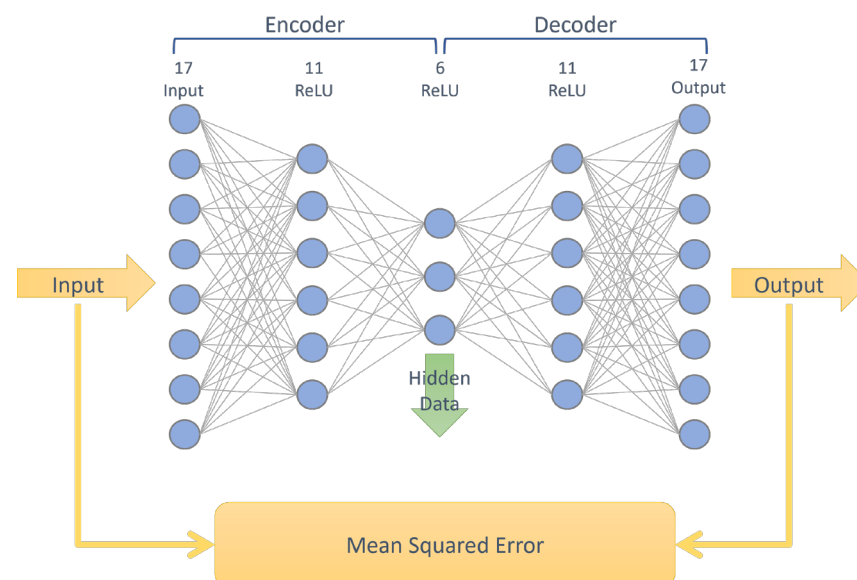


Figure 3. Description of autoencoder architecture layers.

2.4. Clustering

To obtain knowledge from the data and to classify courses, we applied clustering using the K-means algorithm, chosen because of its simplicity, fast convergence and its ease of understanding and visualization [42]. Moreover, it is not necessary to discretize the input values with the consequent loss of information as in other methods, such as LCA (Latent Class Analysis) [43].

The first step to build clusters is to choose the K-value or the number of classes. The silhouette coefficient, gap statistic, elbow method and Canopy are good methods to determine the optimal number of clusters for K-means and provide good results [44]. Subsequently, we measured the performance of the clusters through the following two measures: homogeneity and heterogeneity. The objective is to obtain clusters with low variability within clusters (homogeneity) and a high degree of separation between them (heterogeneity). We used the average distance between clusters as a measure of heterogeneity and the average distance within clusters as a measure of homogeneity [45].

2.5. Research Ethics

In learning analytics, the main ethical issues are related to data ownership and student privacy [46], with the protection of privacy being one of the fundamental values and rights on which ethical artificial intelligence is based [47]. In this study, all human participants' data were fully anonymized before we accessed them. However, in some cases, course identifiers could reveal information about the identity of teachers and students. Therefore, course identifiers were re-codified to minimize possible ethical issues. Since the unit of analysis was the course, there was no potential problem of student identification in this study. In any case, student anonymity was always preserved by removing all personal identifiers from the data. In addition, we did not collect any sensitive data such as racial origin, religious beliefs or health data (according to the Spanish Law of Personal Data Protection).

3. Results and Discussion

As expected, when reducing dimensionality, smaller sizes led to higher reconstruction errors. We empirically found an optimal value for the bottleneck layer that is lower than the dimension obtained from manual data transformation [11]. The autoencoder reduced the input dimension to six features in the latent space, while our previous hand-transformed set consisted of nine variables. As previously mentioned, the defined encoder contains dense layers and the leaky ReLU activation function, and its output layer defines a latent vector size with six nodes. We trained the model for 5000 epochs and calculated the loss and accuracy for each epoch to confirm that it converged. Figure 4 shows loss and accuracy vs. epoch. We can observe that the model converged and how well the autoencoder could reconstruct its input.

From the six latent features, we applied K-means to obtain the different clusters. For this, we determined the optimal number of classes. We calculated the silhouette coefficient and gap statistic, obtaining seven as the optimal number of clusters with both methods. Figure 5 shows the seven different clusters after applying K-means from the six latent (hidden) features in a 2D space.

