A Neural Network-Based System to Predict Early MOOC Dropout

https://doi.org/10.3991/ijep.v12i5.33779

Soukaina Sraidi, El Miloud Smaili, Salma Azzouzi, Moulay El Hassan Charaf(✉)
Ibn Tofail University, Kenitra, Morocco
my.charaf@uit.ac.ma

Abstract—In recent years, the MOOC (Massively Open Online Courses) revolution has transformed the distance learning landscape. Based on the distribution of educational content, this type of education is expected to undergo the same revolution as all the traditional sectors of content and service sales, such as music, video and commerce, due to the emergence of new technologies. However, the completion rate remains a key measure of a MOOC's success, as the number of students enrolling in a MOOC typically drops during the course. This rate can reach 2–10% at the end of the course. Therefore, dropout prediction is an excellent way to identify at-risk students and make timely decisions. In this study, a prediction model is developed using one of the most widely used methods, the artificial neural network (ANN). As a result, our model can be considered as an optimal option in terms of accuracy and suitability for predicting dropouts in MOOC.

Keywords—dropout, MOOC, neural network, prediction

1 Introduction

Already today, distance learning has received a great deal of attention, and it is likely that MOOC (Massively Open Online Courses) will coexist with traditional face-to-face education in the future. In fact, the use of virtual modes of learning has grown rapidly due to technological developments that provide learners around the world with an interactive and productive space for knowledge sharing [1].

At the time of the COVID-19 pandemic, when online education became the exclusive mode of learning in institutions around the world, researchers were faced with an unexpected challenge: the transformation of the educational landscape through the incorporation of technological tools that had previously remained inert or undiscovered by students and instructors [2]. However, it has also been identified as a real opportunity to collect massive electronic data to track student educational performance [3].

Similarly, many online learners are attracted to the benefits of the online learning environment and many universities are now adopting the online learning environment provided by MOOC as an alternative to the traditional learning environment in an effort to reduce the barriers (e.g. health, financial, geographic...) associated with traditional learning. Furthermore, a variety of learners may enroll and participate in the same course. Thus, participants in MOOC typically have heterogeneous levels of knowledge as well as diverse objectives, geographical origins and motivations, causing high drop-out rates during and after the course among participants.

In this case, distance learning can be a very useful alternative method for mature, self-disciplined students, but it is not suitable for more dependent students. An online program requires students to be well-organized, self-motivated, and have excellent time management skills to keep up with the course schedule. In addition, rewarding students with incentives can also be a great opportunity to promote distance learning.

In addition, students often use digital tools (in particular, social media) to interact with their teachers. Thus, researchers are able to analyze the traces left by learners through their interactions with these environments. In this context, big data analytics, data mining, and learning analytics can identify student characteristics, assess learning processes, predict future performance, and quickly identify potential problems [3].

Many studies have attempted to predict the success or failure of students based on their traces using different algorithms. The basic idea is to use a set of observations (traces) and a response variable to determine whether a student has dropped out. Once the model has been trained, it can automatically predict the outcome of a new observation (dropout or not). The most commonly used methods for building predictive models are tree and linear approaches (e.g. decision trees, generalized linear models). Although a large number of studies have been conducted to identify the reasons for dropouts, little scientific research has focused on the use of machine learning algorithms. For example, neural networks can uncover knowledge based on the nature of variables obtained from students' behavior to predict their likelihood of dropping out.

Throughout this paper, we propose to collect data on learners' interactions in a MOOC over the course of a week. The data is then analyzed using deep learning techniques to develop a prediction model. Based on the platform data and student interactions, the model is able to predict for the coming week whether students will stay in the course or leave. This information will help decision makers and educators understand learners' needs and make timely adjustments.

The remainder of the paper is organized as follows: Section 2 presents the problem statement and a review of some related work. In Section 3, we discuss the mate-riels and explain our method. Then, we detail the main results in Section 4, and finally, we present some concluding remarks in Section 5. A conclusion and future work are given in Section 6.

2 Background theory and literature review

2.1 Problem statement

Today, MOOC are becoming increasingly popular among educators as a flexible education model that allows learners to interact despite distance and attendance issues. However, this model faces challenges such as increasing dropout rates between registration and course completion. This raises the following two research questions (RQ):

- **RQ1**: What are the characteristics of changes in learning progression associated with students who drop a MOOC course?
- **RQ2**: Based on a set of characteristics of a student's learning progression, can we predict that the student will eventually drop out of school in the coming weeks?

Therefore, answering these research questions is crucial to: (1) Provide personalized content based on the needs of individual learners; (2) Extract and determine the rate of learner engagement over the course of each week; (3) Identify groups of learners who share the same characteristics; (4) Keep learners in virtual contact by conducting online discussions; (5) Predict when a student may drop out of the MOOC; and (6) Assist in the development of an intervention plan.

2.2 Related work

Recently, many studies have been conducted on the factors that promote learning in MOOC. These works mainly focus on the characteristics of the MOOC packaging, the demographics of the learners or the capacity for self-regulation. One line of research dedicated to MOOC aims to address the problem of dropout rates.

