Multimodal Technologies and Interaction

MDPI

Review

A Comprehensive Review of Dropout Prediction Methods Based on Multivariate Analysed Features of MOOC Platforms

Saad Alghamdi [1], Ben Soh [1,*] and Alice Li [2]

[1] Department of Computer Science and Information Technology, School of Computing, Engineering and Mathematical Sciences, La Trobe University, Bundoora, VIC 3086, Australia; s.alghamdi@latrobe.edu.au
[2] La Trobe Business School, La Trobe University, Bundoora, VIC 3086, Australia; a.li@latrobe.edu.au
* Correspondence: b.soh@latrobe.edu.au

Abstract: Massive open online courses have revolutionised the learning environment, but their effectiveness is undermined by low completion rates. Traditional dropout prediction models in MOOCs often overlook complex factors like temporal dependencies and context-specific variables. These models are not adaptive enough to manage the dynamic nature of MOOC learning environments, resulting in inaccurate predictions and ineffective interventions. Accordingly, MOOCs dropout prediction models require more sophisticated artificial intelligence models that can address these limitations. Moreover, incorporating feature selection methods and explainable AI techniques can enhance the interpretability of these models, making them more actionable for educators and course designers. This paper provides a comprehensive review of various MOOCs dropout prediction methodologies, focusing on their strategies and research gaps. It highlights the growing MOOC environment and the potential for technology-driven gains in outcome accuracy. This review also discusses the use of advanced models based on machine learning, deep learning, and meta-heuristics approaches to improve course completion rates, optimise learning outcomes, and provide personalised educational experiences.

Keywords: MOOCs; dropout prediction; meta-heuristics; deep learning; accuracy optimisation; feature selection

check for updates

Academic Editor: Julius Nganji

Received: 20 November 2024
Revised: 2 January 2025
Accepted: 3 January 2025
Published: 7 January 2025

Citation: Alghamdi, S.; Soh, B.; Li, A. A Comprehensive Review of Dropout Prediction Methods Based on Multivariate Analysed Features of MOOC Platforms. *Multimodal Technol. Interact.* **2025**, *9*, 3. https://doi.org/10.3390/mti9010003

1. Introduction

Massive open online courses (MOOCs) have brought about a new era of learning by providing easier access to high-quality educational materials worldwide [1]. This approach aims to address the challenges of accommodating large numbers of students, especially in developing countries, where traditional infrastructure is often inadequate. However, despite the initial high levels of engagement with MOOCs, many participants unfortunately drop out as the courses progress [2]. The dropout rate is a crucial metric for evaluating the effectiveness of cutting-edge approaches and frameworks designed to improve MOOC adoption and performance. A thorough understanding of the factors driving continuous MOOC acceptance and enhancing learning outcomes is essential for improving course completion rates. Continued use of MOOCs reflects a cognitive state involving a learner's personal attitude and willingness to adopt new technologies in education [3].

MOOC dropout prediction enhances the design of MOOCs and plays a crucial role in improving the user experience. The rise of machine learning (ML) and deep learning (DL) has introduced powerful methodologies for evaluating large datasets and extracting valuable insights [3]. These techniques can be applied to massive open online courses to build predictive models that not only forecast student behaviour but also increase the accuracy of learning outcome predictions. Given the complexity and diversity of learner interactions in MOOCs, the development of intelligent models to predict dropout rates is essential [4]. These models must capture complex patterns in learner data, adapt to varying learning contexts, and provide actionable insights for course designers and instructors to address the challenges associated with MOOC dropout prediction, as shown in Figure 1. In this regard, the selection of relevant MOOC features and feature selection methods are key factors that significantly impact the performance of dropout prediction models.

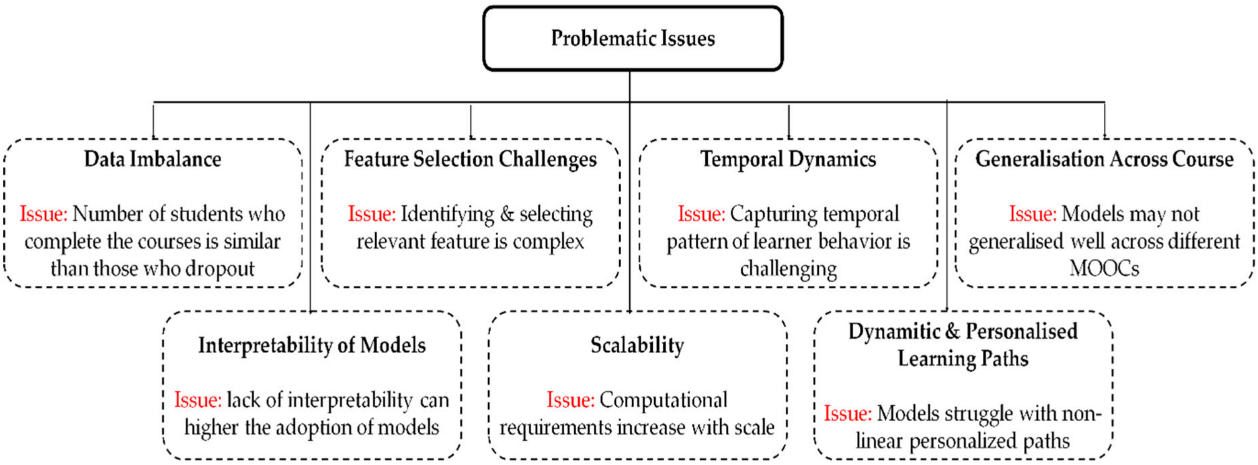

Figure 1. General problematic issues in MOOCs dropout prediction model.

MOOCs' feature types and selection methods are crucial components that significantly influence the performance of dropout prediction models [4]. The vast and diverse data generated by MOOC platforms include various features such as learner demographics, engagement metrics, interaction patterns, and course-specific variables. Identifying and selecting the most relevant features is essential for building accurate and robust predictive models. An effective feature selection process ensures that the model focuses on the most impactful variables, reducing noise and improving the ability of the model to generalise across different learning contexts [5]. Consequently, a well-considered approach to feature selection can lead to more precise dropout predictions and, ultimately, more effective interventions to enhance learner retention and success. There are specific features that must be carefully addressed in building MOOCs, which are reviewed according to the following objectives:

- Identify and select the most relevant learner attributes to ensure that the model accurately reflects diverse learning behaviours.
- Adapt to different learning contexts in MOOCs which may impact learner retention and dropout rates.
- Deal with the vast amount of data generated by MOOCs, which require robust attribute selection methods to filter out noise and focus on the most influential variables.
- Ensure that the model can generalise across different MOOC platforms and learner populations, improving its predictive accuracy and effectiveness.

1.1. Motivations

The purpose of this comprehensive review is to critically explore and evaluate the current literature on predicting dropout in MOOCs and to identify the key factors that contribute to learner dropout and the most effective AI-based approaches to address these challenges. By summarising findings from different studies, this review aims to provide a comprehensive understanding of the current state of research in this area and highlight the need for more sophisticated and intelligent models that can effectively mitigate dropout rates and improve learner outcomes in MOOCs. The motivations of this paper are listed as followings:

- Provide a compilation of MOOCs dropout prediction methods and their significance.
- Classify the potential MOOCs features, selection methods, and architecture.
- Identify and describe the key requirements to improve and optimise the MOOCs dropout prediction.
- Propose a thematic taxonomy of literature according to the most important dropout prediction features and provide a comprehensive review on recent advances MOOCs dropout models.
- Discuss the optimisation challenges related to MOOCs dropout prediction models, highlighting the most related methods and schemes.
- Illustrate the research gaps on MOOCs dropout prediction optimisation-based meta-heuristics and DL approaches, in addition to future works.

1.2. Contributions

This comprehensive review begins with a review of the methodologies used in many studies on MOOCs, with a particular focus on the interaction between their continued acceptance and the accuracy of predicting learning outcomes, through a review and critical analysis of current research on the dynamics of acceptance of MOOCs, with an emphasis on the role of the technology acceptance model. Moreover, this review will investigate the integration of predictive modelling, specifically ML, DL, and meta-heuristics, in improving the accuracy of outcome predictions in the MOOC environment. By deeply focusing into the complexities of these intersecting methods, this review aims to shed light on the mechanisms that underpin MOOC acceptance as well as the possibility for technology-driven advances in outcome accuracy.

This review offers readers a structured overview of the current body of research, providing a comprehensive synthesis of studies on MOOC acceptance and the application of predictive models. Furthermore, this study will review in depth and explore the challenges related to MOOCs optimisation, and will also provide insights into an integrated strategy that has the potential to shape the future of online learning approaches. In this study, numerous articles related to dropout prediction methods have been reviewed and analysed. The selected papers were evaluated based on criteria such as their titles, abstracts, contributions, resource quality, and journal impact factor. The review methodology is structured around the following paper selection and investigation criteria:

- Provide a taxonomical review for different MOOC dropout prediction methods by classifying the proposed studies according to methodologies, important impacts, contributions, and study directions.
- Review dropout prediction methodologies to efficiently increase the student completion rate and optimise the accuracy.
- Provide the main research direction questions related to dropout prediction optimisation methodologies.
- Highlight the weaknesses and expected improvements in the future with respect to the efficient optimisation in MOOCs' completion rate and accuracy.
- Present various approaches in tabular format and charts to help researchers clearly understand the key techniques and methodologies used in MOOC dropout prediction.

The rest of this paper is organised as follows: Section 2 provides an overview of the process and highlights the contributions of our comprehensive review. Section 3 reviews the existing literature on MOOC dropout prediction, highlighting important ideas like feature selection, architectural models, and the difficulties faced in MOOC optimisation. The many dropout prediction methods analysed are presented in detail in Section 4, along with comprehensive evaluations of their uses and limitations. We discuss the remaining issues and research gaps in MOOC dropout prediction and optimisation in Sections 5 and 6, respectively. In Section 7, future prospects in this field are reviewed and the limitations of existing research are reflected upon. This paper is finally concluded in Section 8, which provides a summary of our results and recommendations for further research.

2. The Review Methodology

MOOCs are regarded as a rapidly evolving educational technology that needs substantial attention and additional study on a number of topics, particularly dropout prediction across multiple application scenarios [5]. The integration of online learning platforms with advanced data analytics has led to significant technological advancements by combining various types of data from smaller and smarter sources, allowing for flexible interaction and operation across diverse educational environments. The MOOCs framework enables the use of cloud-based systems and ML techniques to provide real-time monitoring of student engagement and learning behaviours [6]. In addition, through these capabilities, MOOCs can effectively manage and operate various learning resources, and student retention can be significantly impacted by factors such as course design and learner motivation. According to this viewpoint, additional knowledge and approaches to enhance dropout prediction in MOOCs are needed. These approaches and insights may provide new technologies or techniques that guarantee stable and dependable MOOC deployments while enhancing learner results.

Throughout this article, readers will encounter many images, charts, and diagrams used as methodologies to analyse and evaluate past studies and techniques targeted at improving dropout prediction in MOOCs [6]. These review procedures serve to understand the benefits and drawbacks of the proposed prediction models and approaches provided in the field of MOOC dropout reduction. More precisely, this technique offers a thorough comprehension of the conditions pertaining to the existing MOOC dropout prediction systems, acting as a model for an all-encompassing methodology for examining and assessing different experimental outcomes in this field [7]. Future approaches and models to enhance learner retention and performance in MOOCs can be developed using the innovations and related difficulties in dropout prediction that have been presented as a roadmap.

A number of important questions were established to guide the research process and ease the discussion of material received from various investigations. The uniqueness of the models, the guidelines established for data gathering and analysis, and guaranteeing the caliber of the outcomes are the main topics of these inquiries [7]. The methodology of this study, which is shown in Figure 2, helped create an organised framework for evaluating pertinent studies on MOOC dropout prediction.

The sources of different publications were analysed as part of the study methodology that was chosen by looking for titles, abstracts, and keywords that were pertinent to enhancing dropout prediction in MOOCs. Many other studies were disregarded because they only addressed generic e-learning applications without addressing the unique difficulties associated with MOOCs. Consequently, the following methodological steps were followed in order to study and assess the chosen papers thoroughly.

2.1. Sources of Literature

The peer-reviewed research publications and conference proceedings were gathered from university libraries and internet research platforms including Research Gate and Google Scholar. Along with other sources like Tylor & Francis, ACM, Nature, Frontiers, and conferences, the majority of the articles gathered are from famous indexed journals like IEEE, MDPI, Springer, Elsevier, and Wiley. The papers that have been compiled have been published from September 2024 to the year 2018. The number of chosen articles assessed for dropout prediction in MOOCs is indicated in Figure 3, along with a summary of the publication sources.

