

Review

Predicting Student Outcomes in Online Courses Using Machine Learning Techniques: A Review

Areej Alhothali ^{1,*}, Maram Albsisi ^{1,†}, Hussein Assalahi ² and Tahani Aldosemani ³

¹ Department of Computer Science, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 22254, Saudi Arabia; malbsisi0002@stu.kau.edu.sa

² English Language Institute, King Abdulaziz University, Jeddah 22254, Saudi Arabia; hassalahi@kau.edu.sa

³ College of Education, Prince Sattam bin Abdulaziz University, Al-Kharj 16278, Saudi Arabia; t.aldosemani@psau.edu.sa

* Correspondence: aalhothali@kau.edu.sa

† These authors contributed equally to this work.

Abstract: Recent years have witnessed an increased interest in online education, both massive open online courses (MOOCs) and small private online courses (SPOCs). This significant interest in online education has raised many challenges related to student engagement, performance, and retention assessments. With the increased demands and challenges in online education, several researchers have investigated ways to predict student outcomes, such as performance and dropout in online courses. This paper presents a comprehensive review of state-of-the-art studies that examine online learners' data to predict their outcomes using machine and deep learning techniques. The contribution of this study is to identify and categorize the features of online courses used for learners' outcome prediction, determine the prediction outputs, determine the strategies and feature extraction methodologies used to predict the outcomes, describe the metrics used for evaluation, provide a taxonomy to analyze related studies, and provide a summary of the challenges and limitations in the field.

Keywords: MOOCs; SPOCs; student performance; student dropout; machine learning; learning behaviour; learning analytics



Citation: Alhothali, A.; Albsisi, M.; Assalahi, H.; Aldosemani, T. Predicting Student Outcomes in Online Courses Using Machine Learning Techniques: A Review. *Sustainability* **2022**, *14*, 6199. <https://doi.org/10.3390/su14106199>

Academic Editors: Neil Gordon and Han Reichgelt

Received: 27 March 2022

Accepted: 13 May 2022

Published: 19 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

Online education has revolutionized the way people learn and has made education more accessible and affordable to numerous people worldwide. Despite the advantages and increased interest in online and distance learning, educational institutions are becoming increasingly concerned about students' performance and retention rates, particularly low certification/graduation and dropout/completion rates. Failing or dropping out of an online course or program is often considered a key parameter by institutional authorities for assessing program/course quality and allocating resources. Dropout and low certification rates can also pose a potential risk to an institution's reputation, profit, and funding [1]. These outcomes also have significant consequences for a student's self-esteem, well-being, employment, and chances of graduating [1,2]. As a result, finding more efficient approaches to forecasting students' performance as early as possible is critical for institutions, students, and educators to take proactive steps toward improving students' online learning experiences and establishing intervention strategies that target students' needs. With the increased interest in online education and the large amount of data produced by learners through their interactions with online platforms, researchers have proposed methods to analyze learners' behavioral data to predict and improve educational outcomes.

Learning analytics (LA), more commonly known as educational data mining (EDM) [3], the task of analyzing and finding patterns in learners' data for decision-making purposes, has attracted many researchers in recent years. Learning analytic tools enable institutions

to gain an understanding of their students' status, actions and preferences individually, and in relation to their peers and the targeted educational objective. This allows the tailoring of material for individual students based on the projected outcomes and preferred learning styles. In online education, LA systems assess students' learning behavior by utilizing extensive data collection of learners' data, including student enrollment information, past and current academic records, students' online behavior, student surveys via questionnaires concerning courses and teaching techniques, and data from online discussion forums. Scholars have also examined various learning-behavior attributes to predict learning outcomes, such as learners' performance and retention. To predict and analyze students' outcomes in online courses, researchers have examined several machine learning models, including support vector machines (SVMs), linear regression (LR), random forest (RF), and deep learning models such as convolutional neural networks (CNNs) and long short-term memory (LSTM).

1.1. Previous Reviews of Student Outcome Prediction

Several studies have provided a comprehensive analysis of the literature in the field of learning analytics, investigating studies that analyzed learners' behavior to optimize educational outcomes. Three studies presented a review analysis similar to the survey conducted in this research. Moreno-Marcos et al. [4] presented a meta-analysis of state-of-the-art predictive models based on MOOC data. The analysis provides an overview of the features, methods, and metrics used in the literature. The authors of [5,6] presented a survey of learning analytic studies covering related studies and methodologies used in the literature. Several studies have also presented literature reviews on online learners' dropping out from massive open online courses (MOOCs) [7]. In addition, another review paper [8,9] focused on performance prediction in MOOCs and small private online courses (SPOCs).

1.2. Method

This study presents a review of studies that aim to predict student outcomes, which we define in terms of achievement, completion, and continuation in online educational courses. The study also provides an overview and taxonomy of the current related work with a detailed analysis of the features and methods used in the literature. It also covers different types of online learning environments, including MOOCs and SPOCs. In particular, this study aimed to answer the following research questions:

1. What is the process followed by researchers for learner outcome prediction?
2. What are the predictive variables used to predict learner outcome?
3. What are the learner outcomes used in the literature?
4. What are the online learning platforms used in the literature?
5. What are the machine learning methodologies used in the literature?
6. What are the challenges and limitations, and future directions of this field?

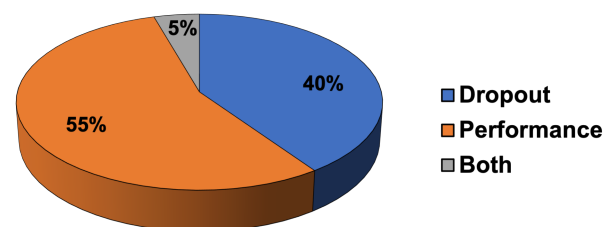
The literature contains numerous studies that examine online learner data to better understand learners' progress and outcomes in online courses. We included all studies published between 2017 and 2021 that utilized users' learning-behaviour data and employed machine learning and deep learning techniques to predict learner outcomes in online courses, including student dropout and student performance (students at risk, student grade, and student certificate acquisition). We searched a number of electronic databases, such as Scopus and Web of Science, and publishers, such as Springer, IEEE, Elsevier, and Sage, using the keywords and queries shown in Table 1. The search was expanded using snowball search methods to identify additional related studies.

Table 1. Search strings used in the web search engine.

Concept	Search Query
Learner Performance	("Grade" OR "Performance" "Success" OR "Failure" OR "Certificate" OR "At-risk")
Learner Dropout	("Dropout" OR "Retention" OR "Completion" OR "Attrition" OR "Withdrawal")
Online Learning	("Online learning" OR "MOOC" OR "Online course" OR "Online Education")
Machine Learning	("Classification" OR "Prediction" OR "Machine Learning" OR " Predictive model" OR "Deep learning")

1.3. Study Selection

The search was conducted between 1 May 2021 and 30 June 2021, and returned a total of 137 research studies. To exclude studies that do not match the inclusion criteria, we examined the research's paper title, abstract, and sometimes the proposed methodologies if needed. We excluded studies that did not employ any machine learning models in their proposed methodology, those that utilized learning features obtained from blended learning or face-to-face learning, and those that used features obtained from mobile-based or e-book platforms. Studies that aimed to detect and analyze other aspects of students' online experiences, such as sentiment, engagement, satisfaction, and learning style, were also excluded from this review. Studies that only utilized features obtained from learning systems, such as previous academic achievements and demographic data, were excluded from the study. We also excluded studies that only looked into linguistic features of students' posts or comments on discussion boards. Studies that predict students' performance in assignments, exams, or quizzes, rather than final grades or final exams, were excluded from this study. After removing irrelevant research, we were left with 67 studies. Figure 1 shows the number of dropout and performance-prediction studies after applying the exclusion criteria.

**Figure 1.** The number of dropout and performance-prediction studies after applying exclusion criteria.

1.4. Student Outcomes Prediction Model Process

The problem of forecasting student outcomes using learner-interaction data is often formulated as a supervised problem that requires a dataset of pairs of values (x_{ij}, y_j) , in which x_{ij} denotes the i th feature or attribute that characterizes the j th student and y_j is the learner outcome (e.g., 1 = dropout or 0 = no dropout). The goal of the predictive model is to learn a function $h = y(x)$ that estimates the relationship between the input or independent variables and the predicted outcome, which is the dependent variable in this problem. The dependent variables can be continuous real values (e.g., grade = 93.0), dichotomous variables that take two possible values (e.g., fail = 1, success = 0), or polychotomous variables that have more than two possible values (e.g., high, intermediate, and low).

Similar to other problems in machine learning, predictive models in this domain commonly follow a five-step procedure, as shown in Figure 2. The first step involves collecting datasets from online platforms. The samples in the dataset represent learners' information and activities in one or more courses over a period of time. A preprocessing stage is then performed to extract and select valuable features. After extracting the features, several machine learning models are trained on the training set and validated on a validation set. Then, hyper-parameter (e.g., number of training epochs, regularization penalty) optimiza-

tion techniques are implemented to choose the optimal hyper-parameters of the model. Then, the trained model is evaluated on an unseen testing set, and the prediction accuracy is estimated using metrics in the last step. The following sections provide information related to each of these steps.



Figure 2. The procedure of student outcome prediction.

The remainder of the paper is organized as follows: Section 2 provides a summary of the online learning environments used in the literature; Section 3 provides an overview of the courses or subjects used in the literature; Section 4 presents a summary of the predictive variables used in the literature, with detailed analysis of feature-extraction and feature-selection techniques in Sections 5 and 6, respectively. Section 7 provides an overview of the models used in previous studies. Section 9 provides details about learner outcomes, a summary of the studies that predict student performance in Section 9.1 and student dropout in Section 9.2. Finally, Sections 10 and 11, give a summary of the related studies and challenges in the field. Online courses that are delivered to a large scale of learners are called massive open online courses (MOOC), while those that target private or specific groups of students are called small private online course (SPOC) [10]. Despite the similarity between these two types of courses, they have distinctive characteristics. The number of enrolled students in SPOCs, for example, is much smaller (15 to 20 students per course) than that of students in MOOCs (up to 10,000 per course). The small number of SPOC enrolments has contributed to enhancing teachers' guidance and increasing retention rates in comparison to MOOCs [10].

2. Online Learning Environment

Several MOOC and SPOC learning systems, such as Coursera, edX, Moodle, and other private online platforms, have been used to predict student outcomes. However, because of their large number of students and courses, the majority of studies leveraged data from MOOCs. MOOCs also have a higher dropout rate; thus, a large number of studies have investigated methods to forecast dropout in MOOCs [10]. Figure 3 shows the number of studies that uses MOOCs and SPOC in dropout and performance prediction. In the literature, we found that datasets are often collected from various courses that range from one course to 56 courses, with several samples/records that range from 104 to 597,692 records. The duration of data collection also varies among studies. The researchers considered two strategies: collecting data over a fixed period, such as four weeks or 30 days, or over different durations, such as two, three, and four weeks.