To this end, the authors of [4] explain how to use student performance data and statistical analysis of online surveys to reveal patterns that can help reduce MOOC dropout rates. Thus, it can encourage learners to both participate in online courses and create their learning plans. The main findings of this study are that by applying innovative teaching methods in extenuating circumstances of distance learning (such as the case of COVID-19), students' spatial conceptions improve, overcoming the lack of a physical learning space. Another study [5] confirms that large number of participants do not complete their studies, resulting in high dropout rates. This study advocates self-regulated learning in MOOC and describes how students can develop these skills through this form of instruction by assessing their cognitive abilities. In addition, the authors of [6] point out that MOOC have higher dropout rates. They conclude that MOOC completion rates will decline unless MOOC platforms are revised to allow interactivity and collaboration.

Furthermore, the authors of [7] suggest reversing the traditional teaching sequence. In the proposed model, teaching begins with practical work, followed by tutorials, and ends with lectures to allow time for questions and discussion. This method is seen as an effective remedial approach to revitalize students, who sometimes feel overwhelmed. As a result, the method significantly reduces failure rates. Moreover, the authors of [8] mention that the use of video conferencing and social networking tools, YouTube channels [9] or educational robotics [10] has led to many ideas, but no consensus has been developed and no common method has been adopted by the majority of teachers. The authors point out that in times of uncertainty (such as a pandemic), it is important to ensure that students can acquire the required knowledge by making the learning process as seamless as possible for all involved.

In general, analyses take little or no account of the social contexts within MOOC and the reasons for their use. In this context, other work has explored algorithmic methods, based on tree and neural approaches, to predict student success or failure from their traces. Regarding tree-based approaches, the authors of [11] use the model of decision trees constructed by partitioning the data set (student traces) according to the value of an input feature. The process is repeated recursively for each subset keeping the variance of the response variable (dropout or not) low until a stopping rule is satisfied or each partition contains only one data point.

Neural networks are able to solve the dropout prediction by performing automatic extraction of relevant features without any human intervention. Furthermore, the authors of [12] propose a hybrid model based on 35 variables that have been filtered from the statistical analysis. For the studied module, these variables are highly correlated with students' academic performance during the first semester of 2020–2021. Their neural network architecture consists of 36 input layers, one hidden layer, and one output layer. The model was trained on the 70% of participants to predict the score of the remaining 30%. With their method, the data fitting process resulted in a 100% success rate in terms of the association of the independent variables with the dependent variable (score). Other studies [13–16] highlight the potential value of using multilayer neural networks to predict MOOC learner dropout. These studies have mainly used various machine learning techniques, different types of data or different data sources. A comparative study between our approach and these works is presented in Section 5. In summary, the use of these approaches illustrates the role of artificial intelligence (AI) as a key component to improve education in the future, not only through dynamic enhancement of MOOC content for students and instructors, but also through the provision of real-time learning indicators for professors and students to reduce dropouts during the learning process.

Finally, this paper is a continuation of previous work [17–19] in which we propose solutions for predicting learner dropout and improving adaptive learning systems using optimization algorithms.

3 Materials and methods

3.1 Overview

Figure 1 describes the workflow of our predictive model.

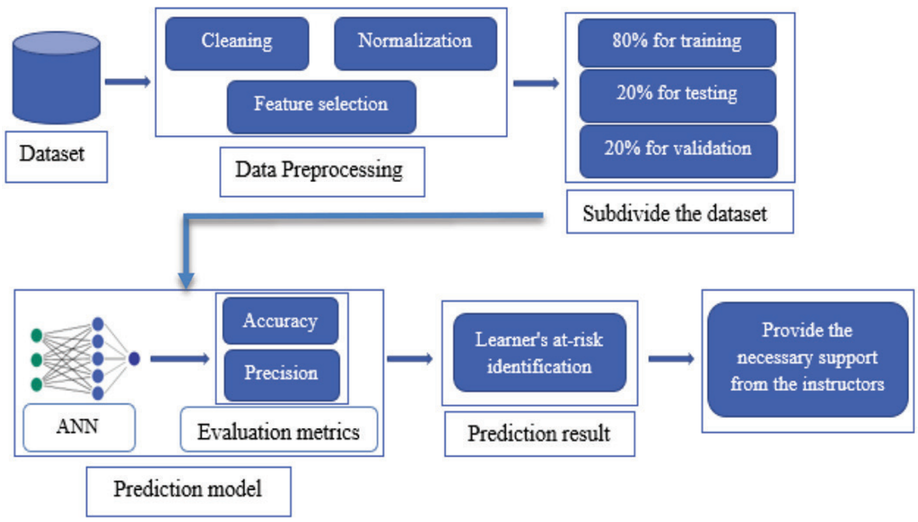


Fig. 1. Workflow of our predictive model

The goal is to extract and determine the rate of learner engagement over the course of each week using a deep learning technique by focusing on learners' interactions with the virtual learning environment (VLE) instead of using characteristics such as student demographic information, gender, or degree [20,21].