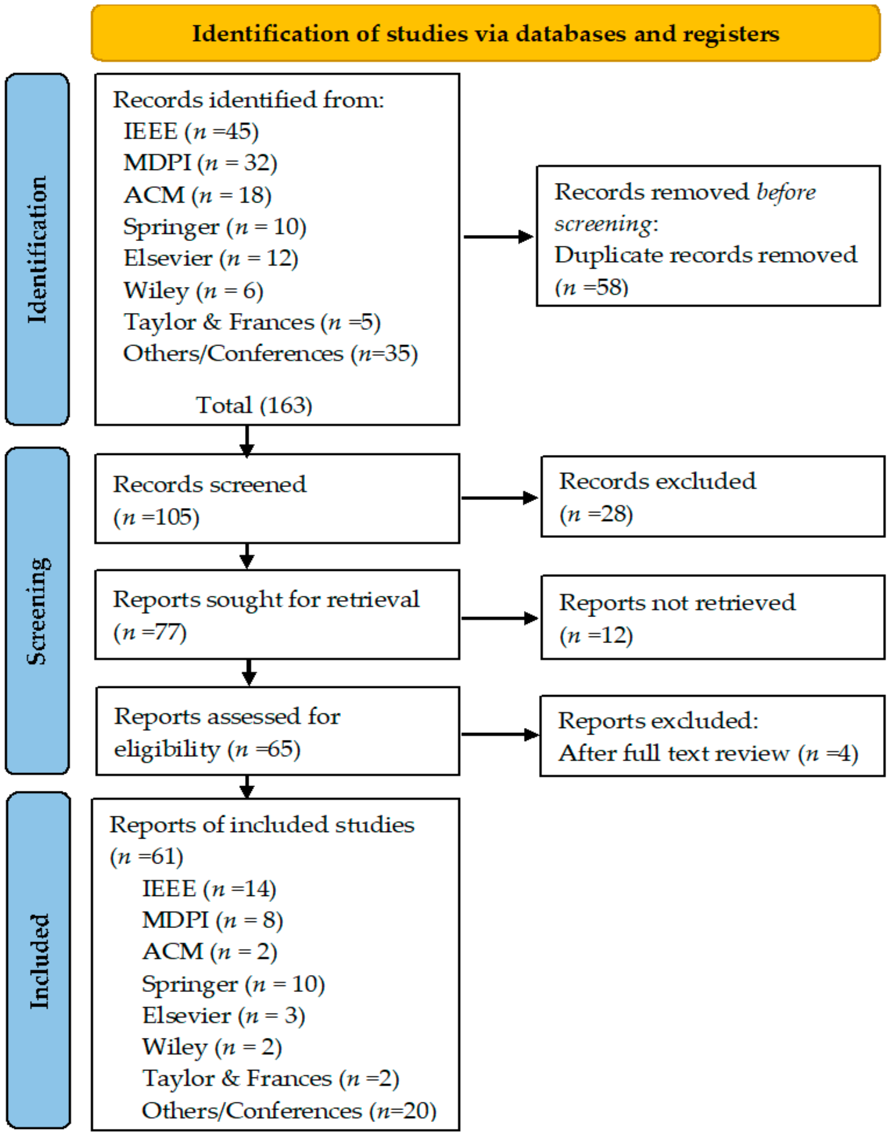

Figure 2. The methodology of studies selection criteria.

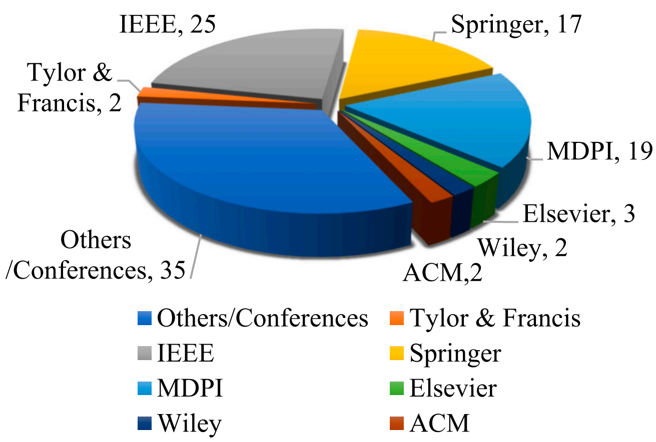

Figure 3. The number of total selected papers for review study (105 articles).

2.2. Key Questions and Inclusion Criteria

A set of guiding questions was formulated to review the papers related to this study, aiming to clarify various parameters used in different dropout prediction techniques for MOOCs. These questions are key to the presented study and were answered through the analysis of several discussed papers.

- Q1. What are the parameters that can impact MOOC dropout rates and learner engagement in general?
- Q2. What are the main factors that influence student retention in MOOCs?
- Q3. What technologies, methods, and algorithms can be used to improve dropout prediction in MOOCs?
- Q4. To what extent can recent methodologies achieve optimisation in dropout prediction accuracy?

Based on the guiding questions and the study selection process in Figure 2, a number of review and survey articles were selected following a specific strategy. First, search keywords or phrases were defined to find relevant articles, including MOOCs dropout prediction, MOOCs prediction optimisation, MOOCs feature selection, ML for dropout prediction, DL dropout prediction model, feature selection optimisation, meta-heuristic for feature selection, student retention in MOOCs, student engagement, MOOC completion rate analysis, predictive modeling for MOOCs, and intelligent dropout prediction models. Throughout the collection processed based on the keywords, Boolean operators such as AND and OR were used to combine the keywords during the search process. Next, duplicate articles were removed, along with irrelevant articles by checking the title, abstract, and full text. Finally, the articles most relevant to the proposed study were selected. This strategy allowed us to accurately present the literature related to this study and to identify other papers that provide methods and approaches for improving dropout prediction in MOOCs. The search of related articles collected from different resources yielded a total 163 records from IEEE (n = 45), MDPI (n = 32), ACM (n = 18), Springer (n = 10), Elsevier (n = 12), Wiley (n = 6), and Taylor & Francis (n = 5), as well as others from other databases and conferences (n = 35). The removed duplicated are (n = 32). A total of 105 articles screened as related to dropout prediction approaches in MOOCs were reviewed, as summarised in Figure 3. After reviewing the articles, we found that only 61 specifically addressed dropout prediction methods and solutions for MOOCs. These articles, summarised in Figure 4, were analysed, leading to the following outcomes:

- The concept of dropout prediction has been explored in the reviewed studies and analysed across various contexts, focusing on reducing dropout rates, enhancing student engagement, and optimising learning outcomes.
- Numerous studies have investigated the use of ML/DL as intelligent methods to improve dropout prediction in MOOCs. However, these studies often focus on select influencing factors and fail to encompass all the considerations required for a comprehensive approach to dropout prediction.
- Many papers were excluded as they focused on methods for improving student retention in general e-learning platforms, without specifically addressing dropout prediction in MOOCs.
- This study is expected to provide researchers with comprehensive coverage of the strengths and weaknesses of various relevant studies conducted between 2018 and 2022.

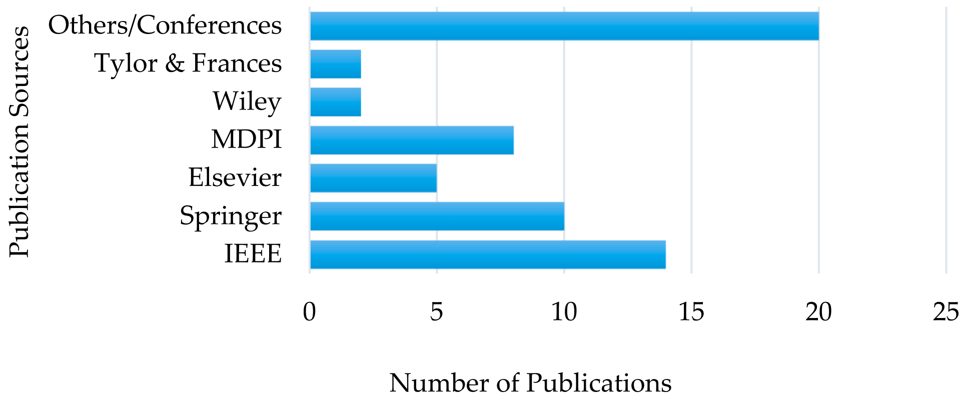

Figure 4. Summary of overall articles related to MOOCs dropout prediction corresponding to the publication sources (61 papers).

3. MOOCs Dropout Prediction Technical Background

MOOCs have received great attention in recent years from all educational levels, but they face major challenges related to a high dropout rate (reaching 92%), which reflects the extent of interaction with these platforms [8]. Various processes are employed to predict student activity on MOOC platforms and analyse their interactions across different courses. This helps in building new platforms or improving existing ones to enhance performance, leading to greater user acceptance and significantly reducing failure rates. Student interactions on a MOOC platform are influenced by the structure and functions of the platform, both of which can impact course completion rates and contribute to dropout levels [9]. The functions of a MOOC platform include managing large volumes of student data, handling independent course data, designing course videos to withstand high traffic, tracking online test scores and title settings, and managing course resources. Accordingly, MOOCs typically consist of four fundamental modules, namely, the user management module, the course outline module, the multimedia resource module, and the test management module, as shown in Figure 5.

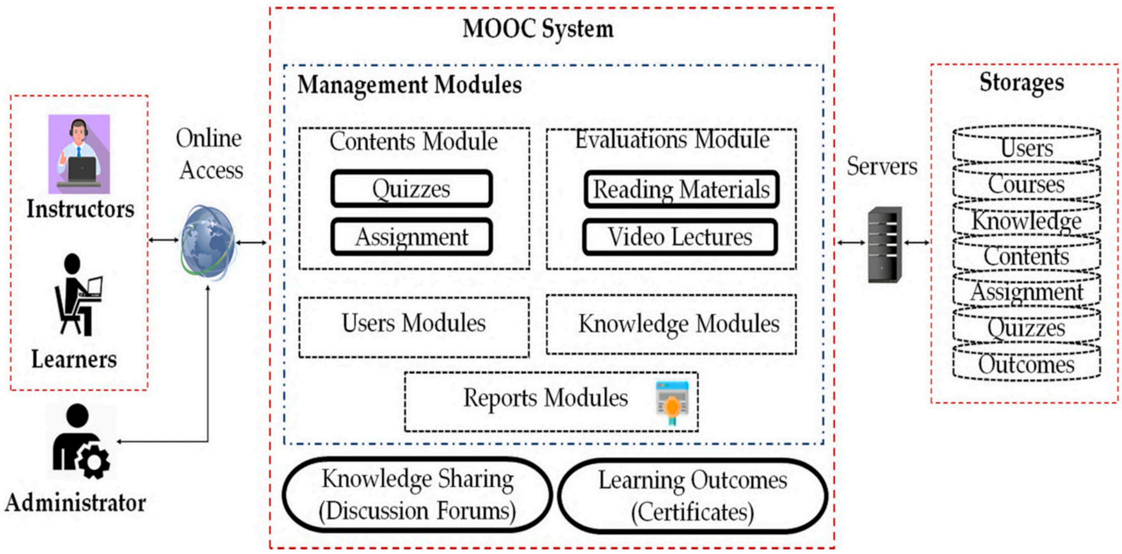

Figure 5. The MOOC structure.

One of the most prominent challenges of MOOCs, after design, is how to ensure acceptance by students and attract high retention rates. Low retention rates are often reflected in the course completion rates or dropout percentages [9]. According to the literature, MOOCs offered by elite universities, such as MIT, Stanford, and UC Berkeley, have reported dropout rates as high as 95%. These dropouts can be attributed to various factors related to both the students and the MOOC platforms themselves, as outlined in the following list.

A. Student factors

- Lack of student motivation, often cited as the most crucial factor contributing to high dropout rates, is influenced by personal and professional identity development, expectations of future economic benefits, the desire for challenge and achievement, and entertainment [9].
- Students may be reluctant to invest the considerable time required to watch videos, complete quizzes, and finish assignments.
- Students often lack the essential background knowledge and skills needed to comprehend the course content.
- Some researchers have noted that many enrolments in MOOCs are casual, i.e., from students who have no real intention to complete them.

B. MOOC factors

- Course design: A weak and non-interactive course design, including course content, structure, and information delivery technology, contributes significantly to high dropout rates.
- Hidden cost: The hidden cost could be another cause of a high dropout rate. Despite MOOCs' reputation as a free resource for online education, students need to pay for their certificates or sometimes buy costly textbooks recommended by lecturers [10].

According to the given factors, the most impactful factors are those related to the MOOCs' design. The way platforms are designed greatly affects the extent of user interaction, regardless of personal factors related to the user. These factors strongly influence the dropout rate, highlighting the need for a strategic approach to designing MOOCs and to assessing dropout predictions continuously for further improvement [10]. To reduce the dropout rate, the quality of instructional design in MOOCs is a crucial determinant and prerequisite for promoting active learning and enhancing learner experiences. Interaction among learners significantly contributes to lower dropout rates. In addition, integrating social tools in MOOCs can effectively assist peers in addressing challenges, encouraging discussions, and facilitating the sharing of new resources [11]. Further, the comfort level with technology for learners must be acceptable. Since MOOCs are online courses, they require technology to deliver content and communicate with students. Evaluation is important and is a crucial step in any educational process. The evaluation period after completing the course must not exceed a period of more than 48 h, and this is important to increasing the learners' connection to the platform [12]. The following subsections outline key details regarding the process of accurately predicting dropout rates in MOOCs, focusing on learning databases, feature selection, dropout prediction methodologies, and optimisation challenges.

3.1. MOOCs Database

The MOOC dataset, which collects extensive data on students' behaviour, engagement, and performance during the course, is essential to dropout prediction. These data contain a variety of information, including students demographics, learning interactions, such as video views and quiz completions, and the amount of time spent on various course activities [12]. In addition, the dataset shows patterns of involvement and disengagement that are important to know in order to identify children who might be at danger of not finishing school. Through the analysis of these patterns, educators and platform developers can acquire valuable insights into the causes that lead to dropout rates, which will facilitate the implementation of focused interventions [13].

The prediction models' performance is directly impacted by the quality of the dataset due to the complexity of dropout causes, which can be attributed to both MOOC-related issues like poorly designed courses and personal factors such a lack of enthusiasm or time [13]. A student's lack of motivation, for instance, could be reflected in the dataset by their decreased interest in the course materials, or their inconsistent participation in quizzes and assignments could suggest that they are pressed for time. It is also possible to deduce information about knowledge gaps from low exam results or little conversation participation. In order to create efficient algorithms that can anticipate dropouts early and give students appropriate support, a rich and well-labelled MOOC dataset is necessary.

The MOOC database structure is organised into three main content areas: course details, specialisation details, and student engagement. These domains are reinforced by crucial standalone blocks for educational institutions, instructors, and learners [14]. The structure of the platform is made up of these elements, which make it possible to handle courses, student data, and interaction effectively. A key component of the platform is the course details block, which arranges details on the materials, structure, and content of the course. Important details like course titles, descriptions, and the connections between courses and specialisations are captured by the model. Every course consists of several parts, each containing related assignments, tests, and videos as learning resources [15]. There is a clear hierarchy of content distribution thanks to the links between these assets and particular course sections.