Once the seven clusters were obtained, we analyzed them to interpretate the obtained results. For this purpose, we studied these seven classes in relation to the users' activity variables described in Table 1, and we checked if it was possible to define a course typology based on these results. Figures 6 and 7 show the normalized mean value of the variables associated with the teachers' and students' activity, respectively. From these results, seven different course typologies were established. Class 2 corresponds to courses with low use of Moodle or inactive courses (type I or Inactive). Class 6 corresponds to courses whose main activities are students viewing content and teachers sending announcements to the forum. In these courses, teachers use the platform as an informative medium or a Web repository (type R or Repository).

Classes 3, 4 and 7 are courses with static content and the use of evaluation based on task assignments. The main difference is that Class 3 matches courses that also have a high use of discussion forums and teacher–student communication, showing a more communicative profile (type C or Communicative). Class 4 has a greater use of assignments and evaluative elements such as manual gradebook items (type E or Evaluative). Class 7, in turn, has a high number of entries in the calendar and in the gradebook; this suggests that the teachers use Moodle as an organizing element and class planner (type P or Planner).

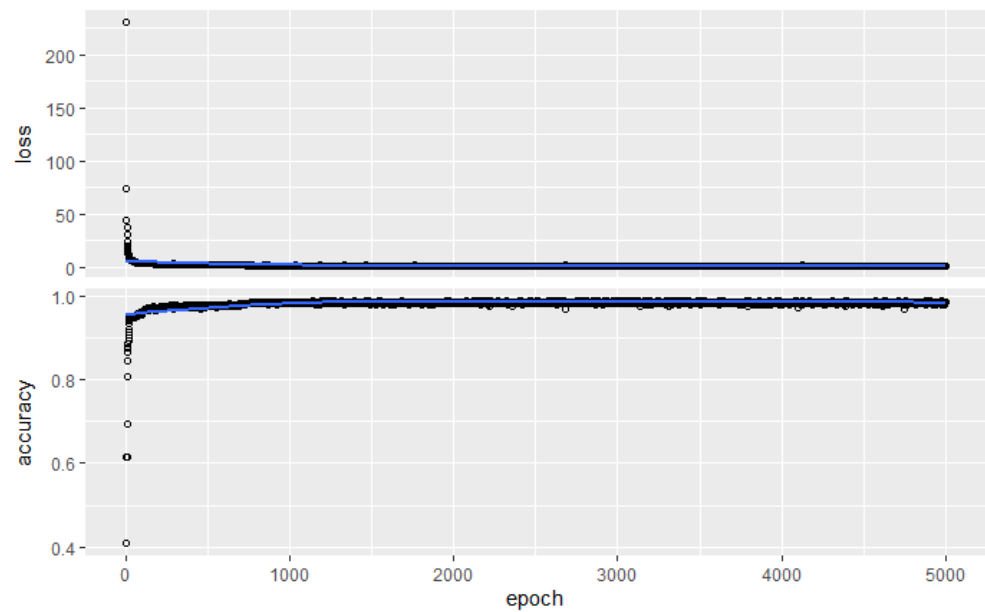


Figure 4. Loss and accuracy vs. epoch.