3.2 Data collection

The online higher education sector has seen a massive increase in the amount of student data. For example, Open University Learning Analytics (OULAD) contains a subset of OU (Open University) student data collected in 2013 and 2014. In this study, we aim to evaluate the performance of the dropout prediction model based on the 2013 OULAD dataset. In addition, it is relevant to note that the dataset contains demographic information as well as aggregate clickstream data from student interactions in the VLE environment. This data can provide insight into student behavior from their interactions.

Table 1. The OULAD database summary

Data File	No of Records	Description	Attributes
Courses	22	Course information	code_module, code_presentation, module_presentation_length
Student-Info	32,593	Provides demographic details of the student	code_module, code_presentation, id_student, gender, region, highest_education, imd_band, age_band, num_of_prev_attempts, studied_credits, disability, final_result
Student-Registration	32,593	Refers to registration for a course presentation	code_module, code_presentation, id_student, date_registration, date_unregistration
Assessments	206	Provides assessments of courses presentation	code_module, code_presentation, id_assessment, assessment_type, date, weight
Student-Assessment	173,912	Provides assessments submitted by students	id_assessment, id_student, date_submitted, is_banked, score
VLE	6,364	Refers to the online learning resources and materials	id_site, code_module, code_presentation, activity_type, week_from, week_to
Student-VLE	10,655,280	Describes the student interaction with the VLE resources	code_module, code_presentation, id_student, id_site, date, sum_click

The dataset contains information on 22 courses, 32,593 students, their assessment results, and logs of their interactions with the VLE summarized in daily summaries of student clicks (10,655,280 entries):

- Demographics: basic student information including age, gender, region, prior education.
- Performance reflects the results and achievements of students during their studies at Open University (OU).
- The learning behavior is the log of student activities in the VLE.

A summary of the OULAD data is presented in Tables 1 and 2 [16]. In the following paragraphs, we present a detailed description of the steps involved in the realization of our prediction model.

- The first step is to integrate and clean the data. The purpose of this step is to determine if redundant information exists or if empty fields or data may affect the prediction process.

– The feature extraction process is a crucial step in determining the importance of variables as well as determining the input layer predictors for our neural network model.
– In the data modeling process, a deep learning (DL) based machine learning technique will be used. This technique is designed to find information about the variables that predict course dropout in MOOC.
– Prediction methods were used to assess the degree of causality between the characteristics.

Table 2. OULAD dataset overview

Feature	Description
id_student	Learner identification number
Gender	Learner gender
Region	Learning geographic area
highest_education	Learner educational level
imd_band	Socio-economic indicator measure of student economic level
age_band	Learner age
Disability	Indicator of student disability
sum_click—	the number of times the student interacted with the material
id_site	the VLE material identification number
activity_type	the role associated with the module material

3.3 Prediction model

Artificial neurons can be defined as structures that receive stimuli (number *n*) from other neurons or from the outside and interconnect them by means of real values called synaptic coefficients or synaptic weights. The weight of these synapses can be either positive (so-called excitatory synapses) or negative (inhibitory synapses). The weighted sum of the input values of the neural network is calculated on the basis of the weights assigned to each of the input functions presented in (1).

$$y = w_1 x_1 + w_2 x_2 + w_3 x_3 + b \tag{1}$$

Where:

– w: the weight of the variable.
– x: neural network entry variable.
– y: output function of the neuronal.
– b: bias.

To improve the model, we will use the activation function called sigmoid. The function converts an output z into a probability $a(z)$ as follows:

$$a(z) = \frac{1}{1 + e^{-z}} \tag{2}$$

Therefore, the function will return values close to zero to predict a learner who is likely to drop out of the MOOC. For example, for $z = 0.4$, $a(z) = 0.8$, which means that the learner has an 80% probability of continuing the course. On the other hand, the closer the probability is to zero, the more the learner is considered at risk.

3.4 Performance analysis (loss function vs. optimizer)

A loss function is used to quantify the errors made by a model. In this case, the function measures the differences between the outputs $a(z)$ and the available data y.

$$L = -\frac{1}{m}\sum_{i=1}^{m} y_i \log(a_i) + (1 - y_i)\log(1 - a_i) \tag{3}$$

The objective is to adjust the weights w and the bias b in order to minimize the errors of the model. In order to determine how the loss function varies with the different parameters, we calculate the corresponding gradient function (derivative). If the gradient is negative, then the function decreases as the weight w increases, so we must increase the weight w to reduce the corresponding errors. Otherwise, if the gradient is positive, the function increases when the weight increases. In this case, we must decrease the weight w to reduce the errors. Therefore, we use the following formula:

$$w_{t+1} = w_t - \alpha \frac{\partial L}{\partial w_t} \tag{4}$$

Where:

- w_{t+1}: Parameter w at the instant $t + 1$.
- w_t: Parameter w at the instant t.
- α: Positive learning step.
- $\frac{\partial L}{\partial w_t}$: Gradient at the instant t.