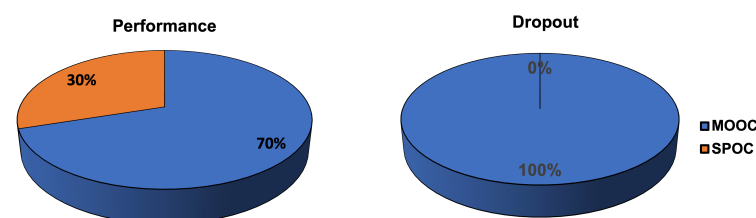


Figure 3. The distribution of the studies that use MOOCs and SPOCs.

Several public datasets have been used in related studies to predict student performance and dropout rates. These datasets can be considered benchmark datasets, allowing researchers to evaluate the performance of the model compared with the others. The public datasets in this domain are the Students' Academic Performance Dataset (SAPData) [11], which has been utilized for grade prediction [12]; Open University Learning Analytics Dataset (OULAD) [13] developed by the Open University (OU) and used for at-risk [14], pass/fail [15], grade [16], dropout [15], and engagement prediction [17]; Center for Ad-

vanced Research Through Online Learning (CAROL) [18] has been used for dropout [19] and fail/success prediction [20,21]; KDD Cup 2015 (KDDcup) [22] has been extensively used in the literature as a whole or subset to predict dropout, such as [23,24]; and HarvardX and MITx dataset (HMedx) [25] have been used for dropout prediction [26] and performance prediction [26,27]. Figure 4 shows the statistics of each dataset used in previous studies. Most of the studies use the KDDcup 2015 dataset, where 19 studies have utilized this dataset, followed by OULAD with 6 studies, CAROL with 3 studies, and HMedx with 2 studies. A summary of the public-dataset usage is provided in Table 2.

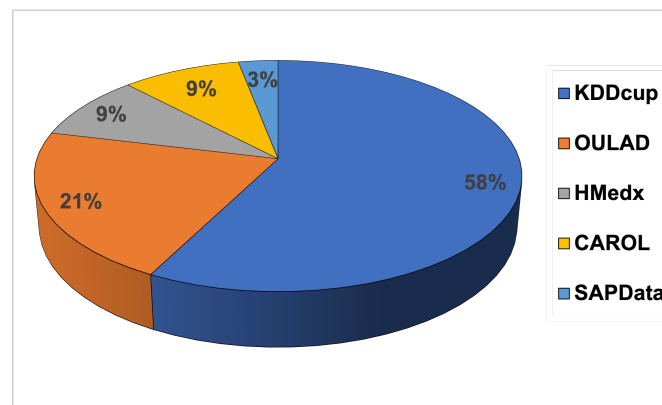


Figure 4. The distribution of the studies that use publicly available dataset.

Table 2. Public learning-analytic datasets used in the literature.

Ref.	Dataset	Platform	Courses	Records	Features	Outcomes
[11]	SAPData	Kalboard 360	12	480	Demographic, Academic Background, Interaction	Performance
[13]	OULAD	OU VLE	22	32,593	Demographic, Registration, Assessment, Interaction	Performance
[18]	CAROL	OpenEdX	5	78,623	Interaction, Assessment	Dropout, performance
[22]	KDDcup	XuetangX	39	120,542	Enrollment, Course, Interaction	Dropout
[25]	HMedx	edX	17	597,692	Academic Background, Video Interaction, Assessment	Dropout, performance

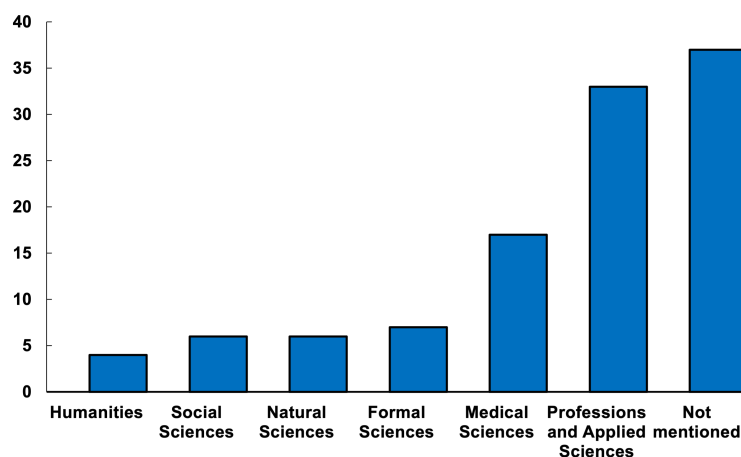
3. Courses

When analyzing students' learning outcomes, researchers use two approaches with respect to the course subject, focusing on subject-independent or subject-dependent features. Most studies in the literature use subject-independent features by examining/analyzing students learning behavior independent of the course subject, such as the number of download sources. However, a few studies have focused on analyzing features related to the subject of the study, such as utilizing variables related to learning mathematics [28].

In the literature, student-outcome predictive systems are often developed based on a single course or multiple subjects. Researchers have identified five subject categories in the literature: humanities, social sciences, natural sciences, formal sciences, professions, and applied sciences [29]. Table 3 provides an overview of some subjects and their categories. As shown in Figure 5, the most widely used subject categories are formal science, and more specifically, computer-related subjects such as programming and computer networking.

Table 3. Subject categories used in the literature.

Subject Category	Example Subject	No. of Studies
Humanities	Online foreign language teaching, understanding language	6
Social sciences	General sociology, Social science	7
Natural sciences	Analytical chemistry laboratory, Physics III	6
Formal sciences	Assembly Language, C programming, Calculus I	33
Medical sciences	First aid general knowledge, Public health research	4
Professions and applied sciences	Circuits and Electronics	17
Not mentioned	-	37

**Figure 5.** Distribution of subject categories used in the related studies.

4. Predictive Variables

Existing methods for forecasting online learners' outcomes, such as dropout, grade, and completion, are based on data-mining techniques, which entail collecting attributes from learners' data provided during their online study and then making predictions using various data-mining tools. Several features have been investigated to predict and estimate students' learning outcomes. These features can be categorized into pre-course information, such as demographic background, previous academic background, course information; in-course information, such as interaction data and learning behavior; and post-course data, such as graded assessment and final grade.

Demographic information, registration, or enrollment data mainly describe the characteristics of the learners, including their name, gender, age, mother tongue, geography/origin, occupation, socioeconomic status, and hobbies. Previous academic background describes a learner's past academic achievements, including information about students' cumulative grade point averages. Learning behavior, also referred to as engagement data or log data, is the behavior students exhibit during their interaction with an online course. Learning behavior features are often represented in terms of frequency or duration, focusing on particular targets such as content, assessment, and forums. Learning behavior also includes multimedia learning behavior (i.e., video-related features), download courseware behavior, text learning behavior (i.e., browsing online content), and exercise-related behavior. The granularity of the extracted pattern can constitute another level, by focusing on more fine-grained interaction data with specific events such as multimedia or video, practice, and participation events. Video or multimedia interaction data, for example, can be pausing, stopping, and replaying. Studies in the literature looked at different types of features. The most widely used features are interaction learning behavior features. Only a small number of studies looked into fine-grained multimedia, practice, and participation features. A summary of the learner data used in the related studies is provided in Table 4. Throughout the analysis of the studies, behavioral (log data) followed by demographic data were the most widely used in both dropout prediction and performance prediction, as shown in Figure 6.

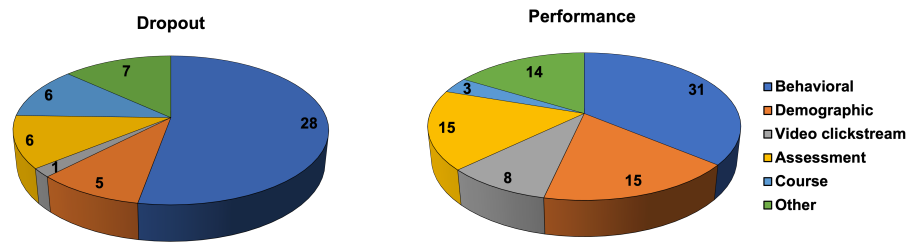


Figure 6. Statistics of feature categories used in dropout prediction and performance prediction.

Table 4. Summary of learner data used in the literature and their categories.

Category	Features
Demographic features	Date of birth, birthplace, age, gender, parent responsible for student, nationality, mother tongue
Academic background	Studying semester, GPA, grade level, section, education
Enrollment data	Information about student’s enrollment on a course
Course data	No. of enrolled students, drop rate, course modules
Attendance data	Student absent from class
Learning interaction (log data)	Activity, visit resources, downloaded resources, play resources, access a piece of content, class participation, logins, starting a lesson, page navigation, page closes
Multimedia/video interactive data	Pause, replay, stop, open, close
Practice data	Questions answered, tests, tries, assessment scores
Participation data	Discussions on forums, polls, messages posted, messages read

5. Features Engineering

The initial form of learning-behavior data is log data or a clickstream, which contains two major variables: the timestamp and events (e.g., opening a page). In the literature, four forms of features have been used to represent learning behavior patterns: statistical, statistical-temporal, raw, and raw-temporal features. Most studies on this subject have relied on a coarse-grained statistical process to represent learning events in terms of frequency, rate, or length computed and accumulated over a specific time frame, such as a day, week, or several weeks. As a result, the learner is represented in the predictive model as a fixed-length flat vector of some continuous values as a whole or as a sequence of features (e.g., the total number of activities in a day or the rate of practicing behavior over time).

Other researchers have considered the statistical-temporal approach by analyzing learner activity over any period (weeks or days). The data of each learner are represented by a dynamic-length flat vector that varies in length according to the chosen duration. Recent research has focused on employing more fine-grained feature analysis to extract temporal properties from raw data using deep-learning algorithms. To automatically extract temporal features from raw textual or categorical data, one first needs to encode these raw data into representations, such as the one-hot encoding scheme. Then, deep-learning models with convolutional layers have been utilized to automatically extract the most significant features in the prediction task. As shown in Figure 7 in both dropout prediction and performance prediction tasks, statistical features have been the most commonly used feature-extraction methodology. In contrast, a small number of studies have looked into raw temporal features.

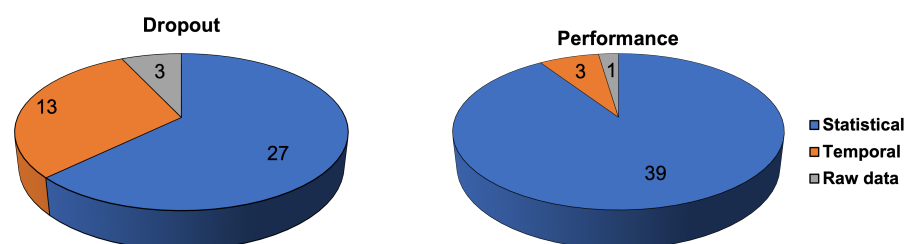


Figure 7. Statistics of feature-extraction methodologies used in the related studies.