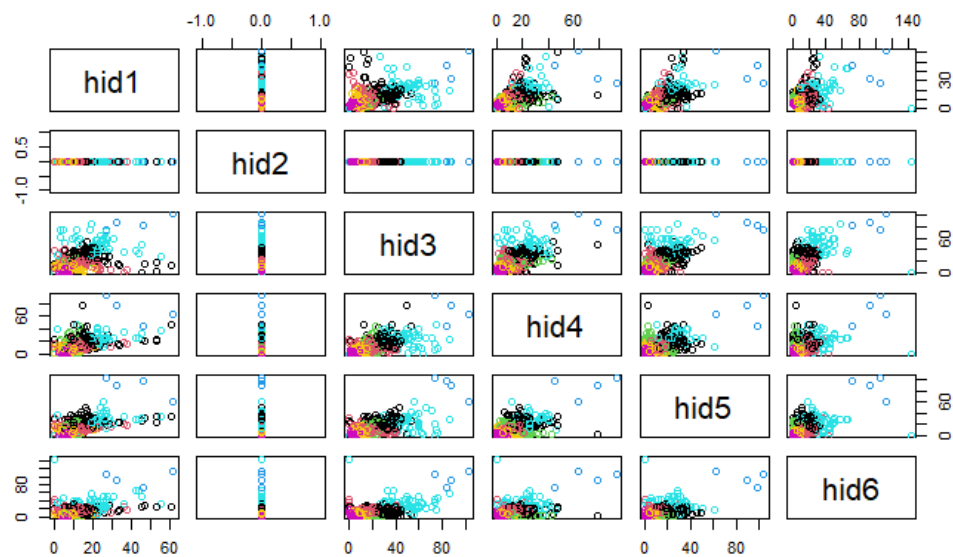


Figure 5. Two-dimensional K-means clustering ($k = 7$) from latent features (hid1 to hid6).

On the other hand, Class 1 corresponds to courses with a mixed use of Moodle tools; these are courses with large amounts of content, forums and assignments, combined with quizzes and evaluative elements (type B or Balanced). Finally, Class 5 consists of courses with very high and wide use of all types of activities on Moodle, including advanced activities (type V or adVanced).

Table 2 shows a summary of the seven course typologies described previously, indicating their main features.

At first, we can observe that these seven classes are similar to the six typologies defined in our previous study [11], although some types have disappeared (like Submission -S) and there are new types. Our aim is to see if there is any connection between them. Figure 8 represents a mosaic plot, where we can visually compare the different groups and obtain a general idea of how they are related to each other. We can observe that the adVanced courses were previously classified as B because the variable associated with advanced activities had been eliminated in the preprocessing of the data, since it was zero in most

courses. Something similar occurs with the Planner courses, which were not present in the previous study when removing the CalendarEvents variable and were spread out across different classes. On the other hand, the Inactive class mainly corresponds to Inactive courses in the previous work, although it also includes some courses previously classified as Submission. The classes Repository, Evaluative, Communicative and Balanced had a more diffuse relationship with the previous classification. Analyzing some of these courses, we observed that some of the variables were near the thresholds separating one class from the other.

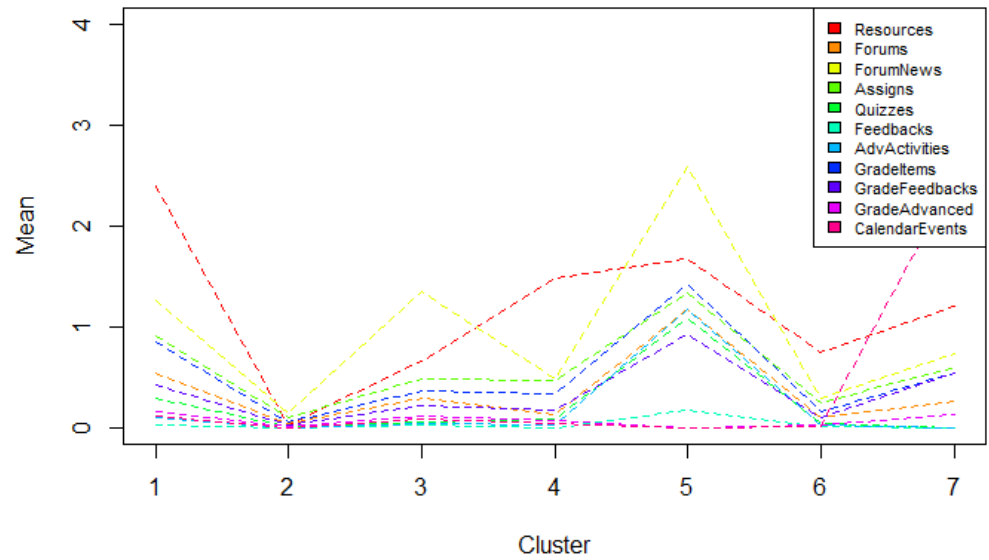


Figure 6. Description of clusters: teachers' actions.