The optimizer is a function or algorithm that modifies the properties (such as the properties of the weights) of a neural network to reduce the overall loss and achieve higher accuracy. The optimization process is performed by the Adam optimizer using the stochastic gradient descent method. In fact, Adam is an adaptive gradient optimizer that implements Adam's algorithm by updating the learning rates based on the first and second order gradients. The program is easy to use and consumes relatively little memory, especially when a large amount of data and parameters are available.

4 Results and discussion

4.1 Feature dependency

Extraction of the most appropriate features from the dataset was performed. By applying the ranking method, relevant features were selected based on their importance. These features will be used to build our model. However, the more features selected, the longer the learning process and the less accurate the results can be. The definition of the variables used is presented in Table 3 and a visualization of the dependency of the features (relationships between our variables) is presented in Figure 2. We note that we defined 90% as a measure to eliminate one of the dependent variables.

Table 3. Definition of variables

ID	Factor
Resource	Usually contains pdf resources such as books
Oucontent	Represents content of assigments, which students should pass during presentation
Forumng	Disscussion forum
Url	Contains links to external or internal resources or for example video/audio content
Homepage	Course homepage
Subpage	Points to other sites in the course together with basic instructions

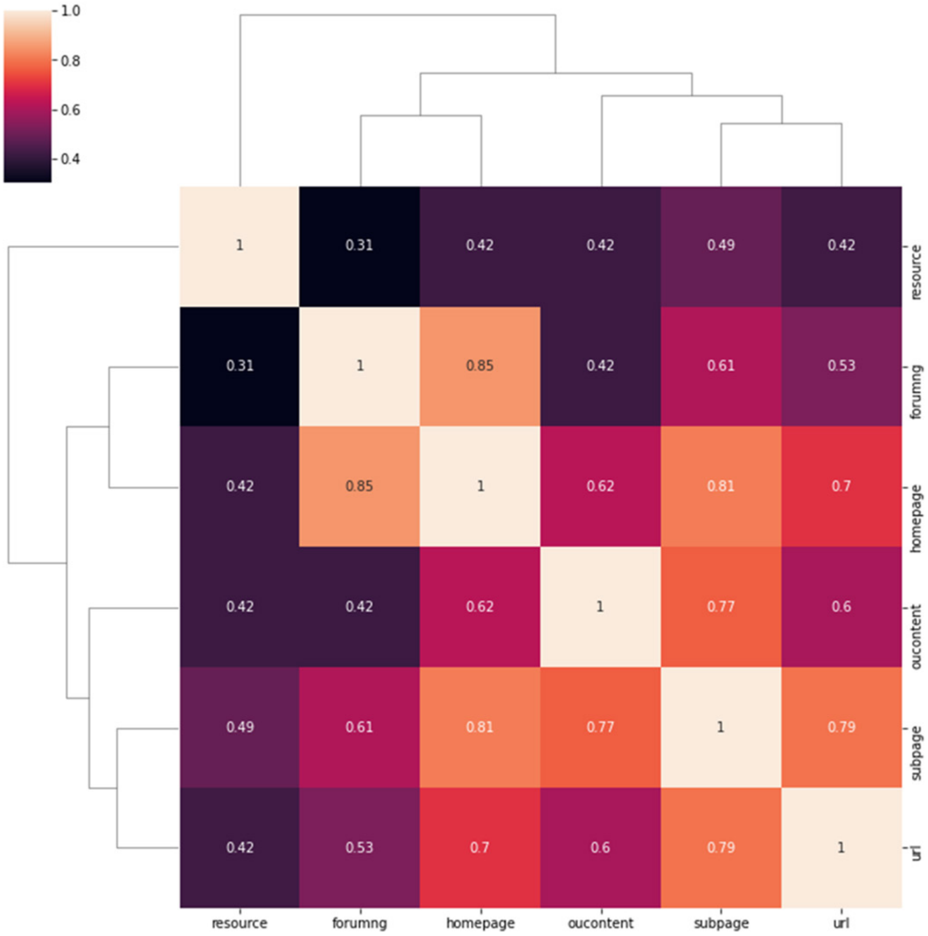


Fig. 2. Feature dependency in our dataset

4.2 Data preprocessing

Preprocessing involves preparing our data to ensure and improve the performance of our model. Typically, this step involves manipulating and reorganizing the data so that it is suitable for information extraction. Several operations were performed on our data.

First, as described in Figure 3, we got the data from two weeks: the first week was from the module AAA version 2013J, which has 383 rows and 18 columns, and the second week was from the module BBB version 2013J, which has 2237 rows and 18 columns. We would like to notice that the total number of weeks is as follows:

– The total number of weeks for module AAA2013J is 31 weeks.
– The total number of weeks for the module BBB2013J is 29 weeks.