6. Feature Selection

To lower the dimensionality of the feature space, two paradigms are often employed in machine learning: feature selection and dimensionality reduction. The feature-selection approach is a preprocessing procedure that is performed before the learning stage to choose the most relevant features for the predictive task and increase the model's accuracy and efficiency. In addition to the strategy of manually selecting pertinent features adopted by most researchers in the field, several automatic feature-selection methodologies have also been used in the field. These methodologies can be categorized into correlation-based, wrapper-based, and ensemble-based feature selection.

Several studies have applied correlation-based feature selection approaches such as chi-square (X^2), mutual information (MI), information gain, fast correlation-based filter (FCBF), relief, and Pearson's correlation coefficient [21,30,31] to select the topmost related features to the targeted values and discard irrelevant and noisy attributes. This approach has the benefit of being adaptable to any machine-learning model.

Wrapper methods, in contrast, search the attribute space and use a classifier to determine the best set of features. Several wrapper-based methods have been used in the literature, including recursive feature elimination (RFE) [32], random hill climbing [27], and WrapperSubsetEval [33]. Ensemble-based feature selection methods have been examined in several studies [21]. Dimensionality reduction approaches such as principal component analysis (PCA) map the feature space to a lower-dimensional space. Several studies have applied this method to reduce the dimensionality of the feature space [20,34,35]. The main drawback of dimensionality reduction approaches is that the newly generated features do not reveal much information regarding the original feature space.

7. Models

Researchers have used a wide range of machine- and deep-learning techniques to predict online student outcomes from their online interaction data. The machine-learning models utilized in the field are categorized into probabilistic models such as Naive Bayes (NB), linear models such as logistic regression (LR), ensemble methods such as AdaBoost, tree-based methods such as decision trees (DT), rule-based methods such as fuzzy-logic-based approaches, and instance-based learning such as K-nearest neighbor (kNN). In addition, some studies have developed optimization methods or heuristic search-based algorithms, such as the Kstar algorithm, to predict student outcomes. Deep-learning models, such as convolutional neural networks (CNNs) and long short-term memory (LSTM), have been investigated in recent studies to predict learners' outcomes in online learning. Table 5 summarizes the machine- and deep-learning methods used in previous studies. Figure 8 shows the predictive-model categories used in relevant studies. As shown in the figure, deep-learning models are the most widely used in the literature. It is worth noting that the term deep learning refers to models that implement neural network models with more than three layers; thus, studies that develop a feed-forward neural network with more than three hidden layers are categorized as deep learning.

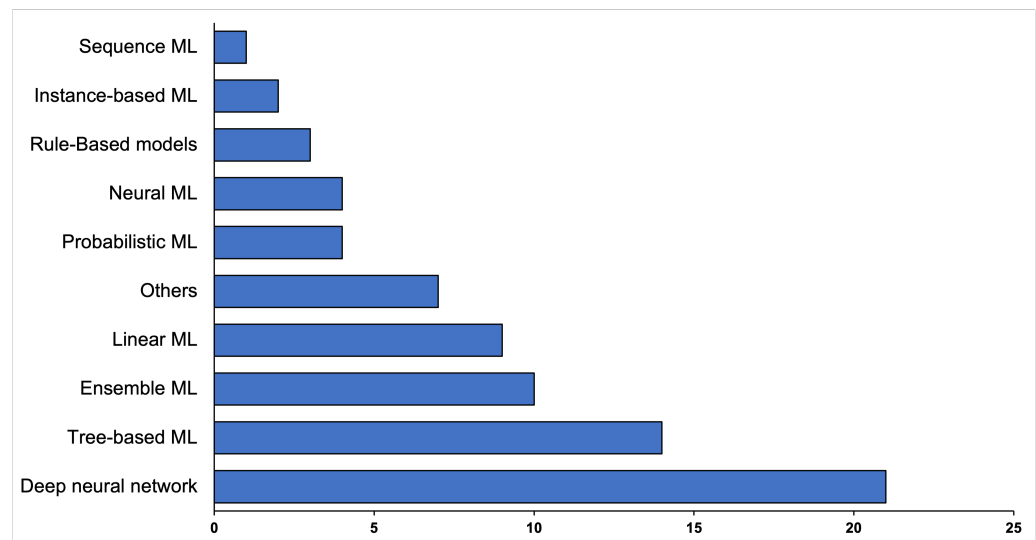


Figure 8. Distribution of machine-learning models in the literature.

Table 5. Summary of the machine-learning models used in the literature and their categories.

Model Category	Models
Probabilistic model	Naive Bayes, Bayes network, Bayesian generalized linear (BGL), Bayesian belief networks,
Linear models	Logistic regression (LR), support vector machine (SVM), linear discriminant analysis (LDA), general- ized linear model (GLM), lasso linear regression (LLG), boosted logistic regression
Ensemble methods	Bagging, boosting, stacking, AdaBoost, gradient boosting (GB), eXtreme gradient boosting (xgbLin- ear), stochastic gradient boosting (SGB)
Tree-based models	Decision tree (DT), random forest (RF), Bayesian additive regression trees (BART)
Rule-Based models	Rule-based classifier(JRip), fuzzy set rules
Instance-based learning	k-nearest neighbors (kNN)
Neural network	Multilayer perceptron(MLP) or artificial neural network (ANN)
Sequence ML models	Conditional random fields (CRF)
Deep neural network	Recurrent neural network (RNN), gated recurrent unit (GRU), long short-term memory (LSTM), convolutional neural network (CNN), squeeze-and-excitation networks (SE-net)
Others	Search algorithms (Kstar), optimization algorithm (pigeon-inspired optimization (PIO)), matrix completion, unsupervised learning model (self organized map (SOM))

8. Evaluation Metrics

Predicting student outcomes is regularly formulated as a supervised learning problem, either as a classification or regression problem. The predictive performance of the proposed classification model has been evaluated using metrics such as accuracy (acc.), sensitivity, specificity, precision, recall, f1-measure (f1), and area under the ROC curve (AUC) [36,37]. The confusion matrix summarizes the prediction outcomes for a given classification problem in terms of true positive (TP), true negative (TN), false positive (FP), and false negative (FN), which are positive examples correctly predicted by the classifier, negative examples correctly classified by the model, negative instances incorrectly classified by the model, and positive examples incorrectly classified by the model, respectively. Accuracy is the most commonly used metric in both performance prediction (55%) and dropout prediction (63%), even though accuracy is not suitable for highly imbalanced datasets.

The performance of the regression model is measured using the error or difference between the actual outcome (y_i) of an instance (i) and the estimated outcome by the model (\hat{y}_i). Several error measures have been used in this task, such as mean absolute error (MAE), root mean square error (RMSE), and root mean square error of approximation (RMSEA) [38]. In addition, statistical measures such as r^2 (R-squared) have been used to estimate the performance of regression models. Table 6 provides the metrics used in the literature.

Table 6. The performance metrics used for classification and regression.

Prediction Task	Metric	Formula
Classification	Precision (P)	$\frac{TP}{TP + FP}$
	Recall (R)	$\frac{TP}{TP + FN}$
	Accuracy	$\frac{TP + TN}{TP + TN + FN + FP}$
	F-score	$2 * \frac{R * P}{R + P}$
Regression	MSE	$\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2$
	RMSE	$\frac{1}{m} \sum_{i=1}^m \sqrt{(y_i - \hat{y}_i)^2}$
	MAE	$\frac{1}{m} \sum_{i=1}^m y_i - \hat{y}_i $
	R ²	$1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2}$

9. Student Outcomes

Researchers have explored and analyzed students' characteristics and learning behavior for various reasons, including providing personalized learning by building students' profiles, understanding students' learning styles, and optimizing educational outcomes. They have investigated a wide range of outcomes that can be grouped into two major categories: learners' performance predictions and retention and completion predictions, as success and retention rate are essential variables that institutions need to measure and assess frequently. Learner-performance prediction systems aim to predict students' performance upon the completion of a given course, such as certificate, grade category, grade value, success/failure, and risk prediction. Student-dropout prediction has attracted many researchers in the field in recent years. The literature has witnessed a large increase in contributions to this issue owing to the increased use of MOOCs. Finding more efficient approaches to mitigate online student withdrawals is of fundamental importance to institutions, students, and staff. These problems are often formulated as classification (binary or multi-class) or regression. In the following subsection, a summary of the proposed studies on student performance and student retention/dropout is presented.

9.1. Predicting Students Performance

To predict student performance, studies have examined four performance measures: certificate acquisition, grade, failure/success, and at-risk prediction. Table 7 summarizes the proposed models for student-performance prediction.

Table 7. Summary of the performance-prediction studies that include type of online education (MOOC or SPOC), platform (the dataset or the platform; blank if not mentioned), sample (the sample of instances in the dataset; blank if not mentioned), courses (number of courses in the dataset; blank if not stated), features (B: learning behavior or log data, D: demographic information, V: video click stream, A: assessment data (grade, answer data), C: course data, O: other, such as a forum, discussion, and academic background, attendance data, features extraction (FE) method, either statistical (S), temporal (T), or raw data (R), classes, output, best-performing model/s, and best performance-accuracy scores. F1-score (F1), AUC, or regression metrics is provided if the accuracy score was not reported. Scores are rounded up to two decimal places and reported with the week if the performance was estimated during different time spans. ** and * indicate the performance using different classes (binary, multi, or regression) or using different features.