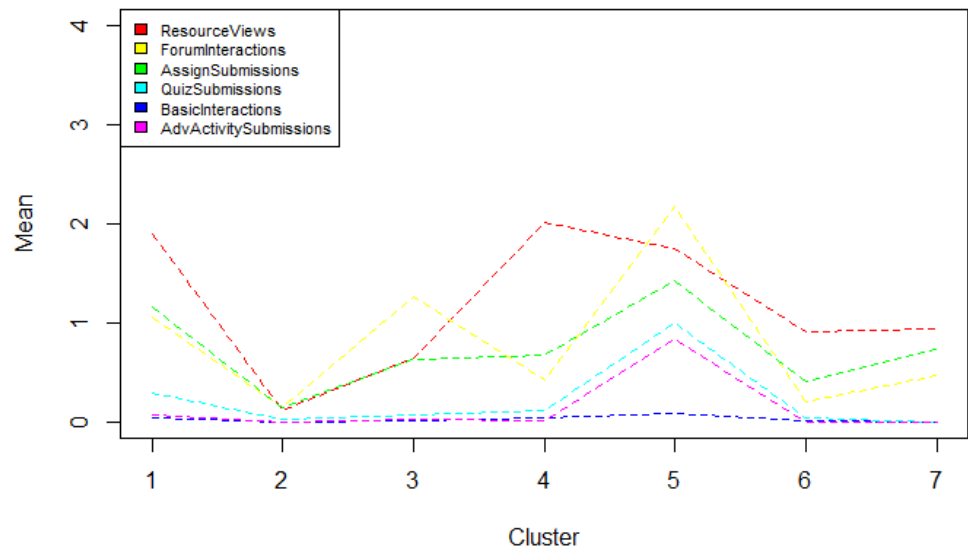


Figure 7. Description of clusters: students' actions.

In our previous study [11], we removed features with a high correlation and/or low variability of values. We discarded information and lost classes, such as the typology described as “Planner course”. Moreover, we discretized the variables in three levels (low, medium, high). That is, we lost information about the real values. This led to some values not being properly considered when being on the threshold between one level and another. Some of these shortcomings may be overcome with this new analysis technique.

Table 2. Description of course typologies.

Typology	Description
Inactive—I	Low use of Moodle
Repository—R	Content and news
Communicative—C	Content, assignments and teacher–student interactions
Evaluative—E	Content, assignments and evaluative elements
Planner—P	Content, assignments and very high use of calendar events
Balanced—B	Heavy and balanced use of Moodle tools
adVanced—V	High and wide use of Moodle tools, including advanced tools

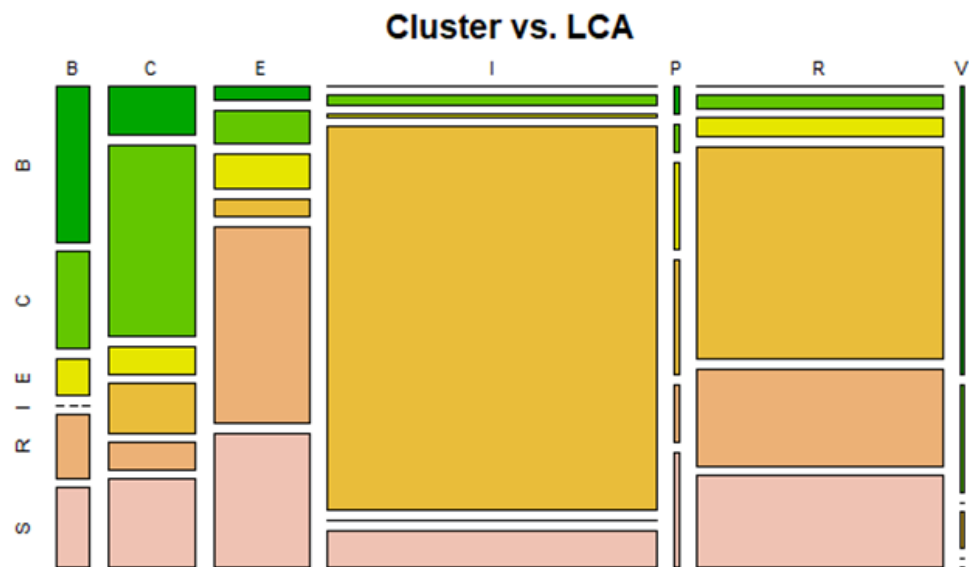


Figure 8. Mosaic between classes' LCA and autoencoder clustering.

To validate the new analysis, we measured the homogeneity and heterogeneity (as the average distance within clusters and between clusters [45], respectively) to analyze the performance of clusters in relation to the previous work [11].