In the specialisation details section, the model manages how courses are arranged into specialisations, each of which consists of several courses and culminating projects. In addition to facilitating the management of discounts and course bundles, this structure offers advanced learning pathways. In addition to enrolment information, session schedules, and performance goals, student participation tracks how students engage with courses and specialisations [15]. Tracking completions, keeping focus on student progress, and recognising dropout concerns all depend on this area. As essential data repositories, stand-alone blocks for educational institutions, instructors, and students offer the necessary cross-model references [16]. These tables ensure accurate associations between courses and their designers, as well as between students and their individual participation records. This database structure efficiently supports the fundamental features of a MOOC platform, allowing for the scalable and effective administration of large-scale virtual learning environments and opening up opportunities for future improvements to the capabilities and user experience of the platform.

3.2. MOOCs Feature Selection

As it entails determining which characteristics from the dataset are most pertinent to making accurate predictions, feature selection is a crucial step in the MOOC dropout prediction process. The effectiveness and efficiency of a prediction model can be enhanced by concentrating on its most noteworthy aspects as not all data points hold equal significance [17]. The model can better identify trends that signal the possibility of dropout by choosing the appropriate parameters, such as quiz scores, time spent on course materials, and student engagement indicators. With less complexity, the model can identify at-risk students more accurately and with less computational overhead due to this improved method.

In addition, feature selection improves the dropout prediction model's interpretability [17]. By focussing on the most important variables, instructors and platform developers can gain a deeper understanding of the reasons for student dropouts and the particular course components or student behaviours that are causing them. Certain aspects, like interaction with course videos or participation in discussion forums, may indicate a higher level of engagement, but low quiz scores may indicate a lack of expertise [18]. A more precise understanding of the reasons for dropouts and the ability to design retention-focused interventions are made possible by feature selection, which is effective in preventing the model from becoming overloaded with irrelevant data.

Figure 6 illustrates the many feature selection techniques that can be applied. The types of feature selection methods used in MOOCs have a direct impact on the effectiveness and performance of the dropout prediction models. Every selection technique, such as filtering, wrapping, and hybrid techniques, has unique properties with regard to optimising MOOC dropout prediction [18]. Statistical measures can be quickly and easily ranked and chosen with the help of filters. However, because they do not take into account the relationships between features, they may be less accurate. While embedded approaches incorporate feature selection into the learning process, wrappers allow for the evaluation of feature subsets depending on model performance and offer improved accuracy [19]. The strengths of the filter, wrapper, and embedded techniques can be used with the hybrid approach to balance prediction accuracy.

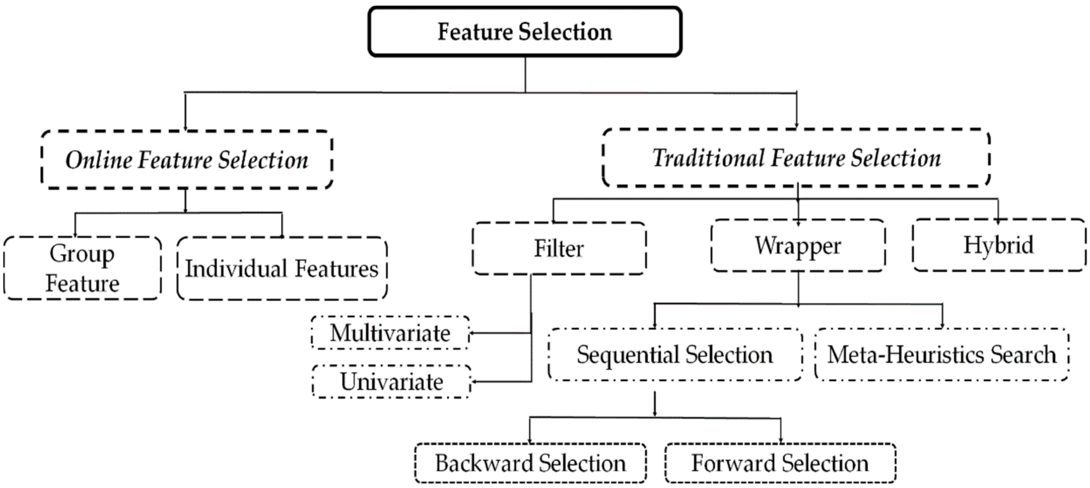

Figure 6. Feature selection approaches.

When the accuracy of wrapper techniques is combined with the filter strategy, the hybrid method proves very beneficial for dropout prediction in MOOCs. With the use of these combinations, it is possible to filter out features that are not important and then refine the features that remain, using strategies based on wrappers and model performance [19,20]. Optimising the dropout prediction process is possible with a hybrid strategy that uses a wrapper based on meta-heuristic algorithms. To identify the subset that optimises the prediction accuracy of the model, the meta-heuristic algorithms cleverly investigate various feature combinations. This can guarantee both efficiency and high accuracy, particularly for large-scale datasets. Dropout prediction is a critical issue that needs to be resolved in order to increase MOOCs' scalability, user engagement, personalisation, predictive accuracy, and high student retention rates [20]. Table 1 shows how these requirements relate to different feature selection factors based on MOOC applications.

Table 1 also shows that every component has an impact on every improvement need. Greater data quantities typically result in increased accuracy and engagement, but they are also needed for sufficient scalability and flexibility. Useful complexity encourages good scalability, prediction accuracy, and adaptability, but because it may make interpretation and use more challenging, it may have a mediocre effect on student engagement [20]. The flexibility of the algorithm offers low influence on student engagement, medium prediction accuracy, moderate scalability, and support for customisation and adaptability. This indicates that while adaptable algorithms may adapt to different data sources, engagement may not always be directly impacted by them. Though it has a small impact on adaptability, higher feature relevance across all criteria ensures that characteristics employed are essential for enhancing personalisation, scalability, accuracy, and engagement [21]. While real-time processing is necessary for fast modifications, it is not always suitable for handling extensive personalisation and adaptability. Finally, while varied data may not always immediately increase engagement, they have a significant impact on personalisation and scalability, which enhances model performance and adaptability.

Table 1. Feature selection factors and their impact on effectiveness and performance requirements.

Factors	Personalisation	Scalability	Predictive Accuracy	Learner Engagement	Adaptability
Data Volume	Moderate	High	High	High	Moderate
Model Complexity	High	High	High	Medium	High
Algorithm Flexibility	Moderate	Moderate	Medium	Low	High
Feature Relevance	High	High	High	High	Moderate
Real-Time Processing	Low	Medium	Medium	Medium	Low
Data Diversity	High	High	Medium	Medium	High

3.3. MOOCs Dropout Prediction Methodology

Many MOOC courses that facilitate the exchange of educational content over the internet are built on data-driven platforms and learning management systems. Along with other applications pertaining to personalised learning and evaluation, these programs are used to track and manage student progress, lower dropout rates, and improve learning outcomes [22]. Adaptive learning systems in educational settings use data-driven insights to modify support systems and course materials, enabling self-paced learning, better student retention through thoughtful feedback mechanisms, and a lower dropout rate.

Figure 7 illustrates the general structure for the dropout prediction approach, which consists of three key stages: dropout prediction model, feature selection and analysis, and dataset pre-processing. The output of the framework, which is dependent on the input, i.e., student learning behaviour data, is the number of dropouts [23]. The attributes of each form of learning record are derived based on the ranking approach after the students' learning records are grouped in the framework depending on the type of learning behaviour. Feature engineering and data pre-processing are essential components of a high-precision stall prediction model.

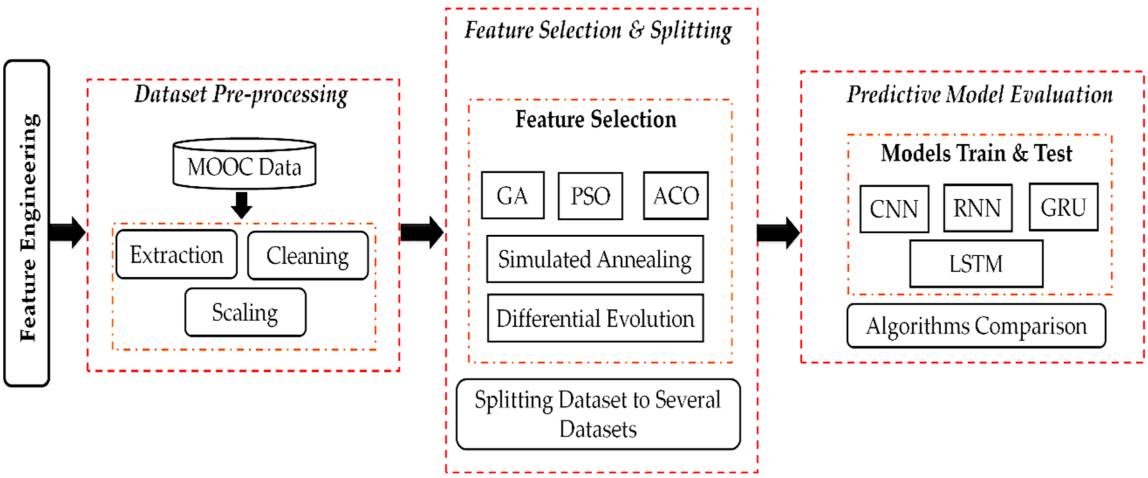

Figure 7. General model for MOOC dropout prediction.

A variety of techniques, including decision trees (DT), random forests (RF), support vector machines (SVM), and even deep learning (DL) models like recurrent neural networks (RNN), may be used in the dropout prediction stages [23,24]. Each student receives a categorisation or risk score from the dropout prediction framework, which indicates which students are most likely to discontinue their studies. For the purpose of assessing MOOC dropout risk, a carefully designed dropout prediction framework comprising feature selection, data pre-processing, and a reliable prediction model is essential. By increasing model correctness and efficiency, the process is further improved by the use of hybrid filtering and wrapper techniques. MOOC platforms are able to better assist students by anticipating dropouts with accuracy, which eventually increases student retention and course completion rates [24].

3.4. Dropout Prediction Optimisation Challenges and Solutions

Predicting dropout in MOOC is generally associated with a number of issues, as seen in Figure 8. The most notable of these is the insufficient sample size, which not only correlates with dropouts but also with graduates in order to produce findings for the classification process that are free of bias. The majority of machine learning algorithms rely on the abundance of positive and negative data [25]. However, a lot of people sign up for different MOOCs merely to try out online self-learning, sometimes even with the goal of obtaining a certificate, and they frequently discontinue the course before the end of the allotted time. Due to the difficulties of confirming the accuracy of the available data, missing data may also result from MOOC big data and unstructured data, which are the most significant prediction problems for ML [25,26]. Average values can occasionally be used to make up for missing data, but they are never a practical substitute for precise feature determination.

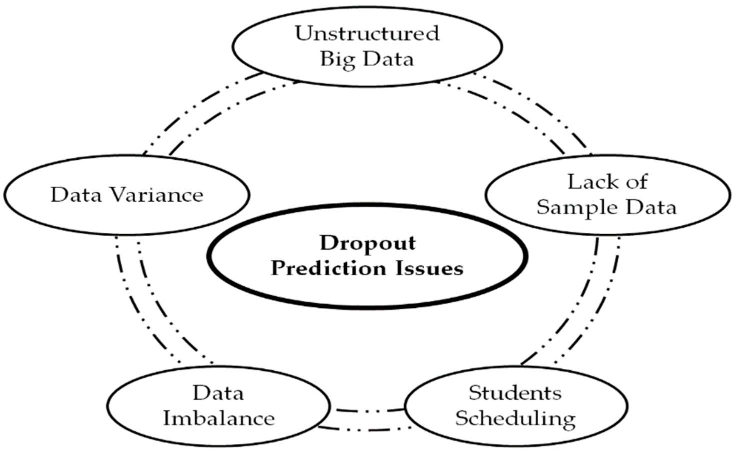

Figure 8. Main issues related to MOOCs dropout prediction.

Data variance and an imbalance in MOOCs are additional factors to take into account. Data variation can have a detrimental effect on the reliability of machine learning models because of self-paced learning. Unbalanced data negatively impact algorithms such as SVM. causing them to produce biased predictions that favour the majority class. This decreases the accuracy and dependability of the model, particularly when it comes to dropout prediction [26]. In addition, there is the problem of meeting the varied scheduling preferences of students. While some prefer to move through courses week by week, others prefer to study at their own speed or have access to all materials up front. While satisfying these diverse preferences is difficult, AI optimisation techniques can help. AI has a place in scheduling; it can aid with time constraints for students and enable more individualised learning.

3.4.1. Feature Availability

The quality and diversity of features have a direct impact on the efficacy and accuracy of machine learning models; hence, feature availability is crucial to the optimisation of dropout prediction models in MOOCs. An enormous quantity of data are produced by students' learning behaviours in MOOCs. Optimisation of dropout prediction models requires careful selection, engineering, and assurance of crucial feature availability [27]. Missing features can significantly impair model performance; therefore, having rich, comprehensive data is essential. Sparse or inconsistent datasets are produced when some characteristics of MOOCs are frequently inaccessible as a result of insufficient student activity data. Unreliable predictions can result from missing features, particularly in situations where behavioural patterns that are not recorded because data gaps cause dropouts [28]. Missing features also introduce noise.

The interpretability of the dropout prediction model is also impacted by feature availability. For educational institutions seeking to understand why specific students are in danger of dropping out, the model can yield clearer and interpretable findings when key attributes are regularly available [28]. The model may rely on less pertinent data if characteristics are absent or inadequately represented, which would reduce its interpretability and make it more difficult to extract useful insights from the forecasts. The efficacy of dropout prediction models in MOOCs is largely dependent on feature availability. Although a large feature set might improve the resilience and accuracy of the model, it can also present difficulties that affect optimisation [29,30]. Model performance can be negatively impacted by elements like noisy data, excessive dimensionality, irrelevant or duplicated features, and insufficient feature sets. It is imperative to comprehend these possible adverse effects in order to create efficient dropout prediction systems. Table 2 bellow lists some feature availability-related aspects and their potential detrimental effects on optimisation [30]. This table sheds light on how these problems may impair the general effectiveness and dependability of dropout prediction models.

Table 2. Feature selection factors and their negative effects.