Fig. 3. Target week selection process

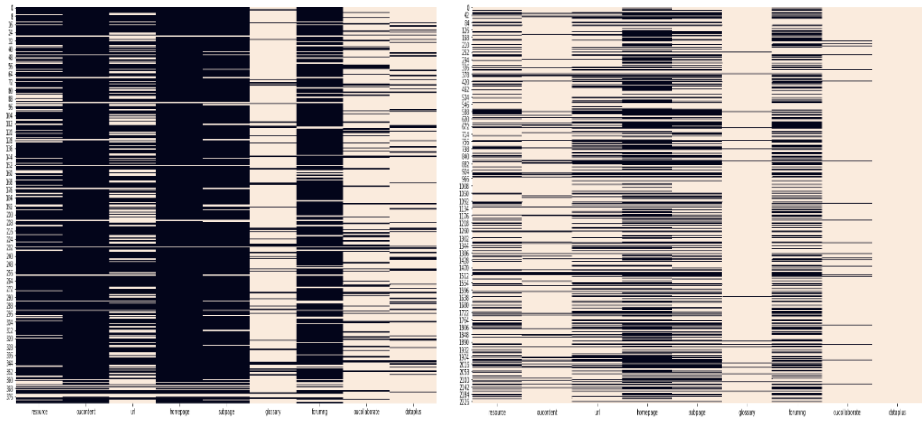
In fact, choosing the first few weeks of instruction in the module yields a higher rate of prediction accuracy because most students leave the course early, despite the high enrollment. We then performed a missing value analysis for both weeks.

One of the criteria we used to select the most useful variables for our model was the missing values measure. To do this, we eliminated all variables that had more than 80% missing values, before proceeding to study the dependence between the variables in our database. This led us to keep only the following variables: (homepage, Subpage, Url, Resource and Forumng).

Table 4 shows that the first week's data contains variables with more than 87% of missing values. In addition, the missing values for week two represent more than 96% for variables reporting learners' interactions with our e-learning platform (see the white boxes in Figure 4).

Table 4. Analysis of missing values

	Week 1	**Week 2**
homepage	0.031332	0.411265
oucontent	0.052219	0.527939
subpage	0.088773	0.566384
forumng	0.154047	0.620474
resource	0.203655	0.673670
url	0.386423	0.878409
oucollaborate	**0.809399**	**0.965132**
dataplus	**0.840731**	**0.965132**
glossary	**0.879896**	**1.000000**



(a) (b)

Fig. 4. Missing values for week 1 (a) and week 2 (b)

Once all unnecessary columns have been removed, view the target variable and determine the relationship between the variables. The last step before moving to the prediction phase is the data encoding phase to prepare the training and test data (see Table 5).

Table 5. Training and test data

	Training Data	Test Data
Week 1	306	77
Week 2	1,789	448

4.3 Predicting student dropout from MOOC

The dropout rate of MOOC has been identified as a key metric for assessing the effectiveness of distance learning. This section outlines our approach for predicting student dropouts by week.

The idea is to use a deep neural network for binary classification with five hidden layers, and an output layer for early prediction of student engagement for each week for a given course. Moreover, the model was evaluated using different activation functions (Sigmoid, ReLu) and optimizers (Adam) with a number of epochs = 100 and batch size = 10.

Table 6. Model summary and output shape for each layer

Layer (Type)	Output Shape	Param #
dense_230 (Dense)	(None, 9)	90
dense_231 (Dense)	(None, 8)	80
dense_232 (Dense)	(None, 7)	63
dense_233 (Dense)	(None, 5)	40
dense_234 (Dense)	(None, 2)	12
dense_235 (Dense)	(None, 1)	3

According to the table above (Table 6), the neural network we built performed better on the basis of the training data with respect to the accuracy metrics as follows:

- Week 1: An accuracy rate of 97.96% and a cost function of 2.90 % (the module AAA version 2013J).
- Week 2: An accuracy rate of 75.14% and a cost function of 24.86% (the module BBB version 2013J).

We notice that our model's accuracy rate differs between weeks 1 and 2 as a consequence of missing values concerning the learners' interactions in week 2.

Table 7. Training/test results

		Week 1	Week 2
Accuracy	Training	0.9796	0.7514
	Test	0.935	0.763
Cost	Training	0.0290	0.2486
	Test	0.0624	0.236

Based on the analysis of the test data (see Table 7), it can be seen that the proposed model gives promising results. Indeed, the accuracy metric shows 93% for week 1 and 76% for week 2, with a cost of 6% for week 1 and 23% for week 2.

5 Discussion

The present study aims to identify at-risk students using features of the deep learning approach. To this end, we implement a multilayer perceptron (ANN) algorithm to predict weekly dropout rates on the MOOC platform.

The proposed model processes the raw click stream data by considering the individual characteristics of learners, the VLE system as well as social aspects. The analysis of the results revealed that the proposed model outperforms the leading methods based on other deep learning algorithms (see Table 8). Indeed, with an accuracy rate of 97% in the first week and 75% in the second week, we can state that the proposed model provides reliable results for identifying learners likely to drop out of MOOC. Overall, the results illustrate the effectiveness of the proposed model. Indeed, the results reveal that interactions with the course homepage (Homepage, Subpage), links to external or internal resources (Url), pdf resources (Resource) as well as the discussion forum (Forumng) are the most important features to assess learners' performance and identify students at risk. These results have led to recommendations that we intend to pursue, as perspectives of this work, to improve E-learning in the context of higher education, especially in the Moroccan educational system.