Ref.	Type	Platform Dataset	Sample/Course	Features						FE			Model	Class	Output	Accuracy	
				B	D	V	A	C	O	S	T	R					
[39]	SPOC		336/1	✓							✓		SVM	Multi	Grade	0.95	
[12]	MOOC	SAPData	500/12	✓	✓					✓	✓		MLP	Multi	Grade	0.84	
[40]	SPOC		202/-	✓						✓	✓		LR	Binary	At risk	F1 = 0.66	
[41]	MOOC	edX	3530/1	✓							✓		RF	Binary	Certificate	w5:0.95	
[27]	MOOC	HMedx	597,692/15	✓	✓						✓		RF	Binary	Certificate	0.99	
[42]	MOOC		9990/1	✓	✓						✓		LSVM	Binary	Certificate	w7:0.99	
[43]	SPOC	Moodle	-/1	✓							✓		DT	Binary	Grade	w7:F1 = 0.85	
[44]	MOOC	Coursera	-/2			✓	✓				✓		RNN	Reg.	Grade	RMSE = 0.058	
[45]	MOOC	edX	18,927/15	✓					✓	✓	✓		BGL	Binary	Certificate	w3:AUC = 0.90	
[31]	MOOC		603/1	✓		✓				✓	✓		RF+ Bagging	Binary	Certificate	0.79	
[46]	MOOC	edX	5537/9	✓							✓		LR	Binary	Pass/fail	w5:0.94	
[47]	SPOC	Blackboard	-/-	✓	✓					✓	✓		GP	Binary	At risk	0.89	
[48]	MOOC		300/1	✓	✓						✓		CART	Multi	Grade	0.90	
[20]	MOOC	CAROL	3585/1	✓	✓		✓				✓		DL	Multi	Pass/fail	w8:0.98	
[49]	SPOC	Moodle	6119/1	✓**			✓*				✓		FURIA	Multi	Grade	0.76 *, 0.99 **	
[50]	SPOC	Fanya	5542/1	✓	✓		✓			✓	✓		LR* DNN**	Reg.* Multi**	Grade	MSE = 20* 0.88 **	
[34]	MOOC	UCATx, coursera	24,789/5	✓						✓	✓		SMOTE SSELM	Binary	Complete	0.97	
[51]	MOOC	HMedx OULAD	8000/6	✓	✓						✓		GBM	Binary	At risk	0.95	
[30]	SPOC		-/3	✓					✓	✓	✓		RF	Binary	Pass/fail	w1:AUC = 0.85	
[52]	MOOC		1528/1	✓			✓				✓		LSTM + DSP	Binary	Pass/fail	0.91	
[15]	MOOC	OULAD	22,437/22	✓	✓		✓				✓		GBM	Binary	Pass/fail	AUC = 0.93	
[53]	MOOC	HPU LMS	1073/1	✓	✓		✓				✓		SSL Regression	Reg.	Grade	MAE = 1.146	
[54]	SPOC	edX	124/1	✓			✓			✓	✓		TrAdaboost	Binary	At risk	AUC = 0.70	
[55]	MOOC	XuetangX	12,847/1	✓							✓		GRU-RNN	Reg.	Complete	r ² = 0.84	
[26]	MOOC	HMedx	641,138/-	✓							✓		DNN	Binary	Certificate	0.89	
[56]	SPOC		122/1	✓			✓				✓		Regression analysis	Reg.	Grade	0.85	
[57]	MOOC	Lagunita	130,000/1	✓								✓	✓	RNN** RNN*	Multi** Reg.*	Grade	w5:0.55** RMSE = 8.65 *
[58]	MOOC	HPU LMS	1073/1	✓	✓		✓			✓	✓		Multiview SSL Regression	Reg.	Grade	MAE = 1.07	
[59]	MOOC		1075/2			✓	✓				✓		SVT	Reg.	Grade	RMSE = 0.30	

Table 7. Cont.

Ref.	Type	Platform Dataset	Sample/ Course	Features					FE			Model	Class	Output	Accuracy	
				B	D	V	A	C	O	S	T					R
[60]	MOOC	edX	6241/2			✓					✓		JRIP	Binary	Grade	0.70
[33]	MOOC	edX	-/3	✓			✓				✓		RF	Binary	Grade	0.79
[21]	MOOC	CAROL	49,551/4	✓	✓		✓				✓		Ensemble	Multi	Pass/fail	w7:0.93
[61]	SPOC	Moodle	69/1							✓	✓		ARM	Binary	Grade	
[14]	MOOC	OULAD	32,593/7	✓	✓		✓	✓			✓		RF*, GB**	Binary* Multi**	At risk, Grade	0.91* 0.73**
[62]	MOOC	Moodle	66/1			✓					✓		KNN	Binary	At risk	0.65
[63]	SPOC		104/ 1			✓	✓			✓	✓		NB	Multi	Grade	0.86
[16]	MOOC	OULAD	32,593/-	✓	✓						✓	✓	CNN + LSTM	Multi	Grade	0.61
[64]	SPOC	Moodle	150/1			✓	✓			✓	✓		KNN	Multi	Grade	0.87
[38]	MOOC	Moodle	802/4	✓	✓						✓		LR	Reg.	Grade	RMSEA = 0.13
[65]	MOOC		2556/2			✓				✓	✓	✓	DNN	Reg.	Grade	w5:MAE = 6.8

9.1.1. Certificate Acquisition Prediction

Student certificate acquisition is one of the effective measures of learner performance in professional courses. In MOOCs, learners earn a certificate of completion if they finish the course and meet the course requirements. Other specialty credentials are available upon request and are often expensive. All research in this category has focused on predicting the attainment of a certificate of completion. Thus, predicting certificate and course completion attainment are considered the same.

Korosi et al. [31] developed a certificate-prediction algorithm based on features related to learning behavior, mouse behavior, video-watching attitudes, and text inputs. The gain-ratio feature selection approach was implemented using different classification algorithms, and the best performance was obtained by random forest and bagging. Similarly, Al-Shabandar et al. [27] proposed an approach that uses a random forest algorithm and hill climbing to automatically select statistical features from learners' demographical and behavioral attributes. The model was evaluated on the HMedx dataset [66] with undersampling of the majority class to solve the class imbalance problem. Imran et al. [26] proposed a deep neural network model to predict student dropout and certificate acquisition based on learner behavioral data. Liang et al. [42] proposed an approach to improve online learning by offering a personalized profile to guide student behavior. The authors explored different classification models to predict whether a student will obtain a certificate based on the Jaccard coefficient similarity of student profiles and learning behavior. The model was tested over different weeks (5–7 weeks), and the best results were obtained using the SVM model at 7 weeks. Ruipérez-Valiente et al. [41] proposed an approach that uses statistical learning behavior and progress features with several machine-learning methods to predict whether a student will obtain a certificate or not.

9.1.2. Grade Prediction

Another perspective on student success is students' grades. Several publications have proposed models for predicting student grades in various assessments, such as the course grade or final-exam grades, quizzes, or assignments in a course. Researchers have approached student-grade predictions through classification and regression models. Several studies have used binary classification to predict learner failure and success [33,43,54,60,61]. Other studies categorized grades into multiple categories, such as excellent, good, moderate, and fail [48]; withdrawn, fail, pass, and distinction [14]; and good, medium, and low [12].

Mainstream studies examined several machine-learning models with manually selected statistical learning behavior features to predict learners' grades. Xiao et al. [48] employed the classification and regression tree (CART) algorithm on statistical demo-

graphic and learning behavior data to classify students' final grades into one of four classes. Similarly, Rahman and Islam [12] developed an approach based on demographic, academic background, behavioral, and parents' participation data to classify learners' grades into three categories. They employed different algorithms with ensemble filtering, and the artificial neural network (ANN) model provided the best results. In addition, Villagra et al. [39] proposed a machine-learning model that utilized normalized learning-behavior data to classify students into the five categories of the standard grading system. An SVM model outperformed other baseline models in this task.

Adnan et al. [14] employed several machine models that use learners' demographics, and learning-behavior attributes to classify student performance into binary and multiclass categories. The results show that gradient boosting (GB) and RF scored the highest performance for multiclass and binary classification, respectively. In addition, Singh and Sachan [64] utilized learning behaviors, along with some secondary features, such as intermediate quiz performance and class attendance, to predict learner grade category. SVM and kNN were employed, and kNN showed better accuracy. Huang et al. [33] proposed a binary academic-performance prediction model using behavioral data. Several machine-learning models were tested, and the best performance was achieved using RF.

Chi and Huang [63] proposed a naive Bayes algorithm with the Laplace smoothing method to predict student performance and produce midterm-stage warnings. Using assessment, attendance, and other learning-behavior features, the model was evaluated on a one-course dataset and achieved the best accuracy in comparison with other techniques. Yu [50] developed a regression and classification models for student performance. They used statistical features from log data with a linear regression model (LR) to estimate students' grades and a deep neural network (DNN) to classify students into three classes. KokoÅg et al. [61] utilized assignment submission pattern and association rule mining (ARM) analysis to determine whether a student will pass or fail a course. Karlos et al. [58] developed a multi-view semi-supervised regression model to predict distance-learner grades based on learning activities, demographic, attendance, and assessment information. The model was trained on data obtained from Hellenic Open University (HOU) LMS.

Other studies have employed feature-selection strategies to select the most-relevant features for predictive tasks. Chen et al. [30] proposed an early predictive model that utilized clickstream events to capture student interactions with course content and other students in social learning networks, and classify students into pass or fail. Feature selection based on correlation analysis were employed for each course, and different classifiers were examined in this task. In addition, Mourdi et al. [20] developed a multiclass classifier that utilized demographic, interactive, and assessment features to determine whether a student will pass, fail, or dropout. The dimensionality was reduced using principal component analysis (PCA). In comparison to other machine-learning models, the deep neural network had the greatest accuracy. Chiu et al. [46] proposed a model to predict student performance based on behavioral data. Logistic regression and linear regression models were employed to determine whether a student would pass or fail and estimate their final grade, respectively.

Recent studies have examined temporal features and deep-learning models to predict learners' grades. Kőrösi and Farkas [57] proposed a recurrent neural network (RNN) to predict students' final grade scores and category from raw log data. For both tasks, the RNN model showed better accuracy than baseline models. In addition, Qu et al. [52] employed an LSTM model trained on temporal assignment-related attributes, such as submission order and completion time, to describe students' learning behavior and used a discriminative sequential pattern (DSP) adapter to predict students' achievement. Similarly, Song et al. [16] developed a CNN-LSTM model based on demographic and engagement patterns to predict learners' academic performance.

Fine-grained behavioral features were also examined to predict learners' grades. Lee et al. [65] presented a deep neural network model for student-performance prediction based on video-watching and exercise-answering behaviors. Similarly, Lemay and Doleck [60]

presented a grade-prediction model for weekly assignments based on statistical video-viewing behavior (e.g., the number of fast-forwards, pauses, and playback). Different classifiers were employed, and JRip achieved the best performance. Similarly, Yang et al. [44] proposed an RNN grade-prediction model utilizing video-watching data and assessment grades, such as the number of pauses, rewinds, and average quiz grades. Fuzzy-based approaches have also been investigated for this problem. Hussain et al. [49] employed FURIA [67], among other models, to predict learners who will perform poorly on tests based on learning activities and grades.

Several studies have considered grade prediction as a regression problem. Chunzi et al. [56] explored factors that affect college students' online foreign-language learning performance. The author focused on learning-ability attributes reflected by students' grades of text recitation and translation before and after the course, duration of learning, login times, and the scores obtained in online quizzes. Students' online answers followed by login times were found to be the most-significant variables for estimating students' grades. Xiao et al. [59] developed a matrix-completion-based model to predict online and offline unit tests and final grades. They organized the MOOC data into a matrix with missing values and then applied the singular value thresholding (SVT) algorithm to obtain the values of missing grades. Kostopoulos et al. [53] developed a model to predict students' grade values using semi-supervised learning (SSL) regression.

9.1.3. Students-at-Risk Prediction

Identifying students who are at risk of failing a course is a prevalent objective in the literature. Kondo et al. [40] used learning interaction and the attendance attributes of SPOCs dataset to identify off-task and at risk students. Cano and Leonard [47] presented an early-warning system based on genetic programming (GP) rules to address students with socioeconomic disadvantages. The system automatically extracts classification rules based on student demographic, learning interaction, academic background, registration, and family and socioeconomic data. Similarly, Al-Shabandar et al. [51] presented an approach that identifies failure and withdrawal among MOOC online learners based on demographic and learning-behavior data. The synthetic minority oversampling approach (SMOSE) was applied to address the class imbalance issue, and several machine-learning models were examined, including gradient boosting machine (GBM), which provided the highest accuracy. Wan et al. [54] proposed a transfer-based model to predict at-risk students on a weekly in-class project test. Statistical behavioral features, such as the total time spent on video resources and percentage of total correct submissions, were extracted from the data and combined with previously learned weights from former courses to predict students at risk of failing in the current course. El Aouifi et al. [62] proposed a model to predict final grade based on students' interactions with pedagogical sequences of educational video behaviors such as play, pause, jump forward, jump backward, and end. The author extracted sequences as features according to the pedagogical sequences in which the student performed an action. K-nn and MLP were used, and the best results were achieved using K-nn.