Factors	Negative Impact Description
Missing Certain Features	Reduces model dependability by causing erroneous predictions based on missing data.
Variance	Accuracy may suffer from models' inability to generalise due to high variability.
Irrelevant Features	Low-impact or irrelevant elements can weaken the model's focus and cause it to perform lower.
Noisy Data	Features with noisy data might cause errors and lower prediction optimisation's dependability.
Feature Redundancy	Redundant features can make a model less efficient by increasing computational cost and causing overfitting.
Imbalance	Imbalanced data can still cause models to be biased towards the majority class, leading to poor performance on minority classes.
Overloaded Feature Set	A model that is overloaded with characteristics may find it difficult to determine which predictions are the most important and may require longer training times.

For dropout prediction models in MOOCs to be optimised, feature availability is essential. By giving a more accurate picture of student behaviour, comprehensive and pertinent features like temporal data and student engagement indicators greatly improve model accuracy. These features, which capture a wide variety of behaviours, alleviate data imbalance, facilitate effective feature engineering, and enhance model robustness [31]. They also help with the investigation of temporal engagement patterns, improve the interpretability of the model, and enable more precise feature selection. Having a wide range of high-quality characteristics makes it easier to handle missing data, make more accurate predictions, and ultimately improve learning outcomes and student retention.

3.4.2. Thematic Taxonomy of Optimisation Solutions

Numerous earlier studies that addressed how to improve dropout prediction in MOOCs have presented a range of techniques to increase predictability and lower dropout rates. These strategies include hybrid approaches, data-driven insights, adaptive learning techniques, and sophisticated algorithms [31]. As illustrated in Figure 9, which rates dropout prediction technologies and solutions, these techniques can be roughly divided into multiple groups.

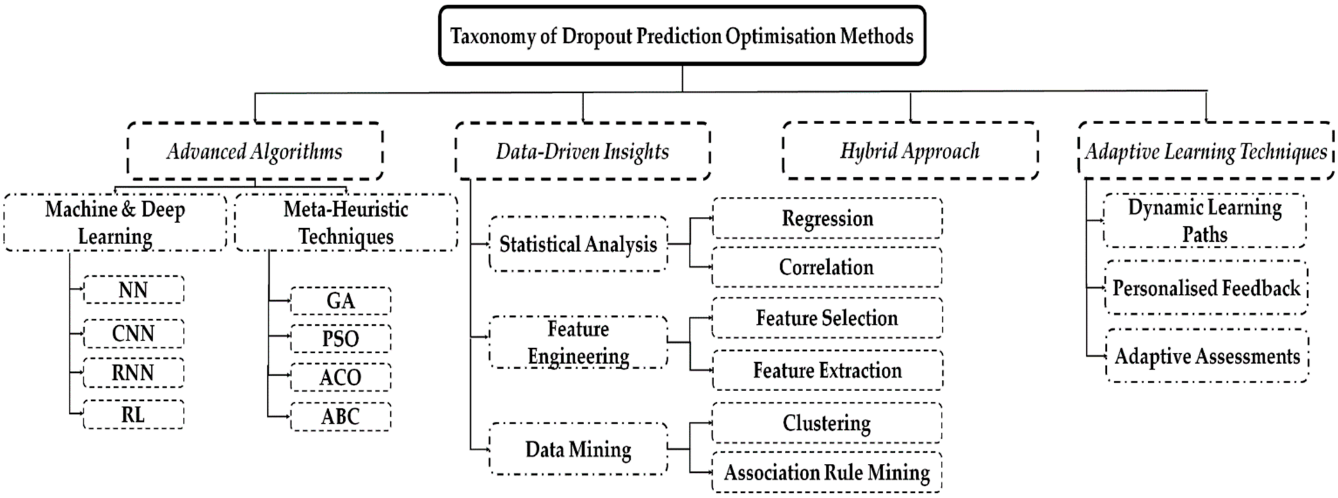

Figure 9. Thematic taxonomy of dropout prediction optimisation solutions.

Optimising dropout prediction in MOOCs can benefit greatly from the hybrid strategy that combines sophisticated algorithms with feature selection-based meta-heuristic methods. Meta-heuristic techniques, such particle swarm optimising (PSO), genetic algorithm (GA), and ant colony optimisation (ACO), are excellent at sifting through intricate feature spaces to find the most pertinent characteristics, improving model efficiency and predictive accuracy [32]. Feature selection is made easier, dimensionality is reduced, and model robustness is increased when these methods are combined with sophisticated ML algorithms like SVM, RF, NN, or deep learning techniques like CNN or RNN. By improving the feature set and better capturing dropout risks across several classes, this combination also solves data imbalance.

The hybrid approach also makes it easier to scale and adapt, which makes it appropriate for a variety of MOOC scenarios and expanding datasets. By concentrating on a pertinent subset of features, it improves interpretability and yields useful insights. In addition, it ensures effective model performance by optimising computational resources by lowering the number of features needed for training [32]. In general, meta-heuristic feature selection and sophisticated algorithms work in concert to maximise each other's advantages, producing dropout prediction models that are more precise, reliable, and scalable. With an emphasis on MOOC platforms, this review will examine numerous papers that have used hybrid methodologies along with ML/DL/Meta-heuristics applied to educational datasets. The discussion will focus on how these methods have been applied to forecast various student outcomes, such as learning performance, course completion rates, and dropout rates [33].

The focus will be on comparing methods, including feature selection techniques, model types used, and datasets used. The review will also address the current gaps in the literature, namely, the scant investigation of hybrid models that combine ML/DL and meta-heuristics in the framework of educational big data. The basis for comprehending how predictive models might be enhanced to enhance MOOC results and lower dropout rates will be provided by this analysis.

4. MOOCs Dropout Prediction Approaches

With the goal of increasing prediction accuracy and lowering dropout rates, a wide range of methods have been presented to improve MOOC dropout prediction. In order to retain high prediction performance while controlling computational resources, these techniques include tactics for modifying learning experiences, improving data analysis, and optimising algorithms [33]. In order to increase their efficacy and accuracy, dropout prediction models also need to make optimal use of data and computational resources. Therefore, preserving high prediction accuracy and enhancing overall MOOC efficacy require good resource utilisation. Enhancing retention rates and dropout prediction accuracy in MOOCs has been the subject of recent research. These studies, however, frequently ignore important elements unique to MOOC applications. A number of studies have recently examined and analysed dropout prediction techniques.

The current review papers on dropout prediction in MOOCs are summarised in Table 3. Although these surveys provide an overview of fundamental MOOC designs, models, and some treatments, they do not fully explore the particular needs for dropout prediction, current research directions, or technological developments.

Table 3. Comparison of proposed review papers with existing surveys on dropout prediction in MOOCs.

Contributions	Review Papers										Proposed Study
	[34]	[35]	[36]	[37]	[38]	[39]	[40]	[41]	[42]	[43]	
MOOC Architecture	✓	×	×	×	×	×	×	×	×	×	✓
Dropout Prediction	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓
Feature Selection	✓	×	×	×	×	✓	✓	×	✓	✓	✓
Learning Models	✓	✓	✓	✓	✓	✓	×	×	×	×	✓
Predictive Analytics	✓	×	✓	×	✓	×	×	✓	✓	×	✓
Scalability Challenges	×	×	×	×	×	✓	×	✓	✓	×	✓
Open research Issues	✓	×	✓	✓	×	✓	×	✓	×	✓	✓
Prediction Optimisation	×	×	×	×	×	×	×	×	×	×	✓

While some studies have addressed dropout prediction in MOOCs, the comparison in Table 3 demonstrates that many have not sufficiently addressed important criteria associated with optimisation prediction accuracy and retention rates. The subsection that follows will cover a variety of approaches and strategies in the field of dropout prediction, offering a thorough analysis of prediction methodologies, cutting-edge research, and answers to major problems. A number of strategies are investigated to improve dropout prediction in MOOCs, including intelligent algorithms, feature selection techniques, and optimisation models. The following subsections will provide an in-depth analysis of the most recent approaches and strategies for boosting student retention and accuracy in MOOCs that have been researched for dropout prediction.

4.1. Data-Driven Approaches

Dropout prediction for MOOCs has always relied on statistical models and traditional methodologies because of their theoretical underpinnings, ease of interpretation, and simplicity. These methods work well for assessing the connections between different elements such as course parameters, student involvement levels, demographics, and the chance of student dropout because they are based on statistical principles [44]. One of the most popular methods for binary classification issues is logistic regression, which calculates the likelihood of a student dropping out of a course based on independent variables like time invested in the course or exam results. Considered transparent in determining crucial elements, this method continues to be a mainstay in dropout prediction models. Regression analysis is one of the most popular prediction methods used in assessing and forecasting student outcomes. Regression metrics like as mean square error (MSE), root mean square error (RMSE), R^2, and mean absolute error (MAE) are commonly employed in relevant studies. Each of the metrics are described in following equations [45]:

$$\text{RMSE} = \sum_{i=1}^{n} \sqrt{(x_i - \hat{x}_i)^2} \tag{1}$$

$$\text{MSE} = \sum_{i=1}^{n} (x_i - \hat{x}_i)^2 \tag{2}$$

$$R^2 = \frac{1}{n} \sum_{i=1}^{n} |x_i - \hat{x}_i| \tag{3}$$

$$\text{MAE} = 1 - \frac{\sum_{i=1}^{n} (x_i - \hat{x}_i)^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2} \tag{4}$$

where n represents the total number of non-missing data points used in the analysis; x_i represents the actual observed value at data point I; and $\overline{x}$ and $\hat{x}_i$ represent the mean value of the observed data series and the predicted value corresponding to the observed value x_i, respectively.

Another statistical technique, the Naïve Bayes (NB) algorithm, is based on Bayes' theorem, which calculates the posterior probability of each class given a set of features. It assumes that features are conditionally independent of the target class. Despite this assumption, which may not always be true in actual scenarios, the technique is computationally efficient and generates fast, scalable predictions [45]. An alternative approach is survival analysis. It is a statistical method used to analyse and model time-to-event data. In the context of MOOCs, the event typically refers to student dropouts. The primary goal is to predict whether and when this event is likely to occur. This approach is particularly valuable because it can handle censored data, where dropout has not yet occurred by the end of the observation period. A key component of survival analysis is the survival function, which estimates the probability that a student will remain enrolled until a given time t. This methodology provides a powerful framework for studying patterns of student engagement and dropout over time [46]. For predicting student dropout, a variety of strategies have been created, ranging from sophisticated ML methods to statistical models. Because statistical approaches are transparent and easy to grasp, they are still frequently employed as basic tools for understanding the major determinants impacting student dropout [46]. In this framework, relevant research that has utilised both conventional and contemporary methods for dropout prediction in MOOCs is summarised in Table 4 and discussed in the section that follows.

A study by Anaíle et al. aims to forecast and reduce student dropout rates by using a model to enhance dropout prediction based on the characterisation of the dropout problem and the implementation of a knowledge discovery process [47]. The outcomes of three models used in the study—DT, NN, and LR (linear regression)—are combined in the ensemble model. The proposed model is able to effectively detect 98% of dropouts and correctly classify the student's enrolment and dropping. When the suggested model was compared to the RF ensemble approach, it showed positive traits that would help the management suggest measures to keep students in sessions. Neema focusses on how well different data balancing strategies work to increase machine learning models' predictive accuracy for student dropout rates [48]. The study highlights the difficulties caused by unbalanced datasets, which frequently contain an under-representation of dropout instances and hence skewed forecasts. The study shows notable gains in predicting accuracy by using techniques like SMOTE and assessing their effects on various ML algorithms. According to the findings, applying these balancing techniques can improve model accuracy, which makes them a useful addition to dropout prediction and educational data mining strategies.

Table 4. Summarised papers related to data-driven approaches for MOOC dropout prediction.

Citation	Contributions	Advantages	Shortcomings	Prediction Accuracy	Efficiency	Best Algorithm
[47]	Ensemble model integrates statistical approaches for dropout classification	High classification accuracy and enhancing reliability in predicting at-risk students	Does not capture all the factors that influence student dropout	89%	High	LR, DT, NN
[48]	Data balancing techniques	Use of SMOTE to enhance accuracy in predicting student dropouts	Weak selected features, and noise from the synthetic samples	83.2%	Moderate	LR, MLP and RF
[49]	Probabilistic technique for predicting university student dropout	Enhance prediction accuracy and handle binary outcomes effectively	Unable to capture complex relationships between variables	90%	High	LR
[50]	Predictive model for forecast dropout rates for ideological and political MOOC learners	Early identification of at risk students to improve retention rates	Potential biases in the data	Up to 83.7%,	Moderate	LR
[51]	DL-based temporal prediction	Accurate the personalised interventions for at-risk students	Heavy data dependency and complex implementation	Approx. 86 to 98%	High	DT, DL

Efficiency keys: high efficiency (>85%); moderate efficiency (70–85%); low efficiency (<70%).

B. Ujkani et al. investigate the probability of student dropout in higher education using LR as a predictive method [49]. Academic achievement, demographic characteristics, and socioeconomic position are among the important elements the authors identify as influencing dropout rates. To develop a model that can accurately identify students who are at risk of dropping out, the study analysed these factors using LR. According to the findings, educational institutions can improve student retention and success rates by using this prediction model to assist them in executing timely interventions and support measures. In tackling the issue of academic dropout, the study emphasises the value of data-driven strategies. Zhang, Y. et al. [50] use big data analytics to present a predictive model that predicts dropout rates among students in MOOC. Learner involvement, motivation, and demographics are the variables the authors examine when analysing dropout rates. The study creates a model that accurately predicts which students are at risk of dropping out by using the logistic regression method on massive datasets. This methodology can help MOOC providers execute timely interventions to improve student success and retention, according to the results.

A study by Wanli Xing et al. conducted research on how to improve dropout prediction models for MOOC by utilising DL algorithms [51]. The authors contended that high attrition rates result from traditional educational systems' frequent inability to promptly identify at-risk students. Their DL algorithm accurately forecasts the likelihood of individual dropouts by utilising a temporal prediction mechanism, which enables customised interventions based on the needs of the students. It enables customised interventions by predicting when students are most likely to drop out based on their engagement patterns over time. By analysing temporal data, such as student activity logs, participation frequency, and engagement levels throughout the course, this mechanism can identify students who may drop out at specific time points during the course. The research shows that this strategy not only outperforms baseline algorithms in prediction accuracy but also makes targeted support for learners who are at risk easier, with the ultimate goal of increasing student retention in MOOCs.