Here are some concluding remarks:

- Model overfitting: As shown in the graphs (see Figure 5), the additional number of epochs does not appear to be of much benefit to the model up to approximately 20 epochs. Furthermore, we do not recommend additional training in order to prevent overfitting. Therefore, we are able to generalize the model since it has sufficient validation accuracy to solve our particular problem.
- Tensorflow Framework: To carry out this study, we use the most popular deep learning framework Tensorflow. A particular characteristic of TensorFlow is the ability to model computations as an execution graph where nodes represent operations to be executed, and links represent Tensors. Furthermore, TensorFlow automatically manages the creation of the graph once the tensors and operations have been implemented (instantiated) which provides a mechanism for optimizing and parallelizing the execution of the code during the startup process. As well as providing extensive support for DL-specific operations, TensorFlow also facilitates the creation of neural network training algorithms by using commonly associated mathematical operations.
- Dataset Configuration: The database used provides the option of working on each week of a given module or all modules at the same time. In our case, we chose the first week from the AAA version 2013J module, and the second week from the BBB version 2013J module. As a result of the deep learning model developed, at the end of each week, the model will be able to predict students at risk of dropping out. Furthermore, we divided the dataset into 3 parts: Training set (data used to train the model), Test set (reserved for the evaluation of our model), and validation set to find the parameters that offer the most optimal performance. The partition dataset is 80% for training, 20% for testing and 20% for validation extracted from the training set.
- Accuracy and Precision: In this study, accuracy was used as a measure of precision, Adam as an optimization algorithm, and MAE as a measure of evaluation. The main advantage of the "adam" optimizer is that the learning rate does not need to be specified, as is the case with gradient descent. Finally, the results of the survey are convincing, as a high level of participation was observed during the first week. As for the other periods, a systematic intervention will be carried out regularly to minimize the risk of dropouts in the upcoming weeks.
- Post-COVID-19 challenges: We assume that the determinants of success of our prediction model applied on post-COVID-19 data should take into account individual characteristics of learners, teachers, the VLE system as well as social aspects. In this case, we plan to improve our model in a progressive manner by taking into account recent data collected regularly from both online surveys and MOOC log files. The idea is to closely monitor the post COVID-19 impact of other parameters that may challenge the dropout factors initially identified in this study.

Table 8. Comparison of the main AI methods (DL) for dropout prediction

	Feng et al [13]	Shabandar et al. [16]	Kanetaki, Z et al. [12]	Our Approach
Methods	DL CFIN (Context-aware Feature Interaction)	Deep Decision tree forest, NN, Generalized linear model	SPSSv20, PCA, ANOVA, (ANN), MATLAB	(ANN) – Tensorflow
Data Types	Demo-graphics, traces	Demo-graphics, traces	Interactions with the VLE, surveys	Learners' interactions in a MOOC
MOOC platforms / data source	XuetangX-KDDCUP	OULAD and a course from Harvard & MIT	1st semester: course, "Computer Aided Mechanical Design CAD".	OULAD and a course from Harvard & MIT
Results	XuetangX: AUC (86.71%) – F1 (90.95%) KDDCUP: AUC (90.93%) – F1 (92.87%)	The model shows that the motivational trajectory is an important factor in predicting dropout.) F–Measure (95.2%)	Several factors were found to have a significant relationship with students' performance during COVID-19 restrictions	Accuracy rate: Week 1 (97.96%) Week 2 (75.14%)