9.2. Student Dropout and Retention Prediction

Dropping out and failing to complete a course are serious challenges for educational institutions, which is why several studies have focused on these topics. Various methods for feature extraction have been proposed. Most studies examined different machine-learning methods trained on manually selected statistical features to predict dropout in MOOCs [68] or used temporal features obtained from raw or statistical data. Table 8 summarizes the dropout prediction approaches.

9.2.1. Statistical Features

Most of the proposed approaches utilize learning-behavior data with demographic or academic background to predict dropouts. Monllaş Olivá et al. [69] presented a

framework that explored multiple prediction analyzers (students, users, or courses), classification targets (at-risk students, course effectiveness), indicator features (e.g., the number of clicks made by the student), and machine-learning models (ANN or LR). To predict student dropout, learning behavior indicators were extracted and used to train the ANN model, which yielded a good performance. Similarly, Zhou and Xu [24] developed a multi-model stacking ensemble learning (MMSE) approach to predict student dropout in MOOCs. The classification was performed in two stages; first, five base classifiers were employed and produced an output that an XGBoost model examined to produce the final prediction. Assessment data has also been examined as indicators of learner dropout rates. Burgos et al. [70] developed a course-specific logistic regression model to predict student dropouts. Similarly, Jha et al. [15] implemented several machine-learning algorithms with a different set of features to predict student dropout and performance. The authors examined demographics, assessment, and learning-interaction data separately and learner interaction provided the highest accuracy.

Several studies have incorporated course data to predict dropout in MOOCs. Cobos and Olmos [45] utilized course and log data to predict dropout and certificate acquisitions. Bayesian generalized linear (BGL) and stochastic gradient boosting (SGB) achieved the best accuracy among other classifiers for certificate and dropout prediction, respectively. Laveti et al. [71] proposed a stacked ensemble model with features extracted from course, log, and enrollment data for dropout prediction. Similarly, Coussement et al. [72] developed a logit leaf model (LLM) trained on data obtained from student demographics, cognitive, engagement, and academic attributes in addition to classroom characteristics for dropout prediction. Hong et al. [73] proposed a two-layer cascading classifier that combines three classifiers for student dropout prediction in MOOCs.

Early predictions of learner dropout have also been studied in the literature. Panagiotakopoulos et al. [74] presented an early-dropout-prediction model based on students' demographics, academic background, and first week interaction collected from a single MOOC course. Several machine-learning models were examined, and LightGBM outperformed the other models. In addition, Alamri et al. [75] proposed a next-week dropout-prediction model that predicts students who do not access 80% of the course content in the following week based on their learning behavior and expressed opinions in the discussion forum. Multiple classification models were tested and AdaBoost provided the highest accuracy. A study by Xing and Du [76] used the weekly accumulated statistical log data with a deep neural network model to predict student dropouts on a weekly basis.

Likewise, Liu et al. [77] proposed a temporal-behavior features weighting approach that can be incorporated with other machine-learning models to predict dropout at any point during the course. The proposed approach was compared with other feature generation methodologies, such as non-stacked [78,79], stacked [78,80,81], aggregated [79,82] and recursive-aggregated [54] approaches. The proposed method achieved a higher F-score and lower false-positive rates in early-weeks prediction. Studies have also examined feature-selection strategies in dropout prediction. Qiu et al. [83] developed a dropout-prediction model that integrates ensemble feature ranking based on MI, RF, and RFE. The model also incorporates a forward search method that gradually examines features that improve prediction. The selected features were tested with a LR model. Ardchir et al. [84] used different machine-learning models with feature selection based on information gain RF to weight features based on their importance. Neighborhood component analysis feature selection was used with several machine-learning models to predict student dropout [85].

Optimization techniques have also been used for dropout prediction. Jin [86] proposed an approach that initiates the weight of predictive models according to the maximum neighborhood definition of the training sample using an improved quantum particle swarm optimization (IQPSO) algorithm. To overcome the class imbalance issue, Mulyani et al. [87] proposed the SMOTE-ensemble learning (SEL) model for dropout prediction using the synthetic minority over-sampling technique (SMOTE) for the non-dropout class in addition to the ensemble learning technique with three machine learning algorithms, LR, KNN, and RF.

Gregori et al. [34] also used the synthetic minority oversampling (SMOTE) algorithm with a semi-supervised extreme learning machine (SSELM) model.

9.2.2. Temporal Features

Recent studies on dropout prediction have used temporal features extracted from raw or statistical data and employed deep-learning techniques to estimate learning behavior dynamics over time. Lai et al. [23] developed a random vector functional link neural network (RVFLNN) to predict student dropout in MOOCs. MOOCs generate periodical data that varies in dimensionality; therefore, RVFLNN were used to increase input and update the weight dynamically. Likewise, Xiong et al. [88] proposed an RNN-LSTM model to predict student dropouts in MOOCs based on statistical weekly features aggregated from students' log data. Sun et al. [55] developed a GRU-RNN model to predict the percentage of course content completed in an entire course syllabus. Several dropout-prediction models use CNNs to extract features from raw data. Wu et al. [89] proposed a model that combined CNN, LSTM, and SVM to predict student dropout on a daily basis. The model transforms raw data into a one-hot vector and represents each enrollment in a number of matrices that each describe the learner's behavior in a single day. The proposed model outperformed other standard machine-learning methods, such as SVM, RF, and LR, and temporal deep-learning techniques, such as CNN-RNN.

Several studies have developed CNN model to extract features from statistical temporal features and predict dropout. Studies in [90,91] developed a CNN model to analyze the students' learning behavior represented by a matrix that encodes the daily statistical features over a period of time. Likewise, Zheng et al. [92] presented a model that extracts statistical learning-behavior data, performs feature weighting and selection based on a decision tree, and then uses CNN to predict dropout from students' time-series learning behavior frequencies. Recently, several studies have proposed connectionist approaches to estimate learning outcomes by combining CNN and sequence models (LSTM). Mubarak et al. [19] proposed a hypermodel of CNN and LSTM to extract features from the raw data of MOOCs and predict student dropout. Wang et al. [93] proposed a dropout-prediction model with automatic feature extraction from the raw MOOC data. The raw data of each event were first converted into a one-hot vector and concatenated to produce matrices representing a student's learning behavior over a number of days. Then, a CNN and RNN model was used to extract features and utilize time-series relationships. Similarly, Fu et al. [94] developed a novel approach that employs a CNN to extract local features from raw clickstream data and LSTM to obtain a time-series incorporating vector representation. Yin et al. [95] designed a novel neural network structure that combined the convolutional attention mechanism and conditional random field (CRF) for MOOC dropout prediction. The proposed approach was compared with several baseline classifiers, including CNN-LSTM, and significantly outperformed the baseline methods.

Table 8. Summary of the dropout prediction studies that include type of online education (MOOC or SPOC), platform (the dataset or the platform, blank if not mentioned), sample (the sample of instances in the dataset, blank if not mentioned), courses (number of courses in the dataset, blank if not stated), features (B: learning behavior or log data, D: demographic information, V: video click stream, A: assessment data (grade, answer data), C: course data, O: other, such as a forum, discussion, and academic background, attendance data, features extraction (FE) method either statistical (S), temporal (T), and raw data(R), classes, output, best-performing model/s, and best performance accuracy scores. F1-score (F1), AUC, or regression metrics is provided if the accuracy score was not reported. Scores are rounded up to two decimal places and reported with the week if the performance was estimated during different time spans.

Ref.	Type	Platform Dataset	Sample Course	Features						FE			Model	Class	Output	Accuracy
				B	D	V	A	C	O	S	T	R				
[73]	MOOC	KDDcup	96,529/39	✓				✓		✓			c-RF	Binary	Dropout	0.93
[71]	MOOC	KDDcup	200,000/39	✓	✓			✓		✓			Stacked ensemble	Binary	Dropout	0.91
		CNN + RNN	Binary	Dropout	0.92											
[76]	MOOC	Canvas	3617/1	✓			✓		✓	✓	✓		DL	Binary	Dropout	w7:0.97
[83]	MOOC	KDDcup	120,542/39	✓					✓	✓			RFE + LR	Binary	Dropout	0.87
[88]	MOOC	LMS	-/6	✓						✓	✓		RNN-LSTM	Binary	Dropout	0.90
[89]	MOOC	KDDcup	79,186/39	✓						✓	✓		CNN + LSTM + SVM	Binary	Dropout	F1 = 0.95
[84]	MOOC	KDDcup	112,448/39	✓			✓	✓	✓	✓			GB	Binary	Dropout	AUC = 0.89
[15]	MOOC	OULAD	32,594/22	✓	✓		✓			✓			GBM	Binary	Dropout	AUC = 0.91
[26]	MOOC	HMedx	641,138/-	✓						✓			DNN	Binary	Dropout	0.99
[85]	MOOC	OULAD	32,593/22	✓	✓		✓			✓			ANN	Binary	Dropout	AUC = 0.93
[72]	MOOC	-	10,554/-	✓	✓			✓	✓	✓			LLM	Binary	Dropout	F1 = 0.84
[92]	MOOC	KDDcup	79,186/39	✓						✓	✓		FWTS-CNN	Binary	Dropout	0.87
[77]	MOOC	KDDcup	112,448/39	✓						✓			Gaussian NB	Binary	Dropout	F1 = 0.85
[96]	MOOC	KDDcup	120542/39	✓				✓		✓	✓		CNN + SE + GRU	Binary	Dropout	0.95
[69]	MOOC	Moodle	46,895/8	✓						✓			ANN	Binary	Dropout	0.89
[23]	MOOC	KDDcup	120,542/39	✓						✓	✓		RVFLNN	Binary	Dropout	0.93
[86]	MOOC	KDDcup	53,596/6	✓						✓			SVM	Binary	Dropout	F1 = 0.90
[24]	MOOC	KDDcup	120,542/39	✓						✓	✓		MMSE	Binary	Dropout	0.88
[90]	MOOC	KDDcup	120,542/39	✓						✓	✓		CNN	Binary	Dropout	0.88
[95]	MOOC	KDDcup	12,004/1	✓						✓	✓		Attention + CRF	Binary	Dropout	0.84
[74]	MOOC	Moodle	700/1	✓	✓		✓			✓			LightGBM	Binary	Dropout	0.96
[75]	MOOC	Future Learn	251,662/7	✓						✓	✓		RF, Adaboost	Binary	Dropout	0.95

10. Summary, Challenges and Limitations

As demonstrated in Tables 7 and 8, most of the studies were conducted on a dataset obtained from a MOOC. Only a limited number of studies considered predicting student outcomes in SPOCs. Researchers have utilized data collected for different durations, for example, a month, two months, a semester, or several years. In addition, researchers have developed predictive models using the learner-interaction data from a whole course or across multiple course periods (e.g., three weeks, five weeks, seven weeks). Assessing predictive models on different course periods shows that models give better predictive results of student outcomes when the evaluation duration increases. In addition, the samples and number of courses vary among the studies. Some studies worked on a subset of a publicly available dataset, whereas others worked on the entire dataset. The SPOC sample size was much less than the MOOC sample size, which impacted the performance of the

SPOC-based models. Similar to any machine-learning problem, the size of the dataset (i.e., number of samples) can impact the performance of the model.