In MOOCs, feature engineering, which includes feature extraction and feature selection to improve model performance, is essential for dropout prediction. In order to identify pertinent patterns suggestive of dropout risk, feature extraction entails extracting useful measures from raw data, such as user engagement levels or interaction rates [52]. In order to increase the accuracy and efficiency of the model, feature selection then concentrates on finding the most important predictors from these measurements. Feature selection guarantees that the model employs only the most informative variables, minimising dimensionality and preventing over-fitting. This results in more accurate predictions and deeper insights into the causes impacting student attrition. Together, these processes transform raw data into actionable insights, enabling more effective dropout prediction [53]. The optimisation of feature selection is based on how to properly extract the feature subsets following the process of feature prediction.

DL techniques, like CNNs, make it possible to select the best input data and extract the most important features. Through the given CNN architecture, as shown in Figure 10, data simplification can be facilitated by the network and its automated feature extraction and selection [53,54]. CNNs are trained to selectively draw attention to important aspects while downplaying or disregarding less important or distracting ones. The CNN-based MOOC dropout prediction algorithm makes it possible to effectively capture local patterns and characteristics. Without further guidance, the model can use a filtering mechanism to gather information about the outside world once it has been educated [54]. One of the main benefits of such DL algorithms is their capacity for adaptive learning, feature extraction, and feature selection.

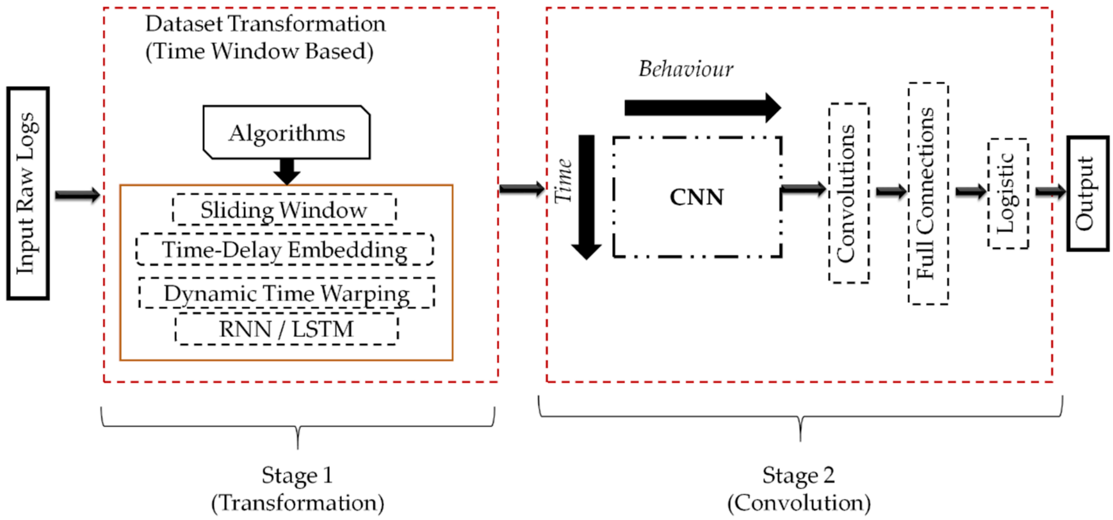

Figure 10. Overall architecture of CNN.

To handle the complexity of the data in MOOCs, DL's classification techniques, such as FNN, RNN, LSTM, and DRL, were applied in different related works. Depending on the type of data, the combination of these methods allows for improving generalisation by managing students' time series data and handling complex data patterns [54]. During data collection, the primary desirable qualities of the hybrid model that were taken into consideration were its capacity to comprehend multimodal input, its dependence on sequential and temporal data, and its reduction in overfitting.

For data with long-term dependencies, LSTM combined with other DL algorithms allows for greater performance. It uses a very complex network of memory cells to recall such small things over such long times. Input, forget, and output gates, as well as memory cells, make up an LSTM. It retains the most pertinent information and discards the remainder after going through these filters [55]. The forget gate is an essential operator that determines which data points can be safely deleted in the last step using a sigmoid function. Through backpropagation, the model adjusts the weights of the network during training by minimising the error in the predicted output. LSTM is trained, enabling it to draw conclusions from the past and modify its current state accordingly. Applying LSTM to time series data makes it a very useful tool. The following equations describe the operation of the LSTM model [56].

$$in\ (\mathrm{t}) = \sigma\ (w_{in} \times [h_s\ (\mathrm{t}-1),\ y(\mathrm{t})] + b_{in}) \tag{5}$$

$$f_g\ (\mathrm{t}) = \sigma\ \left(h_{fg} \times [h_s\ (\mathrm{t}-1),\ y(\mathrm{t})] + b_{fg}\right) \tag{6}$$

$$out\ (\mathrm{t}) = \sigma\ (h_{out} \times [h_s\ (\mathrm{t}-1),\ y(\mathrm{t})] + b_{out}) \tag{7}$$

$$c_s(t) = f_g(\mathrm{t})\ \times\ (h_{out} \times c_s(t) + in\ (\mathrm{t})\ \times\ tanh(w_{cs} \times [h_s\ (\mathrm{t}-1),\ y(\mathrm{t})] + b_{cs} \tag{8}$$

$$h_s(\mathrm{t}) = out\ (\mathrm{t})\ \times\ tanh(c_s(t)) \tag{9}$$

where σ denotes a sigmoid activation function; y(t) denotes an input at time t; $h_s(\mathrm{t}-1)$ is for the hidden state; and b_{in}, b_{out}, b_{fg}, and b_{cs} represent the values of bias vectors; and *in* (t), *out* (t), and f_g(t) represent the gates for input, output, and forget, respectively. The weights of the matrices are represented by w_{in}, b_{in}, b_{out}, and b_{fg}.

In the above equations, the biases and weights are adjusted to pass the data through the neural layers to extract the expected outcome according to the pattern learned. These can help in adapting the LSTM model to accurately predict the dropout with effective feature selection. The model's performance can be strongly impacted by the calibre of the features that are retrieved and chosen [55,56]. Sophisticated methods like ML/DL models, which automatically choose and highlight key traits, improve this procedure even more. The relevant works that have investigated different feature engineering techniques to enhance dropout prediction in MOOCs are summarised in Table 5 and reviewed in the following paragraphs. The percentage of used algorithms and the methodologies' performances in the related papers discussed are shown in Figure 11.

A study by Alghamdi explores the different aspects of MOOCs that influence learner contentment [57]. The study uses ML approaches to discover important factors that influence the dropout, including peer interaction, instructor reputation, and the quality of the course material. Based on learner choices, the findings give educators and course designer's useful information to improve MOOC offerings. Although the study makes a substantial contribution to understanding of learner happiness, its reach is constrained by the dataset's narrow breadth, which might not fully capture the range of MOOC learner experiences. Lopa et al. emphasise providing relevant features to improve prediction results [58]. ML algorithms perform and analyse these predictions. The RFC achieved 86.14% accuracy, outperforming the LR, DT, GBC, and Xgb classifier used in the proposed study.

Blundo et al. present a novel method that uses a smaller collection of characteristics to train a set of dropout predictors [59]. The main concept is to utilise weekly data to identify students who are either likely to drop out or not with a good degree of accuracy. When there is uncertainty, the choice on classification is postponed until the following week, when fresh information becomes accessible. The results show that the method enables us to be cautious and mindful of the course schedule. Additionally, the study employs a rule-mining technique to enable the explicit characterisation of behaviour that leads to dropouts. A study by Psathas et al. uses self-regulated learning (SRL) behaviours to examine how early student dropout rates in MOOCs might be predicted [60]. The authors analyse dropout predictions using five well-known ML models: NB, LR, SVM, DT, and k-nearest neighbours. The study also improves forecast accuracy by combining traditional interaction measurements with SRL behaviours. The results show that adding SRL factors greatly enhances dropout predictions, emphasising certain aspects that are essential for determining students who are at risk.

Table 5. Summarised papers related to feature selection approaches for MOOC dropout prediction.

Citation	Contributions	Advantages	Shortcomings	Prediction Accuracy	Efficiency	Best Algorithm
[57]	Identify and evaluate key factors influencing learner satisfaction	Enhance the understanding of learner preferences and improves MOOC design	Limited scope in terms of the dataset because it does not fully capture the preferences of all MOOC learners	80.9%	Moderate	LDA, SVR
[58]	Proper feature selection for prediction outcomes	Improve prediction accuracy by selecting correct features	No production model is provided; weak performance measurements	86.14%	High	RFC
[59]	Novel process to train a set of dropouts predictors via reduced features	Improve a set of features extraction for efficient predict students' Dropout	Weak in enhancing the MOOC dropout performance	97%	High	RS, 3WD
[60]	Integrating self-regulated learning behaviours into ML	Improved accuracy of dropout predictions	Limited generalisability due to course variability static analysis; does not capture dynamic student behaviours	Approx. between 85 and 90%	High	LR, KNN
[61]	Novel approach to automatically select the optimal subset of features for dropout prediction	Efficient and accurate automatic feature selector	No dropout prediction method is reviewed after feature selection	NA	NA	DL, GA
[62]	Feature selection based on PCA and LDA approaches	Best prediction accuracy	Inherent error rate of 4.28% with a standard error of 0.82	>87.7%	High	SVM
[63]	Reliable prediction-based RF	Stable and accurate predictive methodology	Need for interpretable model to explain factors contributing to student dropout	87.5%	High	RF
[64]	Optimised SVR model	Efficient feature extraction	Focus on quantitative metrics and temporal dynamics	Up to 90%	High	SVR with quantum PSO
[65]	DFS method to predict MOOC students' performance	Better accuracy for generated features	Uninterpretable generated features; limited in understanding student factors	NA	NA	PCA, KNN
[66]	Integrated model for fine-grained feature generation and prediction	Improved prediction accuracy and model interpretability	Higher dropout	Approx. 85%	Moderate	LR

Efficiency keys: high efficiency (>85%); moderate efficiency (70–85%); low efficiency (<70%).

A unique method based on evolutionary algorithms is presented by Cheng, Y. et al. to automatically choose the best collection of features for dropout prediction [61]. The technique is tested using a dataset of about 248 k student records from a university in

Brazil. The findings demonstrate that in comparison to earlier research in the literature, the suggested strategy considerably improves dropout prediction accuracy. A hybrid approach is provided by Poudyal et al. and is used to integrate several techniques such as feature extraction, regressions, and classifications [62]. They integrated principal component analysis and linear discriminant analysis. Two methods for extracting features are used: a linear and kernel SVM classifies attention patterns using extracted characteristics. The results show that the best outcomes in terms of recall, accuracy, precision, F1, and kappa were obtained via the hybrid approach. Given a collection of features designed from the daily learning progress of the students, a model to forecast student dropout from a MOOC course is presented by Dass et al. [63]. The study uses ML algorithms for prediction process to develop an RF model approach. The generated model has the ability to predict, with an accuracy of 87.5% and an F1-score of 87.5%, whether students will drop out or continue in the MOOC course on any given day.

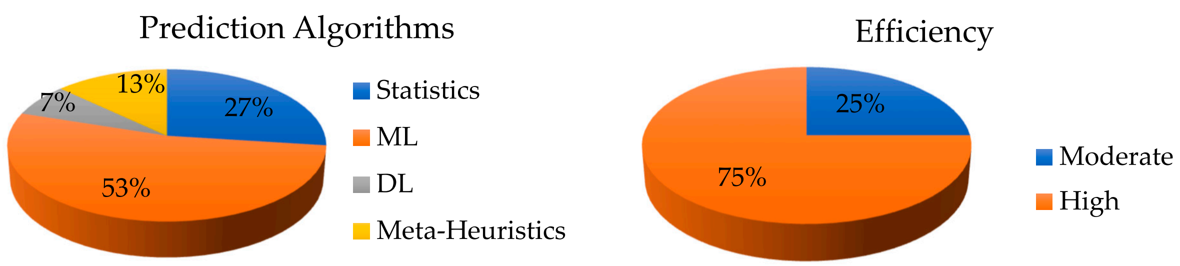

Figure 11. Related works featuring engineering algorithms and performance efficiency.

Cong Jin presented a learning behaviour data-driven student dropout prediction model for MOOCs. An enhanced quantum PSO (IQPSO) algorithm is used to calculate the parameters of an intelligently optimised SVR model, which is subsequently used in the study to reflect weekly student learning characteristics through a feature extraction method design [64]. Results from experiments using publicly available data show that the suggested model performs better in terms of prediction than a number of benchmark models, underscoring the significance of parameter optimisation and learning behaviour aspects for successful dropout prediction in MOOCs. Nadirah et al. describe the process of feature engineering for forecasting MOOC student success using the DFS method. The experiment creates features and the top features picked using PCA are the features generated by the procedure [65]. In terms of prediction, comparing both based and created features, the results demonstrate that generated features proposed using the k-nearest neighbours technique have higher accuracy, indicating that the method has the potential to be used for future prediction models.

The integrated framework with feature selection model proposed by LIN QIU et al., combines' dropout prediction, feature generation, and feature selection to predict the dropout rate in MOOCs [66]. The proposed model generates features using a fine-grained feature generation approach in days, selects valid features using an ensemble feature selection method, and then feeds the selected features into a LR model for prediction. The model obtained good results compared to existing dropout prediction algorithms according to extensive trials conducted on a public dataset. Ultimately, recommendations for building MOOCs are presented by means of an examination of the characteristics of the ultimate choice. The utilisation of data mining techniques is crucial for the analysis and comprehension of intricate datasets pertaining to MOOC dropout prediction. With the use of clustering, it is possible to find patterns and structures in data without the need for labels. Clustering can be used to identify separate groups of students with similar behaviours or traits, such as engagement levels, performance patterns, or demographics, in the context of dropout prediction.