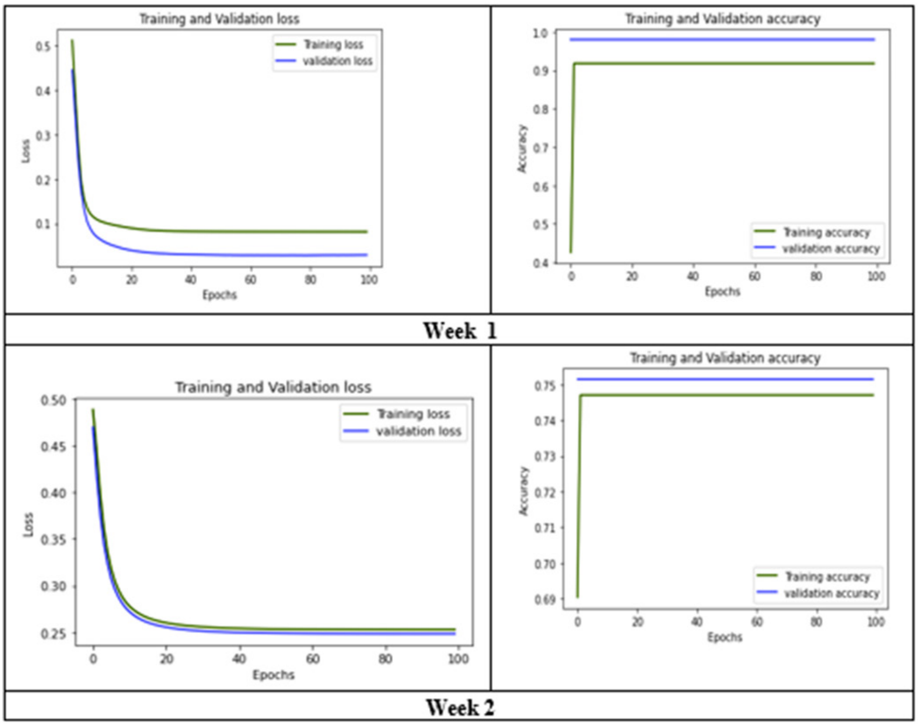


Fig. 5. Training and validation results (loss and accuracy)

6 Conclusions

In summary, a multi-layer perceptron algorithm is implemented to predict the dropout rate on the MOOC platform each week. The analysis results show that the first week had a prediction rate of 97%, while the accuracy rate of the second week corresponded to 75%. Therefore, the proposed model provides reliable results for identifying learners who are at risk of dropping out of MOOC. These results can help MOOC administrators make timely decisions and implement strategies to reduce dropout rates.

As perspectives, we will focus on customizing courses based on identified needs and intend to test the performance and reliability of our approach on a real-world use case. In addition, we intend to further investigate the psychological impact of MOOC dropout on university students, particularly in the Moroccan higher education system.

7 References

[1] R. Bansal, A. Gupta, R. Singh, and V. K. Nassa, "Role and Impact of Digital Technologies in E-learning Amidst COVID-19 Pandemic," *Proceedings – 2021 4th International Conference on Computational Intelligence and Communication Technologies, CCICT 2021*, pp. 194–202, Jul. 2021. https://doi.org/10.1109/CCICT53244.2021.00046

[2] Z. Kanetaki, C. Stergiou, G. Bekas, S. Jacques, C. Troussas, C. Sgouropoulou, and A. Ouahabi, "Grade Prediction Modeling in Hybrid Learning Environments for Sustainable Engineering Education," *Sustainability*, vol. 14, no. 9, p. 5205, 2022. https://doi.org/10.3390/su14095205

[3] Z. Kanetaki, C. Stergiou, G. Bekas, C. Troussas, and C. Sgouropoulou, "Evaluating Remote Task Assignment of an Online Engineering Module through Data Mining in a Virtual Communication Platform Environment," *Electronics*, vol. 11, no. 1, p. 158, 2022. https://doi.org/10.3390/electronics11010158

[4] Z. Kanetaki, C. Stergiou, G. Bekas, C. Troussas, and C. Sgouropoulou, "Analysis of Engineering Student Data in Online Higher Education During the COVID-19 Pandemic," *International Journal of Engineering Pedagogy (iJEP)*, vol. 11, no. 6, pp. 27–49, 2021. https://doi.org/10.3991/ijep.v11i6.23259

[5] D. F. O. Onah and J. E. Sinclair, "Assessing Self-Regulation of Learning Dimensions in a Stand-alone MOOC Platform," *International Journal of Engineering Pedagogy (iJEP)*, vol. 7, no. 2, pp. 4–21, May 2017. https://doi.org/10.3991/ijep.v7i2.6511

[6] D. Gamage, I. Perera, and S. Fernando, "MOOCs Lack Interactivity and Collaborativeness: Evaluating MOOC Platforms," *International Journal of Engineering Pedagogy (iJEP)*, vol. 10, no. 2, pp. 94–111, Mar. 2020. https://doi.org/10.3991/ijep.v10i2.11886

[7] S. Jacques and T. Lequeu, "The Attractiveness of Reversing Teaching Forms – Feedback on an Electrical Engineering Course," *International Journal of Engineering Pedagogy (iJEP)*, vol. 10, no. 3, pp. 21–34, 2020. https://doi.org/10.3991/ijep.v10i3.12361

[8] S. Jacques, A. Ouahabi, and T. Lequeu, "Remote Knowledge Acquisition and Assessment During the COVID-19 Pandemic," *International Journal of Engineering Pedagogy (iJEP)*, vol. 10, no. 6, pp. 120–138, 2020. https://doi.org/10.3991/ijep.v10i6.16205

[9] Z. Kanetaki, C. Stergiou, G. Bekas, S. Jacques, C. Troussas, C. Sgouropoulou, and A. Ouahabi, "Acquiring, Analyzing and Interpreting Knowledge Data for Sustainable Engineering Education: An Experimental Study Using YouTube," *Electronics*, vol. 11, no. 14, p. 2210, 2022. https://doi.org/10.3390/electronics11142210