Various sets of indicators have been used to predict student outcomes in the literature. The most widely employed group of information in predicting student performance was learning behavior (log data), followed by assessment and demographic features. Learning behavior was also the most widely utilized set of features in predicting learner dropout; demographic and assessment features were less-frequently used. Statistically extracted characteristics were studied more than raw features by the researchers. The number of statistically retrieved features differs among research as well. Some studies focused on a small number of features (e.g., five features [42]), whereas others examined a larger number of features (e.g., 79 features [83]). Those features were often selected manually, and automatic feature selection was less-widely utilized. Several machine-learning models were evaluated in this task. RF and ANN were among the best-performing machine-learning models.

Several limitations were observed in the literature, which can be summarized as follows:

- There is no consensus among researchers on the definition of dropout, success, and other related terminologies. For example, some researchers considered a student to be a dropout if they fail to complete a specific percentage of the assessments (e.g., 50%) [70], or if they were not active for several consecutive days [73]. Others considered dropout to be the inability to pass a course, and some did not provide any precise definition. The inconsistency in defining dropout and other related terminologies could be a concern to researchers, since it influences how dropout is assessed, addressed, and investigated [97].
- Most proposed approaches use MOOC datasets in which students are self-motivated and are not required or obligated to participate in the courses. Thus, there is an enormous disparity between learners who are registered for curiosity, who, for example, view some videos and do not complete the required assignments, and those who are registered to finish the course, which may make identifying dropouts or failure relatively easy tasks. Exploring other types of online environments, such as SPOCs, might introduce more challenges, as learners may attempt to finish the course but fail or withdraw due to their lack of knowledge or other reasons.
- Most studies also proposed systems that predict student outcomes when students approach the end of the course. Only a small number of studies considered early predictions of user outcomes, while many studies did not report the duration of data collection. This made it difficult to compare the different proposed methods, as the duration of the extracted features varied significantly. In addition, some studies used a subset of publicly available datasets, making it difficult to compare different methods on benchmark datasets.
- Most studies employed feature-engineering techniques to calculate students' statistical features as a whole. These features were often chosen arbitrarily or using statistical techniques such as correlation analysis. Some recent studies looked into methods that automatically extract temporal features from raw data by mapping raw features into numerical representations, such as one-hot encoding, and then using deep-learning methods, in particular, convolution functions, to extract features. Despite the usefulness of these techniques, the generated representations are often very sparse and less useful for comparing user behavior.
- Several limitations have been addressed by researchers in this field. One of these limitations is the problem of multi-valued instances, in which some instances containing the same patterns have different outcomes [98]. In addition, the dropout-prediction task is well known to be imbalanced because the proportion of the positive class is much larger than that of the negative class. Several studies also use methods to handle the problem of class imbalance by either oversampling of the minority class [34,87] or under-sampling of the majority class [66].

- One of the observed challenges is the quality of the training samples in which a large number of attributes are clickstream data which might be less representative when learners do not interact or engage in learning activities. This practice is common among students who register in MOOCs for curiosity [86].
- Learning analytic systems are based on large volumes of real-time data collected over a protracted period. Analyzing, combining, and linking multiple forms of learners' data to forecast their outcomes or any part of the learning process has raised many ethical challenges that cannot be properly measured using traditional ethical procedures [99]. The risks of de-anonymizing learners' identities [100] and decontextualizing data [101] are some of the potential risks of this practice.

11. Conclusions and Future Directions

With the emergence of MOOCs and the expansion of online education in recent years, the prediction of learners' outcomes in online environments has attracted considerable attention. The spread of COVID-19 has also aided the growth of SPOCs, blended education, and an interest in monitoring student engagement and performance. Therefore, this study reviewed current strategies for predicting online-student outcomes in MOOCs and SPOCs. It summarized the predictive variables, online learning platforms, feature extraction, selection techniques, evaluation metrics, and the predictive models employed in this area. It also provided a thorough analysis and taxonomy for related research. Throughout our analysis, we found that most studies in the field utilized statistical features such as the number of downloaded materials and duration of video watching in a given time period. A small number of studies examined statistical temporal and raw temporal features in predicting learner outcomes. Studies conducted on benchmark datasets showed that statistical temporal features provide better results than raw features. Thus, further investigation of temporal features will provide a valuable understanding of users' learning progress and, eventually, their learner outcomes. Most temporally based LSTM or GRU models learners' time-series features. Further investigation of other recent sequence-based models, such as the attention-based model, is required. Studies using one-hot encoding to represent raw features and different representation techniques for raw features are worth investigating. Different machine-learning and deep-learning models have been used to predict learners' outcomes. RF and ANN are among the most effective machine learning models' performance and dropout prediction, whereas the sequence-based model provides the best performance on the publicly available dropout dataset. Further investigation of deep-learning models is recommended to predict student performance.

Author Contributions: Conceptualization, A.A., H.A., T.A., and M.A.; methodology, A.A. and M.A.; validation, A.A., M.A., and H.A.; formal analysis, M.A.; Data curation, M.A.; writing—original draft preparation, M.A. and A.A.; writing—review and editing, A.A., H.A., and T.A.; visualization, M.A. and A.A.; project administration, H.A. and T.A.; funding acquisition, H.A.; All authors have read and agreed to the published version of the manuscript.

Funding: The authors extend their appreciation to the Deputyship for Research and Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number IFPRC-039-126-2020 and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Arce, M.E.; Crespo, B.; Míguez-Álvarez, C. Higher Education Drop-out in Spain—Particular Case of Universities in Galicia. *Int. Educ. Stud.* **2015**, *8*, 247–264. [[CrossRef](#)]
2. Xavier, M.; Meneses, J. *Dropout in Online Higher Education: A Scoping Review from 2014 to 2018*; ELearn Center, Universitat Oberta de Catalunya: Barcelona, Spain, 2020.

3. Baker, R.S.; Inventado, P.S. Educational Data Mining and Learning Analytics. In *Learning Analytics: From Research to Practice*; Larusson, J.A., White, B., Eds.; Springer: New York, NY, USA, 2014; pp. 61–75. [CrossRef]
4. Moreno-Marcos, P.M.; Alario-Hoyos, C.; Mu  oz-Merino, P.J.; Kloos, C.D. Prediction in MOOCs: A Review and Future Research Directions. *IEEE Trans. Learn. Technol.* **2019**, *12*, 384–401. [CrossRef]
5. Ranjeeth, S.; Latchoumi, T.; Paul, P.V. A survey on predictive models of learning analytics. *Procedia Comput. Sci.* **2020**, *167*, 37–46. [CrossRef]
6. Hamim, T.; Benabbou, F.; Sael, N. Survey of Machine Learning Techniques for Student Profile Modelling. *Int. J. Emerg. Technol. Learn.* **2021**, *16*, 136–151. [CrossRef]
7. Prenkaj, B.; Velardi, P.; Stilo, G.; Distante, D.; Faralli, S. A survey of machine learning approaches for student dropout prediction in online courses. *ACM Comput. Surv. (CSUR)* **2020**, *53*, 1–34. [CrossRef]
8. Gardner, J.; Brooks, C. Student success prediction in MOOCs. *User Model. User-Adapt. Interact.* **2018**, *28*, 127–203. [CrossRef]
9. Katarya, R.; Gaba, J.; Garg, A.; Verma, V. A review on machine learning based student’s academic performance prediction systems. In Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 25–27 March 2021; pp. 254–259. [CrossRef]
10. Filius, R.M.; Uijl, S.G. Teaching Methodologies for Scalable Online Education. In *Handbook for Online Learning Contexts: Digital, Mobile and Open*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 55–65.
11. Amrieh, E.A.; Hamtini, T.M.; Aljarah, I. Mining Educational Data to Predict Student’s academic Performance using Ensemble Methods. *Int. J. Database Theory Appl.* **2016**, *9*, 119–136. [CrossRef]
12. Rahman, M.H.; Islam, M.R. Predict Student’s Academic Performance and Evaluate the Impact of Different Attributes on the Performance Using Data Mining Techniques. In Proceedings of the 2017 2nd International Conference on Electrical Electronic Engineering (ICEEE), Rajshahi, Bangladesh, 27–29 December 2017; pp. 1–4. [CrossRef]
13. Kuzilek, J.; Hlosta, M.; Zdrahal, Z. Open university learning analytics dataset. *Sci. Data* **2017**, *4*, 170171. [CrossRef]
14. Adnan, M.; Habib, A.; Ashraf, J.; Mussadiq, S.; Raza, A.A.; Abid, M.; Bashir, M.; Khan, S.U. Predicting at-Risk Students at Different Percentages of Course Length for Early Intervention Using Machine Learning Models. *IEEE Access* **2021**, *9*, 7519–7539. [CrossRef]
15. Jha, N.; Ghergulescu, I.; Moldovan, A. OULAD MOOC Dropout and Result Prediction using Ensemble, Deep Learning and Regression Techniques. In Proceedings of the 11th International Conference on Computer Supported Education, Heraklion, Greece, 2–4 May 2019; SciTePress: Set  bal, Portugal, 2019; Volume 2, pp. 154–164. [CrossRef]
16. Song, X.; Li, J.; Sun, S.; Yin, H.; Dawson, P.; Doss, R.R.M. SEPN: A Sequential Engagement Based Academic Performance Prediction Model. *IEEE Intell. Syst.* **2021**, *36*, 46–53. [CrossRef]
17. Hussain, M.; Zhu, W.; Zhang, W.; Abidi, R. Student Engagement Predictions in an e-Learning System and Their Impact on Student Course Assessment Scores. *Comput. Intell. Neurosci.* **2018**, *2018*, 6347186. [CrossRef] [PubMed]
18. Stanford, U. Center for Advanced Research through Online Learning (CAROL). Available online: <https://carol.stanford.edu> (accessed on 24 August 2021).
19. Mubarak, A.A.; Cao, H.; Hezam, I.M. Deep analytic model for student dropout prediction in massive open online courses. *Comput. Electr. Eng.* **2021**, *93*, 107271. [CrossRef]
20. Mourdi, Y.; Sadgal, M.; El Kabtane, H.; Berrada Fathi, W. A machine learning-based methodology to predict learners’ dropout, success or failure in MOOCs. *Int. J. Web Inf. Syst.* **2019**, *15*, 489–509. [CrossRef]
21. Mourdi, Y.; Sadgal, M.; Berrada Fathi, W.; El Kabtane, H. A Machine Learning Based Approach to Enhance Mooc Users’ Classification. *Turk. Online J. Distance Educ.* **2020**, *21*, 47–68. [CrossRef]
22. KDD. KDD Cup 2015. Available online: <https://kdd.org/kdd-cup> (accessed on 24 August 2021).
23. Lai, S.; Zhao, Y.; Yang, Y. Broad Learning System for Predicting Student Dropout in Massive Open Online Courses. In Proceedings of the 2020 8th International Conference on Information and Education Technology, Okayama, Japan, 28–30 March 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 12–17. [CrossRef]
24. Zhou, Y.; Xu, Z. Multi-Model Stacking Ensemble Learning for Dropout Prediction in MOOCs. *J. Phys. Conf. Ser.* **2020**, *1607*, 012004. [CrossRef]
25. Ho, A.D.; Reich, J.; Nesterko, S.; Seaton, D.T.; Mullaney, T.; Waldo, J.; Chuang, I. *HarvardX and MITx: The first year of Open Online Courses*; HarvardX and MITx Working Paper No. 1; Harvard University: Cambridge, MA, USA, 2014.
26. Imran, A.; Dalipi, F.; Kastrati, Z. Predicting Student Dropout in a MOOC: An Evaluation of a Deep Neural Network Model. In Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence, Bali, Indonesia, 19–22 April 2019; pp. 190–195. [CrossRef]
27. Al-Shabandar, R.; Hussain, A.; Laws, A.; Keight, R.; Lunn, J.; Radi, N. Machine learning approaches to predict learning outcomes in Massive open online courses. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 713–720. [CrossRef]
28. Liu, K.F.-R.; Chen, J.-S. Prediction and assessment of student learning outcomes in calculus a decision support of integrating data mining and Bayesian belief networks. In Proceedings of the 2011 3rd International Conference on Computer Research and Development, Shanghai, China, 11–13 March 2011; Volume 1, pp. 299–303. [CrossRef]
29. Wu, W.H.; Jim Wu, Y.C.; Chen, C.Y.; Kao, H.Y.; Lin, C.H.; Huang, S.H. Review of Trends from Mobile Learning Studies: A Meta-Analysis. *Comput. Educ.* **2012**, *59*, 817–827. [CrossRef]