Researchers can gain a better understanding of the various learner types and their dropout risks by grouping students into clusters. This allows for more specialised interventions and individualised support plans. Clustering techniques, like K-means, hierarchical clustering, or DBSCAN, can identify at-risk groups and improve the efficacy of dropout prevention programs by offering insightful information about the underlying structure of the data. Classification metrics are particularly useful in data mining when attempting to uncover patterns and create models from the data [67]. Its applications include evaluating how effectively those models predict class labels for tasks such as classifying students as "enrolled" or "dropouts". Table 6 displays the categorisation metrics that have been applied in the relevant literature review research. Through the use of clustering techniques, researchers can improve the accuracy of dropout prevention efforts by gaining a greater understanding of at-risk populations [68]. A review of associated studies that have used classification metrics and clustering approaches to analyse and predicts MOOC dropout is summarised in Table 7 and discussed in the following paragraphs.

Table 6. Calcification metrics and corresponding formulae for production process.

Metric	Formulae	Terms Description
Recall (R)	$\frac{TP}{TP+FN}$	TP: True Positive; FN: False Negative
Accuracy	$\frac{TP+TN}{TP+TN+FN+FP}$	TN: True Negative; FP: False Positive
Precision (P)	$\frac{TP}{TP+FP}$	-
F-score	$2 \times \frac{R \times P}{R+P}$	R: Recall; P: Precision

A study by Xia, X et al. explored certain key aspects related to learning behaviour and achieved the fusion of sophisticated algorithms using data-driven approaches, incorporating lengthy short-term memory mechanisms to build dropout prediction methods and models [69]. Based on the experimental results, the study discussed the relevant problem of dropouts related to STEM learning behaviour, investigated the key dropout temporal sequences of the learning process, identified related factors that have a significant impact on learning behaviour, and proposed intervention measures and early warning suggestions. The study demonstrated effective approaches and decisions for researching MOOC STEM learning behaviour, as well as high research feasibility and urgency.

Xiaona et al. used the multi-temporal sequences of learning behaviours to study dropout prediction in MOOCs. Their study analyses the temporal dynamics of student engagement and creates a dropout prediction model using a multi-behaviour type association framework [70]. It highlights the significance of intervention tactics based on learning behaviour data and adaptive decision feedback. The results show how well the suggested model predicts dropouts, offering a dependable and workable method for selecting short-term learning methods and long-term monitoring. This makes a substantial theoretical and practical contribution to the field of educational technology. The study of C. Xu et al. takes into account the correlation data of students' learning behaviours across a number of days in a row. It is discovered via a thorough statistical analysis of learners' learning behaviour that the following day's learning behaviour is comparable to that of the days prior to it [71]. Drawing from this attribute, they suggest utilising a Lie group region covariance matrix to depict the local correlation data related to learning behaviour. Additionally, they build a convolutional neural network model featuring a multiple expansion assembly unit to extract the high-level local correlation characteristics of learning behaviour for dropout prediction.

Table 7. Summarised papers related to data mining approaches for MOOC dropout prediction.

Citation	Contributions	Advantages	Shortcomings	Prediction Accuracy	Efficiency	Best Algorithm
[69]	Novel behavioural data analysis integrating LSTM with DL	Improve the quality and effectiveness of dropout prediction	Not directly applicable to other online learning environments; complexity	NA	Moderate	LSTM, RNN, CNN
[70]	Multi-temporal learning behaviour for adaptive monitoring mechanism	Improve prediction accuracy and facilitate adaptive interventions	Complex models	Between 88% and 90%	High	DPM-MTS
[71]	Method of extracting the local correlation high-level features of learning behaviour for dropout prediction	Enhance the dropout prediction base on evaluating learning behaviour to improve the completion rate of MOOCs	Misses the consideration of external factors influencing dropout rates, in addition to low accuracy	Approx. 84.6%	Moderate	SVM, RF, GBDT
[72]	MOOC behaviour analysis performance prediction based on entropy	Best correlation between entropy and the corresponding behavioural features; efficient feature prediction	Focused only on feature prediction performance; does not consider MOOC dropout prediction	Up to 97%	Moderate	LR, DT, SVM, RF
[73]	Attention mechanisms via natural language processing	Improve prediction accuracy and enhance interpretability	Complex and contingent upon data quality	Between 85% and 90%	High	Attn-L
[74]	Synthetic minority over-sampling technique (SMOTE) for predictive performance	Improve the accuracy of predicting at-risk students and prediction performance	Limited scalability	97.50%	High	LR, RF, and KNN
[75]	Clustering techniques and tensor completion for predicting course drop-out rates	Promote the quality of education	Low precision and slightly higher dropout	NA	Low	Pigeon-inspired optimisation

Efficiency keys: high efficiency (>85%); moderate efficiency (70–85%); low efficiency (<70%).

A data mining technique is presented by Xiaoliang et al. to handle the substantial amount of learner log data that are available on the MOOCs platform and then analyse the association between learner behaviour and academic performance [72]. An entropy-based indicator is presented to quantify behaviour change trends based on the behaviour log data of learners taking part in 12 MOOC courses. The indicator investigates the connections between behaviour change trends and learners' academic achievement. The authors proceed to construct a collection of behavioural attributes in order to examine the connections between actions and academic achievement in more detail. The results show that the proposed method has a specific association that may be used to accurately depict behavioural change trends. Also, the results demonstrate that the suggested feature selection-based model can successfully pinpoint important characteristics and achieve strong prediction accuracy.

Shengjun et al. investigate the use of attention mechanisms to forecast dropout rates among students enrolled in MOOCs. Their paper presents a unique NN model that uses attention layers to focus on the most important aspects of student behaviour, improving prediction accuracy [73]. The results show that in terms of predictive performance, the attention-based model performs noticeably better than standard models, offering important new information about the aspects that most influence student dropout. A new method for improving student dropout prediction in MOOCs is proposed by Mulyani et al., which advances the field of educational data mining [74]. Using ensemble learning techniques to improve predictive accuracy, the study combines the application of the SMOTE to address class imbalance. Liao et al. propose a combined technique called MOOP which consists of a global tensor and local tensor. Their proposal is to utilise a global tensor structure to represent the data of online courses, and to capture the internal connections between courses, a local tensor structure is clustered [75]. In order to achieve drop-out prediction, a high-accuracy low-rank tensor completion method is utilised, together with a parameter-optimisation strategy inspired by pigeons. Using MOOC data, the suggested approach is empirically assessed and shown to provide a remarkably improved level of accuracy and efficiency compared to other options.

For all previously discussed studies, the fact that the difficulty of adapting models to different learning settings and user behaviours can affect prediction accuracy and effectiveness is one of the main problems with dropout prediction for MOOCs. Different algorithms has been used in these studies, and their percentage uses are illustrated in Figure 12. From the review, prediction improving, model adaption, and data gathering techniques that work well can greatly improve the accuracy of dropout prediction.

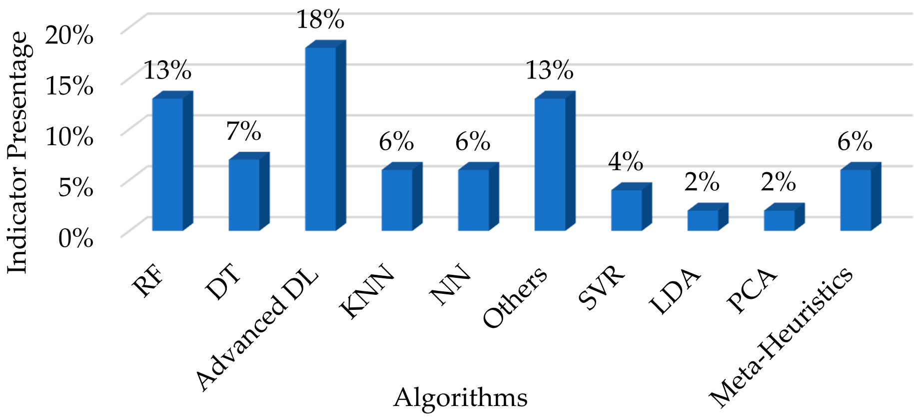

Figure 12. Distribution of data-driven insight models in the literature.

Moreover, dropout prediction faces a major difficulty due to the increasing complexity of MOOCs. Accurately identifying likely dropout learners becomes harder as the courses become larger and more diverse. This complexity is a common problem for traditional feature selection techniques used in dropout prediction algorithms, resulting in poor performance and high false-positive rates. An ensemble model that optimises complex problem spaces using meta-heuristic techniques could be suggested as a solution to this issue. The goal of this combination would be to increase the accuracy of dropout predictions, especially in complex learning contexts.

4.2. Advanced Approaches

ML algorithms are increasingly being used in MOOCs to provide new predictions and insights from data. With large datasets, DL has gained popularity as a predictive analytics technique and can make significant contributions to dropout prediction. When studying complex MOOC datasets, its ability to extract features from raw data automatically makes it very successful [76]. Models with multiple layers of nonlinear information processing and supervised or unsupervised learning techniques for feature representations at more abstract levels will help improve the accuracy of MOOC dropout predictions. The majority of ML-based classification approaches are frequent, where they function as fundamental instruments for classifying data into pre-established labels or classes. These methods use algorithms like LR, RF, DT, and SVM to create models that can forecast results based on input data. To increase the accuracy of the model, ML classification techniques in conventional contexts frequently call for manual feature engineering, in which particular features are chosen or altered from the data [76,77]. Despite their widespread effectiveness, these methods may have drawbacks when working with large, complicated datasets because they mostly rely on domain knowledge for feature extraction. When comparing DL models to traditional ML approaches, the former regularly yield superior predictive accuracy. This is primarily due to the fact that DL algorithms perform more accurately than ML as they can automatically recognise complex patterns and features from massive datasets.

DL models, such as the multi-layer artificial neural network process shown in Figure 13, can handle massive amounts of data and learn abstract representations that improve prediction accuracy, whereas traditional models rely on hand-picking features and simpler structures [77]. Moreover, when compared to ML, DL models regularly produce superior predicted accuracy, and they are able to automatically extract intricate patterns and characteristics from vast datasets. The reason behind DL outperforming ML is that it can handle enormous volumes of data and develop abstract representations that improve prediction accuracy. Prior to exploring the associated literature, it is critical to acknowledge the increasing significance of ML/DL models in attaining improved prediction accuracy in learning environments, especially on MOOC platforms. These sophisticated algorithms present fresh chances to improve the forecasting of student outcomes, such as course completion and dropout rates. The studies that have used ML/DL techniques are examined in the following paragraphs, which also highlight their methods and fill in some of the gaps, particularly with regard to hybrid model creation for large educational datasets. These studies are summarised in Table 8.

A study by Kowsar et al. provides early dropout prediction using ensemble models combining CNN and LSTM networks. The study provides at-risk students with timely treatments and highlights how crucial it is to anticipate dropout in the first week of classes [78]. The authors generate a thirty-day correlation matrix for learners and six novel ensemble classification models. The results show that these models greatly improve prediction accuracy; in addition, prediction errors can be successfully reduced when ensemble classifiers are used in conjunction with the NN approach. The learning behaviour feature-fused deep learning network model is suggested by Liu et al. for the prediction of MOOC dropout. The LBDL model separates the features of the data into two categories, namely, general information features and video learning behaviour characteristics, with time series information [79]. The model mines time series data for video learning behaviour features using Bi-LSTM and attention mechanisms, and it processes general information features using fully connected and embedding layers. To train the gradient boosting framework, LightGBM, the original feature information is integrated. The proposed model achieved greater AUC and F1-scores than other baseline models and performs better when it comes to dropout rate prediction.

Patel et al. investigate ML methods for forecasting MOOC dropout rates. They assess seven algorithms to forecast dropout rates and student outcomes. In order to improve student retention, the authors also suggest tailored treatments based on unique dropout probability [80]. The results highlight machine learning's potential to address MOOC dropout issues and provide guidance to educators on how best to support students who are at risk of dropping out. An innovative dropout prediction model for MOOC learning clickstream data analysis is presented by X. Zhang et al. It is named the Image Convolutional and Bidirectional Temporal Convolutional Network (IC-BTCN). The proposed study tackles the important problem of high dropout rates in MOOCs [81]. Utilising convolutional approaches, the technology extracts local features from a 3D learning behaviour matrix created by converting student learning data. Utilising dilated causal convolution to collect temporal learning features, these features are subsequently processed through a temporal convolutional network to further improve the input. According to the results, the proposed method's accuracy rate in forecasting student dropouts is 89.3%.

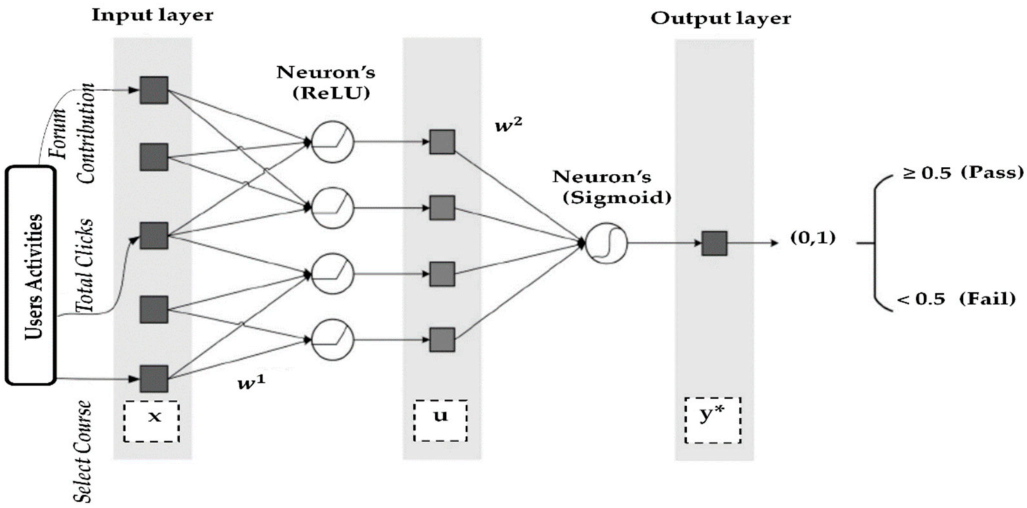

Figure 13. Artificial neural network model for student retention classification performance.