[10] A. Chatzopoulos, M. Kalogiannakis, S. Papadakis, and M. Papoutsidakis, "A Novel, Modular Robot for Educational Robotics Developed Using Action Research Evaluated on Technology Acceptance Model," *Education Sciences*, vol. 12, no. 4, p. 274, 2022. https://doi.org/10.3390/educsci12040274
[11] J. Gardner and C. Brooks, "Student Success Prediction in MOOCs," *User Modeling and User-Adapted Interaction*, vol. 28, no. 2, pp. 127–203, Jun. 2018. https://doi.org/10.1007/s11257-018-9203-z
[12] Z. Kanetaki, C. Stergiou, G. Bekas, C. Troussas, and C. Sgouropoulou, "A Hybrid Machine Learning Model for Grade Prediction in Online Engineering Education," *International Journal of Engineering Pedagogy (iJEP)*, vol. 12, no. 3, pp. 4–24, 2022. https://doi.org/10.3991/ijep.v12i3.23873
[13] W. Feng, J. Tang, and T. X. Liu, "Understanding Dropouts in MOOCs," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 517–524, Jul. 2019. https://doi.org/10.1609/aaai.v33i01.3301517
[14] B.-H. K. Udacity and A. I. Team, "Deep Learning to Predict Student Outcomes," *arXiv preprint arXiv:1905.02530*, Apr. 2019.
[15] B. H. Kim, E. Vizitei, and V. Ganapathi, "GritNet: Student Performance Prediction with Deep Learning," *Proceedings of the 11th International Conference on Educational Data Mining, EDM 2018*, Apr. 2018.
[16] R. Al-Shabandar, A. J. Hussain, P. Liatsis, and R. Keight, "Detecting At-risk Students with Early Interventions using Machine Learning Techniques," *IEEE Access*, vol. 7, pp. 149464–149478, 2019. https://doi.org/10.1109/ACCESS.2019.2943351
[17] E. M. Smaili, S. Sraidi, S. Azzouzi, and M. E. H. Charaf, "Towards Sustainable e-Learning Systems Using an Adaptive Learning Approach," *Advances in Science, Technology and Innovation*, pp. 365–372, 2021. https://doi.org/10.1007/978-3-030-53440-0_38
[18] E. M. Smaili, S. Sraidi, S. Azzouzi, and M. E. H. Charaf, "An Adaptive Learning Approach for Better Retention of Learners in MOOCs," in *The 3rd International Conference on Networking, Information Systems & Security*, pp. 1–5, 2020.
[19] S. Soukaina, E. M. Smaili, M. El, and M. E. H. Charaf, "MOOCs Performance Analysis based on Quality and Machine Learning Approaches," in *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, pp. 1–5, 2020. https://doi.org/10.1109/ICECOCS50124.2020.9314606
[20] A. A. Mubarak, H. Cao, and I. M. Hezam, "Deep Analytic Model for Student Dropout Prediction in Massive Open Online Courses," *Computers & Electrical Engineering*, vol. 93, Jul. 2021. https://doi.org/10.1016/j.compeleceng.2021.107271
[21] M. Hussain, W. Zhu, W. Zhang, S. M. R. Abidi, and S. Ali, "Using Machine Learning to Predict Student Difficulties from Learning Session Data," *Artificial Intelligence Review*, vol. 52, no. 1, pp. 381–407, Jun. 2019. https://doi.org/10.1007/s10462-018-9620-8

8 Authors

Soukaina Sraidi is a PhD student in the field of machine learning/deep learning by applying sentimental analysis based on quality tools and supervised and unsupervised learning algorithms to track learners' emotional tendencies in order to reduce the dropout rate in MOOC. Responsible for the management of the public market IT system – University Ibn Tofail, Kenitra, Morocco (Email: sraidi.soukaina@uit.ac.ma).

El Miloud Smaili is a PhD student in the field of machine learning/deep learning through the prediction of learners at risk of dropping out in MOOC and the proposal of adaptive learning optimization solutions to improve the quality of MOOC. Holder of a state engineering degree in computer engineering at ENSAO, senior state engineer at the Faculty of Human and Social Sciences – University Ibn Tofail, Kenitra, Morocco (Email: miloud.smaili@uit.ac.ma).

Salma Azzouzi is a Professor at Ibn Tofail University – Faculty of Science (Computer Science Department). She is a member of the LaRI laboratory. She is also the general co-chair of the International Conference on Electronics, Control, Optimization and Computing (ICECOCS) and guest editor of a special issue of the journal SOIC. Her main areas of interest are: Distributed testing, e-learning, optimization. (Email: salma.azzouzi@uit.ac.ma).

Moulay El Hassan Charaf is a Professor at Ibn Tofail University – Faculty of Sciences (Department of Computer Science). He is a member and deputy director of the LaRI laboratory. He is also the general co-chair of the International Conference on Electronics, Control, Optimization and Computing (ICECOCS) and guest editor of the special issue on Intelligent Control for Future and Complex Systems in the International Journal of Modeling, Identification and Control (IJMIC). His main areas of interest are: Control, optimization, adaptive learning, distributed testing. (Email: my.charaf@uit.ac.ma).

Article submitted 2022-07-09. Resubmitted 2022-09-21. Final acceptance 2022-09-21. Final version published as submitted by the authors.