30. Chen, W.; Brinton, C.G.; Cao, D.; Mason-Singh, A.; Lu, C.; Chiang, M. Early Detection Prediction of Learning Outcomes in Online Short-Courses via Learning Behaviors. *IEEE Trans. Learn. Technol.* **2019**, *12*, 44–58. [[CrossRef](#)]
31. Korosi, G.; Esztelecki, P.; Farkas, R.; TÁs̄th, K. Clickstream-Based outcome prediction in short video MOOCs. In Proceedings of the 2018 International Conference on Computer, Information and Telecommunication Systems (CITS), Colmar, France, 11–13 July 2018; pp. 1–5. [[CrossRef](#)]
32. Pereira, F.D.; Oliveira, E.; Cristea, A.; Fernandes, D.; Silva, L.; Aguiar, G.; Alamri, A.; Alshehri, M. Early Dropout Prediction for Programming Courses Supported by Online Judges. In *Artificial Intelligence in Education*; Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B., Luckin, R., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 67–72.
33. Huang, A.Y.; Lu, O.H.; Huang, J.C.; Yin, C.J.; Yang, S.J. Predicting students' academic performance by using educational big data and learning analytics: Evaluation of classification methods and learning logs. *Interact. Learn. Environ.* **2020**, *28*, 206–230. [[CrossRef](#)]
34. Gregori, E.B.; Zhang, J.; Galván-Fernández, C.; de Asís Fernández-Navarro, F. Learner support in MOOCs: Identifying variables linked to completion. *Comput. Educ.* **2018**, *122*, 153–168. [[CrossRef](#)]
35. Liu, D.; Zhang, Y.; Zhang, J.; Li, Q.; Zhang, C.; Yin, Y. Multiple Features Fusion Attention Mechanism Enhanced Deep Knowledge Tracing for Student Performance Prediction. *IEEE Access* **2020**, *8*, 194894–194903. [[CrossRef](#)]
36. Faraggi, D.; Reiser, B. Estimation of the area under the ROC curve. *Stat. Med.* **2002**, *21*, 3093–3106. [[CrossRef](#)] [[PubMed](#)]
37. Hutter, F.; Kotthoff, L.; Vanschoren, J. *Automated Machine Learning: Methods, Systems, Challenges*; Springer Nature: Berlin/Heidelberg, Germany, 2019.
38. Bravo-Agapito, J.; Romero, S.J.; Pamplona, S. Early prediction of undergraduate Student's academic performance in completely online learning: A five-year study. *Comput. Hum. Behav.* **2021**, *115*, 106595. [[CrossRef](#)]
39. Villagra-Arnedo, C.J.; Gallego-Duran, F.J.; Compan, P.; Llorens Largo, F.; Molina-Carmona, R. Predicting academic performance from Behavioural and learning data. *Int. J. Des. Nat. Ecodynamics* **2016**, *11*, 239–249. [[CrossRef](#)]
40. Kondo, N.; Okubo, M.; Hatanaka, T. Early Detection of At-Risk Students Using Machine Learning Based on LMS Log Data. In Proceedings of the 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), Shizuoka, Japan, 9–13 July 2017; pp. 198–201. [[CrossRef](#)]
41. Ruipérez-Valiente, J.A.; Cobos, R.; Muñoz-Merino, P.J.; Andujar, Á.; Delgado Kloos, C. Early prediction and variable importance of certificate accomplishment in a MOOC. In *European Conference on Massive Open Online Courses*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 263–272.
42. Liang, K.; Zhang, Y.; He, Y.; Zhou, Y.; Tan, W.; Li, X. Online Behavior Analysis-Based Student Profile for Intelligent E-Learning. *J. Electr. Comput. Eng.* **2017**, *2017*, 9720396. [[CrossRef](#)]
43. Liu, W.; Wu, J.; Gao, X.; Feng, K. An early warning model of student achievement based on decision trees algorithm. In Proceedings of the 2017 IEEE 6th International Conference on Teaching, Assessment, and Learning for Engineering (TALE), Hong Kong, China, 21–14 December 2017; pp. 517–222. [[CrossRef](#)]
44. Yang, T.Y.; Brinton, C.G.; Joe-Wong, C.; Chiang, M. Behavior-Based Grade Prediction for MOOCs Via Time Series Neural Networks. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 716–728. [[CrossRef](#)]
45. Cobos, R.; Olmos, L. A Learning Analytics Tool for Predictive Modeling of Dropout and Certificate Acquisition on MOOCs for Professional Learning. In Proceedings of the 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Bangkok, Thailand, 16–19 December 2018; pp. 1533–1537. [[CrossRef](#)]
46. Chiu, Y.C.; Hsu, H.J.; Wu, J.; Yang, D.L. Predicting student performance in MOOCs using learning activity data. *J. Inf. Sci. Eng.* **2018**, *34*, 1223–1235. [[CrossRef](#)]
47. Cano, A.; Leonard, J.D. Interpretable Multiview Early Warning System Adapted to Underrepresented Student Populations. *IEEE Trans. Learn. Technol.* **2019**, *12*, 198–211. [[CrossRef](#)]
48. Xiao, B.; Liang, M.; Ma, J. The Application of CART Algorithm in Analyzing Relationship of MOOC Learning Behavior and Grades. In Proceedings of the 2018 International Conference on Sensor Networks and Signal Processing (SNSP), Xi'an, China, 28–31 October 2018; pp. 250–254. [[CrossRef](#)]
49. Hussain, M.; Hussain, S.; Zhang, W.; Zhu, W.; Theodorou, P.; Abidi, S.M.R. Mining Moodle Data to Detect the Inactive and Low-Performance Students during the Moodle Course. In Proceedings of the 2nd International Conference on Big Data Research, Hangzhou, China, 18–20 May 2018; Association for Computing Machinery: New York, NY, USA, 2018; pp. 133–140. [[CrossRef](#)]
50. Yu, C. SPOC-MFLP: A multi-feature learning prediction model for SPOC students using machine learning. *J. Appl. Sci. Eng.* **2018**, *21*, 279–290. [[CrossRef](#)]
51. Al-Shabandar, R.; Hussain, A.J.; Liatsis, P.; Keight, R. Detecting At-Risk Students With Early Interventions Using Machine Learning Techniques. *IEEE Access* **2019**, *7*, 149464–149478. [[CrossRef](#)]
52. Qu, S.; Li, K.; Wu, B.; Zhang, S.; Wang, Y. Predicting Student Achievement Based on Temporal Learning Behavior in MOOCs. *Appl. Sci.* **2019**, *9*, 5539. [[CrossRef](#)]
53. Kostopoulos, G.; Kotsiantis, S.; Fazakis, N.; Koutsonikos, G.; Pierrakeas, C. A semi-supervised regression algorithm for grade prediction of students in distance learning courses. *Int. J. Artif. Intell. Tools* **2019**, *28*, 1940001. [[CrossRef](#)]
54. Wan, H.; Liu, K.; Yu, Q.; Gao, X. Pedagogical Intervention Practices: Improving Learning Engagement Based on Early Prediction. *IEEE Trans. Learn. Technol.* **2019**, *12*, 278–289. [[CrossRef](#)]