Pan F et al. address the difficulty of predicting student dropouts by providing a unique approach that uses multi-objective reinforcement learning (MORL) to maximise the trade-off between forecast accuracy and earliness [82]. The suggested method approaches dropout prediction as a partial sequence classification problem using a multiple-objective Markov decision process. The authors also describe an improved envelope Q-learning technique for thoroughly exploring the solution space, with the goal of identifying Pareto-optimal techniques that accommodate a variety of tastes. The model's usefulness is proven through rigorous testing on real-world MOOC datasets.

Amala et al. created an Arabic/English MOOC platform called the "Tadakul" system and assessed the impact of the learning analytics system. The authors also present a novel DL approach that is employed in the system to monitor learners' learning processes based on their interactions with registered courses. The input gained through the DL can be used to enhance the student learning experience [83]. They also study the learners' feedback and offer a new technique that combines a Bidirectional Long Short-Term Memory (BiLSTM) network with a CNN approach. Their results show that the proposed model outperformed traditional ML approaches in predicting learners' learning behaviour and enhanced the student experience by increasing teaching efficiency and effectiveness.

A study by Smaili et al. introduces a novel prediction model that offers at-risk learners the optimum learning path in order to address the problem of low retention rates in MOOCs [84]. The authors validated the efficacy of the proposed strategy with a case study that uses the ACO algorithm to optimise learners' routes and classification techniques for prediction. They suggest two complimentary strategies to lower MOOC dropout rates. The first is to use an adaptive boosting algorithm to forecast whether or not learners are at risk of quitting the MOOC program. The second approach entails developing profiles to assist in material selection and decision-making, thereby customising the educational experience for students identified as at risk. The proposed approach incorporates adaptive learning by dynamically modifying the course content to suit the needs of at-risk learners, drawing on the experiences of other learners with comparable profiles.

Table 8. Recent studies on dropout prediction based on ML/DL/meta-heuristic techniques.

Citations	Contributions	Baseline Algorithms	Proposed Models	Prediction Results	Efficiency	Datasets
[78]	Introduces six novel ensemble models that leverage CNN and LSTM for early dropout prediction in MOOCs	CNN, LSTM	Boosting CNN-LSTM	AC = 85.7% P = 87% R = 97.2% F = 91.8%	High	MOOCCube
[79]	Novel deep learning model integrates learning behaviour features to enhance the prediction of dropout rates	DL	Learning behaviour feature-fused DL network	AUC = 82.4% F = 74.9%	Low	NA
[80]	Build a prediction model for the early identification of at-risk students based on ML approaches	DT, RF, GNB, AdaBoost, ETC, XGBoost, and MLP	Multi-models-based approach	AC = 86% F = 92% R = 92% P = 99%	High	OULAD
[81]	Dropout prediction model based on image Convolutional and Bidirectional temporal Convolutional Network	CNN, MLP	IC-BTCN	AC = 89.3%	Moderate	KDD CUP 2015
[82]	Multi-objective reinforcement learning approach to optimise the trade-off between dropout prediction accuracy and earliness	RL	Multiple-objective Markov decision process	AC = 83.9%	Moderate	KDDCup2015 and XuetangX
[83]	Novel deep-learning approach monitors the learning process in a developing MOOC platform	CNN	Bidirectional LSTM	AC = 78% to 85%	Moderate	Kaggle platform
[84]	Adaptive boosting algorithm for predicting learners at risk of dropping out	DT, XGBost, AdaBoost, GBL	ACO with GBL	AUC = 86.63%	NA	OULAD
[85]	Dropout prediction model based on a fusion of behavioural data and ML	SVM, LR, NB, RT, LDA, RF, CNN	Multi-models	AC = 88.8% R = 96% P = 89% F = 93.1%	High	KDD CUP 2015
[86]	Learning model for dropout rate prediction based on high-dimensional vector features	CNN, SVM	Ensemble DL	AC = 97.4% R = 96% P = 97% F = 97.2%	High	KDD 2015
[87]	Efficient capture of nuanced learning patterns to accurately predict student dropout behaviour	CNN, LSTM, DNN	Federated learning pattern aware	AUC = 97.7% RMSE = 0.271 MAE = 0.2084	NA	XuetangX
[88]	A dropout prediction mode based on multiscale features of student behaviour time series data extraction	SVM, CNN, NB, LDA	MFCN-VIB	AC = 87% R = 96% P = 88% F = 92.2%	High	KDD CUP 2015
[89]	Develop an effective dropout prediction model to improve student retention rates	DNN	DNN-based model	AC = 98% R = 93% P = 82% F = 87% AUC = 99%	Moderate	MITx and HarvardX
[90]	Novel dropout prediction model utilises max neighbourhood and intelligent sample weighting to enhance ML predictive performance	DT, RF, SVM, LR	IQPSO-based model	AUC = 82.39% F = 79%	Moderate	KDD Cup 2015
[91]	Predicting student dropout in a MOOC for smart city professionals at an early stage	RF, LR, Ridge, Extra, GBC, CART, AdaBoost, SVM-SGD	GBDT	AC = 95.5% Kappa = 90.7% P = 97.7% F = 96.3%	High	NA

Table 8. *Cont.*

Citations	Contributions	Baseline Algorithms	Proposed Models	Prediction Results	Efficiency	Datasets
[92]	Hyper-model for MOOC feature extraction and dropout prediction	CNN	CNN- LSTM-based model	AC = 79% to 85% R = 88% to 87% P = 91% to 93% F = 89% to 90% (two datasets)	High	KDD Cup 2015 and XuetangX
[93]	Pipeline model based on CNN and long short-term memory to predict dropout rate	SVM, LSTM, CNN	CLSA	AC = 87.6% F = 86.9% P = 87% R = 86.5%	Moderate	KDD CUP 2015
[94]	Integration of feature weighting and behavioural time series to improve MOOC dropout prediction	LR, NB, RF, DT, SVM, CNN	FWTS-CNN	AC = 87.1% P = 86.3% R = 86.5% F = 86.4%	Moderate	KDD Cup 2015
[95]	Integrated feature extraction and classification for efficient end-to-end dropout prediction model	LR, GNB, DT, RF, GTB, AdaBoost, SVM	2D-CNN	P = 84.5% R = 84.9% F = 84.2% AUC = 87.8%	Moderate	KDD Cup 2015
[96]	Dropout rate prediction model to improve prediction accuracy	RF, GBDT	RNN-based model	R^2 = 0.80	NA	XuetangX

Metrics keys (P: precision; R: recall; F: F-score; AUC: area under the curve; AC: accuracy). Efficiency keys: high efficiency (≥90%); moderate efficiency (75–90%); low efficiency (<75%).

A dropout prediction model based on the combination of behavioural data and SVM is proposed by Yujiao et al. They use a creative model that integrates many behaviour traits with varying weights and uses them as input data. The proposed method is implemented on two trained datasets, which are used to train the SVM classifier [85]. Results show that the proposed model performs better and, according to experimental results on both datasets, gives the behaviour features the same weights. A dropout rate prediction based on an ensemble DL model and on learner behaviour data features was proposed by Gaurav et al. ResNet-50 is used in the proposed study to extract the local features, and a kernel method is then applied to construct feature relations. The high-dimensional vector characteristics are extracted; then, they are fed to a Faster RCNN to produce a vector representation that includes time series data [86]. The vector is then subjected to a static attention mechanism in order to acquire an attention weight for each dimension. The results demonstrated that the suggested model could perform on par with previous dropout prediction techniques.

Zhang et al. present a FLPADPM model as a bidirectional LSTM layer and a one-dimensional CNN integrated into a federated learning framework. The proposed approach is intended to capture complex learning processes and adjust to differences in a range of educational environments [87]. The author's conduct an empirical evaluation with the two datasets to assess the performance of FLPADPM. The results shows that FLPADPM is more effective than cutting-edge models at properly predicting the behaviour of dropout students. A dropout prediction model that combines a variation information bottleneck and a multiscale fully convolutional network is proposed by Shou et al. [88]. By building a multiscale full CNN, the model extracts multiscale features from student behaviour time series data. A variation information bottleneck approach is then used to reduce the impact of noise on the prediction outcomes. According to the results, the suggested method performed the best.

Anjali et al. present a study describing how DL techniques were used to create a dropout prediction system for at-risk MOOC participants. It draws attention to the rising acceptance of MOOCs and the alarming dropout rates that jeopardise their viability [89]. The study uses a DNN model to analyse learning patterns and predict student dropouts with two MOOCs datasets. The goal of this approach is to give teachers early warnings in order that they may step in and help at-risk pupils. According to the results, DL works better than other approaches, presenting a viable way to improve MOOC student retention. Ahmed et al. present a hyper-model of the LSTM and CNN models, known as CONV-LSTM, to automatically extract features from MOOC raw data and predict if each student will finish their courses or drop out [90]. The authors take into account the problem of class imbalance, which implies that models produce good results on most class cases and poor results on some class examples. In their study, the use of loss function employs a cost-sensitive method that takes into account the different misclassification costs for false positives and false negatives in order to reward improved prediction performance. In comparison to baseline approaches, the performance of the suggested model is better.

A pipeline model called CLSA is proposed by Qian et al. to forecast the dropout rate using information on student behaviour [91]. The CLSA model constructs feature relations using a kernel technique after first extracting local features using a convolutional neural network. After that, an LSTM network receives this high-dimensional vector produced by the CNN in order to produce a time series-integrated vector representation. To determine an attention weight for each dimension, authors apply a static attention mechanism to the vector. The proposed model outperforms baseline models by over 2.8%, with an accuracy of 87.6%. Cong Jin suggests a dropout prediction model based on ML to solve the high dropout rates in MOOCs. The study presents a novel idea known as max neighbourhood, which improves prediction accuracy by taking into account both the distance between student samples and their labels [92]. In addition, it creates an alternative to conventional random selection techniques for determining initial weights for every student sample by using this new neighbourhood definition. The study in question investigates the use of clever optimisation strategies to optimise these weights in more detail. The predictive ability of the dropout prediction model is greatly enhanced by the use of intelligent optimisation and sample weighting, as evidenced by the experimental results.

A study by Panagiotakopoulos et al. proposed that student dropout in MOOCs can be predicted using a variety of cutting-edge supervised machine learning methods [93]. Their study makes use of many predictive models that have their hypermeter learning parameters tuned for the best classification performance. Based on data gathered during the first week of the course, the experimental results demonstrate that accuracy surpasses 96%, enabling efficient intervention methods and support steps. Yafeng presents a CNN model that combines behavioural time series and feature weighting. The proposed study takes the learner's log of learning activities and extracts continuous behavioural features from it [94]. It then filters important features and uses a DT algorithm to rank them in order of importance. Finally, it weighs the continuous behavioural features according to their importance and uses behavioural time series and weighted features to create a convolutional neural network model. The suggested approach provides excellent accuracy with over 87% of the results.

Qiu et al. proposed an end-to-end dropout prediction model based on CNNs to predict the student dropout problem in MOOCs [95]. The model integrates feature extraction and classification into a single framework, automatically extracting features to improve feature representation and transforming the original timestamp data according to different time windows. Comprehensive tests on a publicly available dataset demonstrate that the proposed method can produce performance metrics scores that are on par with baseline dropout prediction techniques. Sun D et al. define the dropout prediction problem as forecasting how much content in the entire course syllabus can be finished by the student [96]. To address this issue, an RNN-based dropout rate prediction model is presented by the authors, together with a URL-embedding layer. The results reveal that the model's prediction accuracy outperforms that of the classic ML model.

Several ML and DL algorithms have been used in previous studies to improve dropout prediction and analyse students' learning behaviours in MOOCs. Figure 14 shows the percentage of algorithms used in these studies. In general, the majority of research has been focused on using DL algorithms to predict different student learning outcomes with different MOOCs platform, as illustrated in Figure 15. Though ML/DL is being used more frequently, not much research has integrated the two techniques to predict outcomes with huge data from education. The current study closes this research gap by offering a thorough evaluation of the predictive performance of various DL/ML models, with a particular emphasis on enhancing dropout prediction in MOOCs through the optimisation of feature selection and dropout prediction using meta-heuristic approaches.

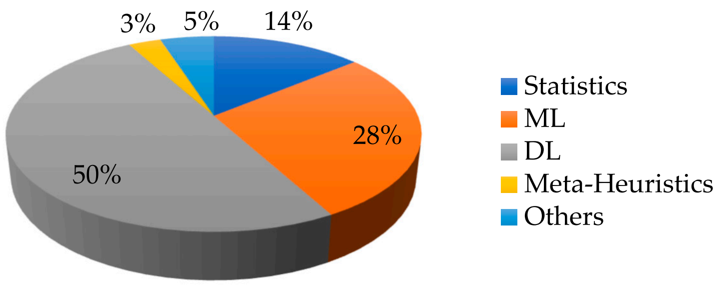

Figure 14. Distributions of algorithms used by related works (advanced approaches).

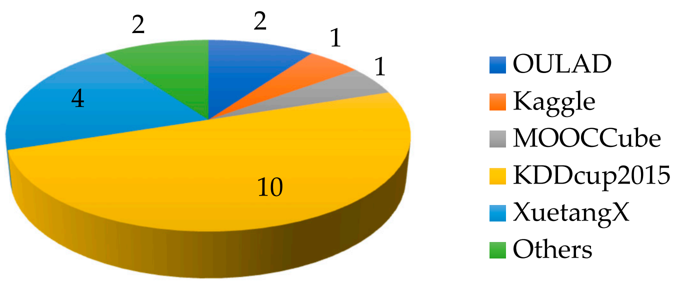

Figure 15. Distribution of MOOCs databases used by related works (advanced approaches).

Another issue with dropout prediction is the large amount of student data that needs to be analysed. Massive course enrolments and ongoing participation have led to a notable increase in the amount of learner interaction data generated every day. At this scale, manual dropout detection becomes impractical, and existing automated techniques often suffer from memory and precision issues. To address this, integrated meta-heuristic algorithms can be used to efficiently analyse large volumes of learner data and identify potential dropout patterns. This minimises the effect of data volume on prediction accuracy and makes the approach suitable for a range of learning scenarios.