Table 3 shows the results obtained. We can observe how K-means with autoencoders provides the best value for both homogeneity (the lowest value is the best) and heterogeneity (the highest value is the best). That is, it offers the most homogeneous clusters, and it is the most effective in separating clusters.

Table 3. Analysis of cluster performance: autoencoder vs. manual preprocessing. Adapted from Ref. [11].

Method	Homogeneity	Heterogeneity
K-means + Autoencoders	0.5354	2.2339
K-means + Manual Preprocessing	0.6548	2.0887
LCA + Manual Preprocessing	0.8097	2.2049

4. Conclusions and Future Work

This paper proposes a method to infer a typology model of courses out of the use of the LMS by teachers and students. We have used an autoencoder to obtain a latent representation of LMS courses and perform clustering in the latent space instead of the input data space, avoiding the complex manual data pre-processing (answering RQ1) and improving clustering performance (answering RQ2), with more homogeneous and differentiated clusters than those obtained with manual pre-processing. Moreover, we have obtained seven typologies of courses that fit the following seven well-defined usage pat-

terns: Inactive, Repository, Communicative, Evaluative, Planner, Balanced and adVanced (answering RQ3).

A limitation of this work is that we used empirical approaches to select the different parameters of autoencoders (activation function, loss function, number of hidden layers, etc.). The use of metaheuristic hyperparameter optimizers may improve the results. In a future work, we plan to iteratively fine-tune the deep network parameters for latent space optimization and an improvement of the clustering effectiveness [7]. We also plan to include in the autoencoder constraints about the distance between latent data and cluster centers to obtain a more stable and compact representation, which is more suitable for clustering, as suggested by some authors [7,41]. In addition, we could compare the efficiency of different types of autoencoders, such as denoising autoencoders [48,49], contractive autoencoders [12] or sparse autoencoders [50,51].

Finally, we would like to exploit the potential of multimodal data [52], as suggested from a systematic review [3], for example, by including teachers' psychological and context data, as well as self-reported data, to provide feedback to the classification system. In addition to enriching the deep learning system, we will try to engage teachers in the learning analytics process. This is one of the last challenges identified by the learning analytics research community [53].

Author Contributions: Conceptualization, M.J.V., L.M.R., J.P.d.C. and E.V.; Data curation, M.J.V., L.M.R., J.P.d.C. and E.V.; Formal analysis, M.J.V., L.M.R., J.P.d.C. and E.V.; Investigation, M.J.V., L.M.R., J.P.d.C. and E.V.; Resources, E.V.; Software, M.J.V., L.M.R., J.P.d.C. and E.V.; Validation, M.J.V., L.M.R., J.P.d.C. and E.V.; Visualization, M.J.V., L.M.R. and J.P.d.C.; Writing—original draft, M.J.V., L.M.R., J.P.d.C. and E.V.; Writing—review and editing, M.J.V., L.M.R., J.P.d.C. and E.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data cannot be shared publicly because of privacy restrictions. Some data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Muhammad, B.A.; Qi, C.; Wu, Z.; Ahmad, H.K. GRL-LS: A Learning Style Detection in Online Education Using Graph Representation Learning. *Expert Syst. Appl.* **2022**, *201*, 117138. [[CrossRef](#)]
2. Rincon-Flores, E.G.; Lopez-Camacho, E.; Mena, J.; Olmos, O. Teaching through Learning Analytics: Predicting Student Learning Profiles in a Physics Course at a Higher Education Institution. *Int. J. Interact. Multimed. Artif. Intell.* **2022**, *7*, 82–89. [[CrossRef](#)]
3. Celik, I.; Dindar, M.; Muukkonen, H.; Järvelä, S. The Promises and Challenges of Artificial Intelligence for Teachers: A Systematic Review of Research. *TechTrends* **2022**, *66*, 616–630. [[CrossRef](#)]
4. Manhiça, R.; Santos, A.; Cravino, J. The Use of Artificial Intelligence in Learning Management Systems in the Context of Higher Education: Systematic Literature Review. In Proceedings of the 2022 17th Iberian Conference on Information Systems and Technologies (CISTI), Madrid, Spain, 22–25 June 2022; pp. 1–6.
5. Bennacer, I.; Venant, R.; Iksal, S. A Behavioral Model to Support Teachers' Self-Assessment and Improve Their LMS Mastery. In Proceedings of the 22nd IEEE International Conference on Advanced Learning Technologies, Bucharest, Romania, 1–4 July 2022.
6. D'Mello, S.K. Emotional Learning Analytics. In *Handbook of Learning Analytics*; Society for Learning Analytics Research: Edmonton, AB, Canada, 2017; pp. 115–127. ISBN 978-0-9952408-0-3.
7. Mousavi, S.M.; Zhu, W.; Ellsworth, W.; Beroza, G. Unsupervised Clustering of Seismic Signals Using Deep Convolutional Autoencoders. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1693–1697. [[CrossRef](#)]
8. Rodríguez, A.L.; Fahara, M.F.; Tecnológico, C.; León, N. Online Teaching Styles: A Study in Distance Education. *Int. J. Univ. Teach. Fac. Dev.* **2010**, *1*, 1–14.
9. Vikas, S.; Mathur, A. An Empirical Study of Student Perception towards Pedagogy, Teaching Style and Effectiveness of Online Classes. *Educ. Inf. Technol.* **2022**, *27*, 589–610. [[CrossRef](#)] [[PubMed](#)]
10. Regueras, L.M.; Verdú, M.J.; de Castro, J.-P. A Rule-Based Expert System for Teachers' Certification in the Use of Learning Management Systems. *Int. J. Interact. Multimed. Artif. Intell.* **2022**, *7*, 75–81. [[CrossRef](#)]