55. Sun, D.; Mao, Y.; Du, J.; Xu, P.; Zheng, Q.; Sun, H. Deep Learning for Dropout Prediction in MOOCs. In Proceedings of the 2019 Eighth International Conference on Educational Innovation through Technology (EITT), Biloxi, MS, USA, 27–31 October 2019; pp. 87–90. [\[CrossRef\]](#)
56. Chunzi, S.; Xuanren, W.; Ling, L. The Application of Big Data Analytics in Online Foreign Language Learning among College Students: Empirical Research on Monitoring the Learning Outcomes and Predicting Final Grades. In Proceedings of the 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), Taiyuan, China, 23–25 October 2020; pp. 266–269. [\[CrossRef\]](#)
57. Kőrösi, G.; Farkas, R. MOOC Performance Prediction by Deep Learning from Raw Clickstream Data. In *International Conference on Advances in Computing and Data Sciences*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 474–485. [\[CrossRef\]](#)
58. Karlos, S.; Kostopoulos, G.; Kotsiantis, S. Predicting and Interpreting Students' Grades in Distance Higher Education through a Semi-Regression Method. *Appl. Sci.* **2020**, *10*, 8413. [\[CrossRef\]](#)
59. Xiao, F.; Li, Q.; Huang, H.; Sun, L.; Xu, X. MOLEAS: A Multi-stage Online Learning Effectiveness Assessment Scheme in MOOC. In Proceedings of the 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), Takamatsu, Japan, 8–11 December 2020; pp. 31–38. [\[CrossRef\]](#)
60. Lemay, D.; Doleck, T. Grade prediction of weekly assignments in MOOCs: Mining video-viewing behavior. *Educ. Inf. Technol.* **2020**, *25*, 1333–1342. [\[CrossRef\]](#)
61. Kokoňg, M.; Akňgapňsnar, G.; Hasnine, M. Unfolding Students' Online Assignment Submission Behavioral Patterns using Temporal Learning Analytics. *Educ. Technol. Soc.* **2021**, *24*, 223–235.
62. El Aouifi, H.; El Hajji, M.; Es-saady, Y.; Hassan, D. Predicting learner's performance through video sequences viewing behavior analysis using educational data-mining. *Educ. Inf. Technol.* **2021**, *26*, 5799–5814. [\[CrossRef\]](#)
63. Chi, D.; Huang, Y. Research on Application of Online Teaching Performance Prediction Based on Data Mining Algorithm. In Proceedings of the 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 15–17 January 2021; pp. 394–397. [\[CrossRef\]](#)
64. Singh, A.; Sachan, A. Student Clickstreams Activity Based Performance of Online Course. In *International Conference on Artificial Intelligence and Sustainable Computing*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 242–253. [\[CrossRef\]](#)
65. Lee, C.A.; Tzeng, J.W.; Huang, N.F.; Su, Y.S. Prediction of Student Performance in Massive Open Online Courses Using Deep Learning System Based on Learning Behaviors. *Educ. Technol. Soc.* **2021**, *24*, 130–146.
66. HarvardX. HarvardX Person-Course Academic Year 2013 De-Identified Dataset, Version 3.0; Harvard University: Cambridge, MA, USA, 2014. [\[CrossRef\]](#)
67. Hühn, J.; Hüllermeier, E. FURIA: An algorithm for unordered fuzzy rule induction. *Data Min. Knowl. Discov.* **2009**, *19*, 293–319. [\[CrossRef\]](#)
68. Mubarak, A.A.; Cao, H.; Ahmed, S.A. Predictive learning analytics using deep learning model in MOOCs' courses videos. *Educ. Inf. Technol.* **2021**, *26*, 371–392.
69. Monllaň OlivÁl, D.; Huynh, D.; Reynolds, M.; Dougiamas, M.; Wiese, D. A supervised learning framework: Using assessment to identify students at risk of dropping out of a MOOC. *J. Comput. High. Educ.* **2020**, *32*. [\[CrossRef\]](#)
70. Burgos, C.; Campanario, M.L.; de la Peña, D.; Lara, J.A.; Lizcano, D.; Martínez, M.A. Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout. *Comput. Electr. Eng.* **2018**, *66*, 541–556. [\[CrossRef\]](#)
71. Laveti, R.N.; Kuppili, S.; Ch, J.; Pal, S.N.; Babu, N.S.C. Implementation of learning analytics framework for MOOCs using state-of-the-art in-memory computing. In Proceedings of the 2017 5th National Conference on E-Learning E-Learning Technologies (ELELTECH), Hyderabad, India, 3–4 August 2017; pp. 1–6. [\[CrossRef\]](#)
72. Coussement, K.; Phan, M.; De Caigny, A.; Benoit, D.F.; Raes, A. Predicting student dropout in subscription-based online learning environments: The beneficial impact of the logit leaf model. *Decis. Support Syst.* **2020**, *135*, 113325. [\[CrossRef\]](#)
73. Hong, B.; Wei, Z.; Yang, Y. Discovering learning behavior patterns to predict dropout in MOOC. In Proceedings of the 2017 12th International Conference on Computer Science and Education (ICCSE), Houston, TX, USA, 22–25 August 2017; pp. 700–704. [\[CrossRef\]](#)
74. Panagiotakopoulos, T.; Kotsiantis, S.; Kostopoulos, G.; Iatrellis, O.; Kameas, A. Early Dropout Prediction in MOOCs through Supervised Learning and Hyperparameter Optimization. *Electronics* **2021**, *10*, 1701. [\[CrossRef\]](#)
75. Alamri, A.; Sun, Z.; Cristea, A.I.; Steward, C.; Pereira, F.D. MOOC next week dropout prediction: Weekly assessing time and learning patterns. In *International Conference on Intelligent Tutoring Systems*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 119–130.
76. Xing, W.; Du, D. Dropout Prediction in MOOCs: Using Deep Learning for Personalized Intervention. *J. Educ. Comput. Res.* **2018**, *57*, 547–570. [\[CrossRef\]](#)
77. Liu, K.; Tatinati, S.; Khong, A.W.H. A Weighted Feature Extraction Technique Based on Temporal Accumulation of Learner Behavior Features for Early Prediction of Dropouts. In Proceedings of the 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), Takamatsu, Japan, 8–11 December 2020; pp. 295–302. [\[CrossRef\]](#)
78. Xing, W.; Chen, X.; Stein, J.; Marcinkowski, M. Temporal predication of dropouts in MOOCs: Reaching the low hanging fruit through stacking generalization. *Comput. Hum. Behav.* **2016**, *58*, 119–129. [\[CrossRef\]](#)
79. Taylor, C.; Veeramachaneni, K.; O'Reilly, U.M. Likely to stop? predicting stopout in massive open online courses. *arXiv* **2014**, arXiv:1408.3382.

80. Kloft, M.; Stiehler, F.; Zheng, Z.; Pinkwart, N. Predicting MOOC dropout over weeks using machine learning methods. In Proceedings of the EMNLP 2014 Workshop on Analysis of Large Scale Social Interaction in MOOCs, Doha, Qatar, 25 October 2014; Humboldt University of Berlin: Berlin, Germany, 2014; pp. 60–65.
81. Halawa, S.; Greene, D.; Mitchell, J. Dropout prediction in MOOCs using learner activity features. *Proc. Second. Eur. Mooc Stakehold. Summit* **2014**, *37*, 58–65.
82. Whitehill, J.; Williams, J.; Lopez, G.; Coleman, C.; Reich, J. Beyond prediction: First steps toward automatic intervention in MOOC student stopout; Available at SSRN 2611750; Elsevier: Amsterdam, The Netherlands, 2015.
83. Qiu, L.; Liu, Y.; Liu, Y. An Integrated Framework With Feature Selection for Dropout Prediction in Massive Open Online Courses. *IEEE Access* **2018**, *6*, 71474–71484. [[CrossRef](#)]
84. Ardchir, S.; Ouassit, Y.; Ounacer, S.; Jihal, H.; EL Goumari, M.Y.; Azouazi, M. Improving Prediction of MOOCs Student Dropout Using a Feature Engineering Approach. In *International Conference on Advanced Intelligent Systems for Sustainable Development*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 146–156.
85. Nazif, A.M.; Sedky, A.A.H.; Badawy, O.M. MOOC's Student Results Classification by Comparing PNN and other Classifiers with Features Selection. In Proceedings of the 2020 21st International Arab Conference on Information Technology (ACIT), Giza, Egypt, 28–30 November 2020; pp. 1–9.
86. Jin, C. Dropout prediction model in MOOC based on clickstream data and student sample weight. *Soft Comput.* **2021**, *25*, 8971–8988. [[CrossRef](#)]
87. Mulyani, E.; Hidayah, I.; Fauziati, S. Dropout Prediction Optimization through SMOTE and Ensemble Learning. In Proceedings of the 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 5–6 December 2019; pp. 516–521. [[CrossRef](#)]
88. Xiong, F.; Zou, K.; Liu, Z.; Wang, H. Predicting Learning Status in MOOCs Using LSTM. In *Proceedings of the ACM Turing Celebration Conference—China*; Association for Computing Machinery: New York, NY, USA, 2019; pp. 1–5. [[CrossRef](#)]
89. Wu, N.; Zhang, L.; Gao, Y.; Zhang, M.; Sun, X.; Feng, J. CLMS-Net: Dropout Prediction in MOOCs with Deep Learning. In *Proceedings of the ACM Turing Celebration Conference—China*; Association for Computing Machinery: New York, NY, USA, 2019; pp. 1–6. [[CrossRef](#)]
90. Wen, Y.; Tian, Y.; Wen, B.; Zhou, Q.; Cai, G.; Liu, S. Consideration of the local correlation of learning behaviors to predict dropouts from MOOCs. *Tsinghua Sci. Technol.* **2020**, *25*, 336–347. [[CrossRef](#)]
91. Qiu, L.; Liu, Y.; Hu, Q.; Liu, Y. Student dropout prediction in massive open online courses by convolutional neural networks. *Soft Comput. A Fusion Found. Methodol. Appl.* **2019**, *23*, 10287. [[CrossRef](#)]
92. Zheng, Y.; Gao, Z.; Wang, Y.; Fu, Q. MOOC Dropout Prediction Using FWTS-CNN Model Based on Fused Feature Weighting and Time Series. *IEEE Access* **2020**, *8*, 225324–225335. [[CrossRef](#)]
93. Wang, W.; Yu, H.; Miao, C. Deep Model for Dropout Prediction in MOOCs. In Proceedings of the 2nd International Conference on Crowd Science and Engineering, Beijing, China, 6–9 July 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 26–32. [[CrossRef](#)]
94. Fu, Q.; Gao, Z.; Zhou, J.; Zheng, Y. CLSA: A novel deep learning model for MOOC dropout prediction. *Comput. Electr. Eng.* **2021**, *94*, 107315. [[CrossRef](#)]
95. Yin, S.; Lei, L.; Wang, H.; Chen, W. Power of Attention in MOOC Dropout Prediction. *IEEE Access* **2020**, *8*, 202993–203002. [[CrossRef](#)]
96. Zhang, Y.; Chang, L.; Liu, T. MOOCs Dropout Prediction Based on Hybrid Deep Neural Network. In Proceedings of the 2020 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), Chongqing, China, 29–30 October 2020; pp. 197–203. [[CrossRef](#)]
97. Ashby, A. Monitoring student retention in the Open University: Definition, measurement, interpretation and action. *Open Learn. Open Distance-Learn.* **2004**, *19*, 65–77. [[CrossRef](#)]
98. Ng, K.H.R.; Tatinati, S.; Khong, A.W.H. Grade Prediction From Multi-Valued Click-Stream Traces via Bayesian-Regularized Deep Neural Networks. *IEEE Trans. Signal Process.* **2021**, *69*, 1477–1491. [[CrossRef](#)]
99. Hakimi, L.; Eynon, R.; Murphy, V.A. The Ethics of Using Digital Trace Data in Education: A Thematic Review of the Research Landscape. *Rev. Educ. Res.* **2021**, *91*, 671–717. [[CrossRef](#)]
100. Oboler, A.; Welsh, K.; Cruz, L. The danger of big data: Social media as computational social science. *First Monday* **2012**, *17*, 3993.
101. Eynon, R.; Fry, J.; Schroeder, R. The ethics of online research. *SAGE Handb. Online Res. Methods* **2017**, *2*, 19–37.