11. Regueras, L.M.; Verdú, M.J.; Castro, J.D.; Verdú, E. Clustering Analysis for Automatic Certification of LMS Strategies in a University Virtual Campus. *IEEE Access* **2019**, *7*, 137680–137690. [CrossRef]
12. Bank, D.; Koenigstein, N.; Giryes, R. Autoencoders. *arXiv* **2021**, arXiv:2003.05991.
13. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
14. Ardelean, E.-R.; Coporăie, A.; Ichim, A.-M.; Dînşoreanu, M.; Mureşan, R.C. A Study of Autoencoders as a Feature Extraction Technique for Spike Sorting. *PLoS ONE* **2023**, *18*, e0282810. [CrossRef]
15. Hinton, G.E.; Salakhutdinov, R.R. Reducing the Dimensionality of Data with Neural Networks. *Science* **2006**, *313*, 504–507. [CrossRef]
16. Casella, M.; Dolce, P.; Ponticorvo, M.; Marocco, D. Autoencoders as an Alternative Approach to Principal Component Analysis for Dimensionality Reduction. An Application on Simulated Data from Psychometric Models. In Proceedings of the Third Symposium on Psychology-Based Technologies (PSYCHOBIT2021), Naples, Italy, 4–5 October 2021; Volume 3100.
17. Mantripragada, K.; Dao, P.D.; He, Y.; Qureshi, F.Z. The Effects of Spectral Dimensionality Reduction on Hyperspectral Pixel Classification: A Case Study. *PLoS ONE* **2022**, *17*, e0269174. [CrossRef] [PubMed]
18. Fournier, Q.; Aloise, D. Empirical Comparison between Autoencoders and Traditional Dimensionality Reduction Methods. In Proceedings of the 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), Sardinia, Italy, 3–5 June 2019; pp. 211–214.
19. Mancisidor, R.A.; Kampffmeyer, M.; Aas, K.; Jenssen, R. Learning Latent Representations of Bank Customers with the Variational Autoencoder. *Expert Syst. Appl.* **2021**, *164*, 114020. [CrossRef]
20. Bobadilla, J.; Ortega, F.; Gutiérrez, A.; Alonso, S. Classification-Based Deep Neural Network Architecture for Collaborative Filtering Recommender Systems. *Int. J. Interact. Multimed. Artif. Intell.* **2020**, *6*, 68. [CrossRef]
21. Wen, T.; Zhang, Z. Deep Convolution Neural Network and Autoencoders-Based Unsupervised Feature Learning of EEG Signals. *IEEE Access* **2018**, *6*, 25399–25410. [CrossRef]
22. Spencer, A.P.C.; Goodfellow, M. Using Deep Clustering to Improve fMRI Dynamic Functional Connectivity Analysis. *NeuroImage* **2022**, *257*, 119288. [CrossRef]
23. Amrutha, E.; Arivazhagan, S.; Jebarani, W.S.L. Deep Clustering Network for Steganographer Detection Using Latent Features Extracted from a Novel Convolutional Autoencoder. *Neural Process. Lett.* **2022**. [CrossRef]
24. Shinde, K.; Itier, V.; Mennesson, J.; Vasiukov, D.; Shakoor, M. Dimensionality Reduction through Convolutional Autoencoders for Fracture Patterns Prediction. *Appl. Math. Model.* **2023**, *114*, 94–113. [CrossRef]
25. Hurtado, S.; García-Nieto, J.; Popov, A.; Navas-Delgado, I. Human Activity Recognition From Sensorised Patient’s Data in Healthcare: A Streaming Deep Learning-Based Approach. *Int. J. Interact. Multimed. Artif. Intell.* **2023**, *8*, 23–37. [CrossRef]
26. De Oliveira, H.; Martin, P.; Ludovic, L.; Vincent, A.; Xiaolan, X. Explaining Predictive Factors in Patient Pathways Using Autoencoders. *PLoS ONE* **2022**, *17*, e0277135. [CrossRef] [PubMed]
27. Basnet, R.B.; Johnson, C.; Doleck, T. Dropout Prediction in Moocs Using Deep Learning and Machine Learning. *Educ. Inf. Technol.* **2022**, *27*, 11499–11513. [CrossRef]
28. Liu, S.; Liu, S.; Liu, Z.; Peng, X.; Yang, Z. Automated Detection of Emotional and Cognitive Engagement in MOOC Discussions to Predict Learning Achievement. *Comput. Educ.* **2022**, *181*, 104461. [CrossRef]
29. Moridis, C.N.; Economides, A.A. Prediction of Student’s Mood during an Online Test Using Formula-Based and Neural Network-Based Method. *Comput. Educ.* **2009**, *53*, 644–652. [CrossRef]
30. Tomasevic, N.; Gvozdenovic, N.; Vranes, S. An Overview and Comparison of Supervised Data Mining Techniques for Student Exam Performance Prediction. *Comput. Educ.* **2020**, *143*, 103676. [CrossRef]
31. Sarwat, S.; Ullah, N.; Sadiq, S.; Saleem, R.; Umer, M.; Eshmawi, A.A.; Mohamed, A.; Ashraf, I. Predicting Students’ Academic Performance with Conditional Generative Adversarial Network and Deep SVM. *Sensors* **2022**, *22*, 4834. [CrossRef]
32. Tao, T.; Sun, C.; Wu, Z.; Yang, J.; Wang, J. Deep Neural Network-Based Prediction and Early Warning of Student Grades and Recommendations for Similar Learning Approaches. *Appl. Sci.* **2022**, *12*, 7733. [CrossRef]
33. Aljaloud, A.S.; Uliyan, D.M.; Alkhalil, A.; Elrhman, M.A.; Alogali, A.F.M.; Altameemi, Y.M.; Altamimi, M.; Kwan, P. A Deep Learning Model to Predict Student Learning Outcomes in LMS Using CNN and LSTM. *IEEE Access* **2022**, *10*, 85255–85265. [CrossRef]
34. Zhang, H.; Huang, T.; Liu, S.; Yin, H.; Li, J.; Yang, H.; Xia, Y. A Learning Style Classification Approach Based on Deep Belief Network for Large-Scale Online Education. *J. Cloud Comput.* **2020**, *9*, 26. [CrossRef]
35. Ghatak, A. *Deep Learning with R*; Springer: Berlin/Heidelberg, Germany, 2019; ISBN 978-981-13-5849-4.
36. Allaire, J.; Chollet, F. Keras: R Interface to “Keras”; R Package Version 2.8.0.9000. 2022. Available online: <https://tensorflow.rstudio.com/> (accessed on 18 June 2023).
37. Palinkas, L.A.; Horwitz, S.M.; Green, C.A.; Wisdom, J.P.; Duan, N.; Hoagwood, K. Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Adm. Policy Ment. Health Ment. Health Serv. Res.* **2015**, *42*, 533–544. [CrossRef] [PubMed]
38. Whitmer, J.; Nuñez, N.; Harfield, T.; Forteza, D. *Patterns in Blackboard Learn Tool Use: Five Course Design Archetypes*; Blackboard: Washington, DC, USA, 2016; Available online: https://www.blackboard.com/sites/default/files/resource/pdf/Bb_Patterns_LMS_Course_Design_r5_tcm136-42998.pdf (accessed on 18 June 2023).

39. Park, Y.; Yu, J.H.; Jo, I.-H. Clustering Blended Learning Courses by Online Behavior Data: A Case Study in a Korean Higher Education Institute. *Internet High. Educ.* **2016**, *29*, 1–11. [[CrossRef](#)]
40. Cole, J.; Foster, H. *Using Moodle—Teaching with the Popular Open Source Course Management System*, 2nd ed.; O’Reilly: Beijing, China, 2007.
41. Song, C.; Liu, F.; Huang, Y.; Wang, L.; Tan, T. Auto-Encoder Based Data Clustering. In Proceedings of the Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Havana, Cuba, 20–13 November 2013; Ruiz-Shulcloper, J., Sanniti di Baja, G., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 117–124.
42. Jo, I.-H.; Park, Y.; Lee, H.; Song, J.; Kang, S. Clustering Analysis of Academic Courses Based on LMS Usage Levels and Patterns: Gaussian Mixture Model, K-Means Clustering and Hierarchical Clustering. In Proceedings of the Fourth International Conference on Data Analytics, Nice, France, 19–24 July 2015; pp. 130–137.
43. Rindskopf, D. Latent Class Analysis. In *The Sage Handbook of Quantitative Methods in Psychology*; Sage Publications Ltd.: Thousand, CA, USA, 2009; pp. 199–215. ISBN 978-1-4129-3091-8.
44. Yuan, C.; Yang, H. Research on K-Value Selection Method of K-Means Clustering Algorithm. *Multidiscip. Sci. J.* **2019**, *2*, 226–235. [[CrossRef](#)]
45. Charrad, M.; Ghazzali, N.; Boiteau, V.; Niknafs, A. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *J. Stat. Softw.* **2014**, *61*, 1–36. [[CrossRef](#)]
46. Slade, S.; Prinsloo, P. Learning Analytics: Ethical Issues and Dilemmas. *Am. Behav. Sci.* **2013**, *57*, 1509–1528. [[CrossRef](#)]
47. Rivero, A.J.L.; Beato, M.E.; Martínez, C.M.; Vázquez, P.G.C. Empirical Analysis of Ethical Principles Applied to Different AI Uses Cases. *Int. J. Interact. Multimed. Artif. Intell.* **2022**, *7*, 105.
48. Vincent, P.; Larochelle, H.; Bengio, Y.; Manzagol, P.-A. Extracting and Composing Robust Features with Denoising Autoencoders. In Proceedings of the 25th International Conference on Machine Learning, New York, NY, USA, 5 July 2008; Association for Computing Machinery: New York, NY, USA; pp. 1096–1103.
49. Vincent, P.; Larochelle, H.; Lajoie, I.; Bengio, Y.; Manzagol, P.-A. Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. *J. Mach. Learn. Res.* **2010**, *11*, 3371–3408.
50. Makhzani, A.; Frey, B. K-Sparse Autoencoders. *arXiv* **2014**, arXiv:1312.5663.
51. Zeng, N.; Zhang, H.; Song, B.; Liu, W.; Li, Y.; Dobaie, A.M. Facial Expression Recognition via Learning Deep Sparse Autoencoders. *Neurocomputing* **2018**, *273*, 643–649. [[CrossRef](#)]
52. Chango, W.; Lara, J.A.; Cerezo, R.; Romero, C. A Review on Data Fusion in Multimodal Learning Analytics and Educational Data Mining. *WIREs Data Min. Knowl. Discov.* **2022**, *12*, e1458. [[CrossRef](#)]
53. Kollom, K.; Tammets, K.; Scheffel, M.; Tsai, Y.-S.; Jivet, I.; Muñoz-Merino, P.J.; Moreno-Marcos, P.M.; Whitelock-Wainwright, A.; Calleja, A.R.; Gasevic, D.; et al. A Four-Country Cross-Case Analysis of Academic Staff Expectations about Learning Analytics in Higher Education. *Internet High. Educ.* **2021**, *49*, 100788. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.