4.3. Hybrid Approaches

It is important to emphasise that hybrid approaches that can combine ML, DL, meta-heuristic algorithms, and other techniques to maximise the accuracy of dropout prediction in MOOCs are increasingly being used. These hybrid models overcome the drawbacks of individual models and improve prediction accuracy by combining the advantages of several techniques. Researchers have obtained more robust predictions of student dropout and course completion by integrating multiple strategies, including intelligent learning techniques and improved feature selection. The following section discusses research that has used these hybrid approaches, compared their effectiveness, and examined potential avenues for further improvement of the dropout prediction model, as summarised in Table 9.

Table 9. Summary of most-relevant studies for dropout prediction hybrid approaches.

Citation	Year	Algorithms/ Approaches	Contributions	Dataset	Proposed Model	Prediction Results	Performance
[97]	2024	LightGBM, LMP	Utilises the learning process model to analyse learning behaviour and generate feature vectors to predict dropout	MOOCCube	LPM-LightGBM	AUC = 86.4%	NA
[98]	2024	DNN, LSTM, KNN, MLP, SVM, RF	Effectively identifies students at risk of dropping out early in their learning process, enabling timely interventions to improve retention rates	MOOCCube	DNN-LSTM	AUC = 82.4% F = 74.9%	Low
[99]	2024	NN, CNN, LSTM, SVM	Feature selection approach for prediction; opinion mining of learner comments	Coursera and MSIT datasets	Hybrid mutation-based earthworm	P = 90.9% R = 90.1% F = 91.2%	High
[100]	2023	CNN, LSTM	Hybrid NN for effective dropout prediction incorporating periodic learning behaviours and time distribution information	KDD Cup 2015	CGDC-LSTM	AC = 90.3% P = 91.2% R = 97.5% F = 94.25% AUC = 85.67	High
[101]	2023	DL, IQPSO	Hybrid novel dropout prediction model to enhance prediction accuracy effectively	MOOCCube video	IQPSO-PLSTM model	AUC = 88.45% F = 78.34%	Moderate
[102]	2022	GA, DL	Hybrid prediction model for early detection of students at high risk of drop-out in MOOC	KDD cup 2015	DL/GA-based models	AC = 94% AUC = 86%	High
[103]	2022	ANN, MLP, LR, DT, KNN	Highly accurate dropout prediction model to improve feature selection and learner retention	KDD Cup 2015	FIAR-ANN	AC = 92.42%	High
[104]	2022	CNN, TCN	Enhance and extract useful feature information from students' learning records for dropout prediction	KDD CUP 2015	CNN-TCN	NA	NA
[105]	2020	CNN, Squeesed/ excitation network (SE-Net)	Develop a DL-based model to predict learners' dropout behaviour	KDD Cup 2015	CNN-SE-gated recurrent unit (GRU)	AC = 94.5% P = 95.9% R = 97.2% F = 96.5%	High
[106]	2019	DT, ELM	Hybrid algorithm to improve dropout prediction process	KDD 2015	DT-ELM	AC = 89% AUC = 85% F = 91%	High

Metrics keys (P: precision; R: recall; F: F-score; AUC: area under the curve; AC: accuracy). Efficiency keys: high efficiency (≥90%); moderate efficiency (75–90%); low efficiency (<75%).

A MOOC dropout prediction approach utilising the LightGBM algorithm and learning process model is proposed by Nie, H. et al. To examine learning behaviour and produce feature vectors that will allow for a clear interpretation, the proposed method makes use of the learning process model [97]. To predict dropout, the LightGBM algorithm is used based on these feature vectors. The dropout prediction approach used in this study shows good performance when compared to baseline methods. El Aouifi et al. offer a hybrid deep learning model for predicting school dropout that is based on LSTM and DNN techniques [98]. The suggested model was contrasted with earlier research and a number of different ML techniques. According to the results, the accuracy and efficiency of the suggested DNN-LSTM model are superior to other evaluated models. A hybrid bio-inspired meta-heuristic feature selection strategy was presented by Jatain et al. for the analysis of learner comments on a course [99]. Experimental work is conducted over a real-world education dataset comprising of 110 K learner comments acquired from Coursera and learner data from academic institution MSIT. With respect to the gathered dataset, the experimental results show that the suggested model attains an accuracy of 92.24%.

A study by Y. Zhou introduces the notion of a periodic feature and discovers that varying the learning duration has distinct impacts on the prediction outcomes [100]. The authors provide a hybrid neural network model by using LSTM and CNN to simulate and forecast users' dropout behaviour in light of the new insights. CNNs enable the preservation of local correlations in student behaviour. The model combines dilated causal convolution and group convolution to capture students' periodic patterns. LSTM is integrated into the proposed model to extract learning time distribution information and assess the impact of different learning periods on outcomes. Experimental results show that the suggested model outperforms baseline techniques in prediction accuracy.

An improved quantum PSO technique and DL model are combined in a study by Xiao et al. The proposed methodology uses the derived behavioural feature importance as the initial weights of the features. It first pre-trains the data using random forests [101]. Then, weight finding is carried out using an improved quantum PSO technique to optimise the initial feature weights. Lastly, LSTM deep learning modelling is applied to the weighted data. The results show that the proposed model provides an acceptable performance compared to other baseline models. A hybrid feature selection and dimensionality reduction strategy were introduced by ALJ et al. to extract useful features and decrease model complexity and computation time [102]. Authors use supervised learning techniques to develop two models based on DL and GA. Their results show that both GA and DL outperform baseline algorithms employed in comparable investigations.

Nithya et al. focused on using video interaction analysis to describe important MOOC assignments such as forecasting student success and dropout rates [103]. To evaluate the dropout measure, the proposed method generates and selects the optimal characteristics depending on learner behaviour. The authors used frequent itemset-3 in the NN-based MLP for feature selection and calculated the evaluation metrics. Then, they contrasted their results with those of ANN for dropout prediction using the feature importance association rule. The proposed approach evaluated is compared to different ML and DL models, and the results show that it provides a high accuracy and improves the student retention and decreases dropout rates. In addition, the model outperformed other models by 18%. A hybrid model to predict learners' dropout behaviour is provided by H. Liu et al. [104]. This hybrid model uses an attention method to extract the crucial information from the learning records of students, rather than manually extracting features. It does this by designing a CNN model to automatically extract useful features. The authors also use TCN to capture the interactions between hidden features at various time scales. The proposed model was then evaluated, and the results show that it outperforms alternative dropout prediction techniques.

A hybrid DNN model for the prediction of learner dropout behaviour was proposed by Yan et al. The authors encrypt student recorded data, turn it into a 2D matrix, and then apply the CNN technique with squeesed and excitation networks to extract local features [105]. By obtaining the relationship between the learning behaviours, they are able to increase prediction performance. The results demonstrate that the suggested strategy outperforms the baseline in terms of performance. Chen, Jing et al. introduced an integrated DT algorithm with an extreme learning machines (ELM) hybrid model to eliminate the need for iterative training. Features that are well-suited for classification are chosen by the DT algorithm [106]. Moreover, the study establishes improved weights for the chosen characteristics to improve their capacity for classification. The authors translate the DT to ELM using entropy theory in order to improve ELM structure and obtain accurate prediction results. The results show that the proposed method outperforms baseline algorithms.

The related works discussed are the most relevant articles that focus on hybrid methodologies for MOOC dropout prediction. However, between 2018 and 2024, several other articles have proposed various hybrid solutions using ML/DL techniques, along with meta-heuristic approaches for addressing feature selection challenges and improving dropout prediction. Figure 16 shows the distributions of these publications. The primary focus on feature engineering represents the first stage as crucial to any ML/DL algorithm-based classification challenge since it involves extracting and choosing features from the initial data that are pertinent, meaningful, and information-rich. Data extraction based on the features selected in the first phase is the second most important stage. After that, these recovered data go through a scaling and cleaning processes to produce high-quality, standardised data that are prepared for the training phase.

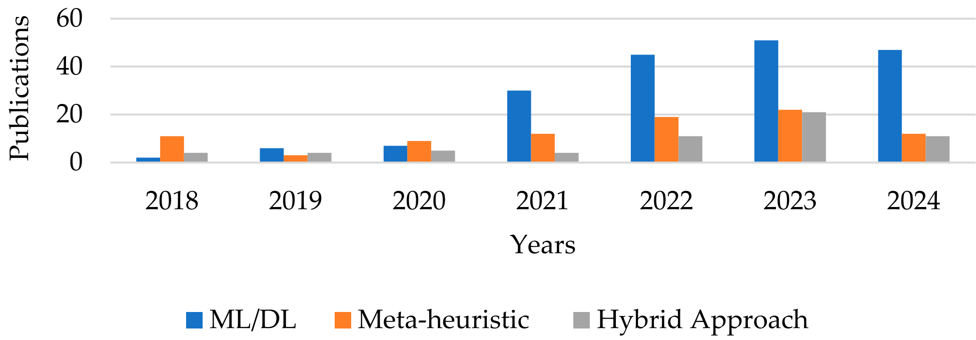

Figure 16. Distributions of hybrid ML, DL, and meta-heuristic approaches for MOOCs feature selection and dropout predictions (resource: Google Scholar).

To ensure an optimising prediction for MOOCs, the third phase involves creating data for every learner; then, the next stage looks for the best predictive traits based on wrapping techniques such as meta-heuristics algorithms. At this stage, many advanced algorithms can be used for optimisation, as shown in Figure 17. This methodology enables the classification of the original dataset into multiple subsets based on the features most relevant to each ML/DL method, which will then serve as input for the proposed prediction module. The final design stage of a hybrid model depends on the previous stages, after which the proposed model and other ML/DL algorithms are trained to compare the performance of each of them. This comparison will pit the presented predictive model against standard prediction models. Figure 17 shows the stages related to developing a hybrid dropout prediction model.

The framework illustrated in Figure 17 can be used for optimising MOOC dropout prediction models, combining a hybrid deep learning and meta-heuristic approach. This framework outlines the procedures necessary for achieving precise learning and improving dropout prediction. Initially, the collected dataset is pre-processed to eliminate errors and anomalies. Various imbalanced data learning techniques, such as SMOTE, boosting, ensemble methods, cluster-based approaches, and adaptive synthetic sampling, have been used in related works to resample the data during pre-processing [107]. The feature selection process for the next stage of the feature engineering procedure can be optimised according to meta-heuristics. To improve the prediction model, the prediction step depends on a DL-inspired hybrid learning mode. The efficacy of the model can be increased by using an optimisation function to minimise an ML/DL model's loss function. Many optimisers can be used, and some of them have been introduced in related works, such as the stochastic gradient descent optimiser, adaptive moment estimation, root mean square propagation, and the momentum optimiser. Q-learning functions as an optimiser and can be used to find the best action or series of actions to minimise or maximise a certain outcome.

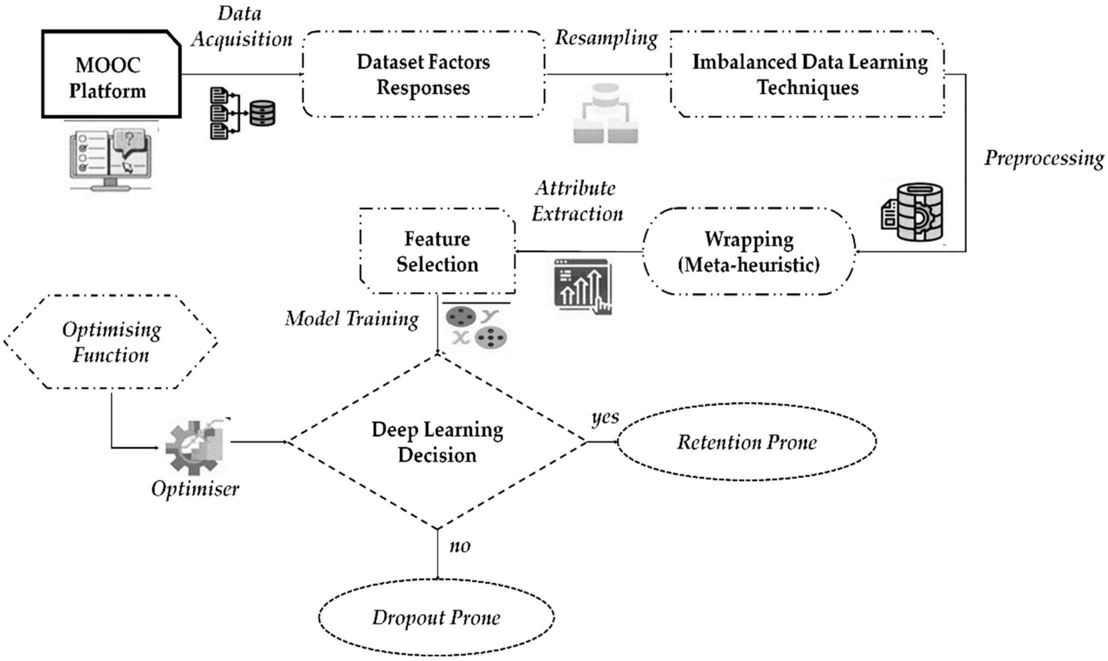

Figure 17. A hybrid framework of DL and meta-heuristic techniques for MOOC dropout prediction-based feature selection optimisation.

In summary, the various earlier studies pertaining to hybrid techniques for MOOC dropout prediction optimisation were covered in this section. We found a few papers linked to this topic since MOOCs have a specific criterion related to increasing the dropout prediction based on feature selection optimisation. More research is still needed on the MOOC performance and dropout prediction problems, which could lead to further advancements in this field. Nonetheless, numerous research projects based on ML/DL solutions are put out to improve dropout prediction. Prediction performance and accuracy can be raised by combining the optimisation of feature selection using meta-heuristic techniques with dropout prediction enhancement using ML/DL methods.

4.4. Adaptive Learning Approach

Three main components make up an adaptive learning model; these are the learner module, the learning environment module, and the decision-making module. The learning environment module offers specialised instructional materials designed to meet each learner's unique requirements. Enhancing learner motivation is mostly dependent on a carefully designed learning environment that encourages skill development and makes job completion easier [108]. Designing instructional materials and activities in a way that best supports and promotes learning objectives and learner development is therefore crucial. As shown in Figure 18, the learner module is responsible for learners' profiles, which are recognised based on data collected from their interactions with the system. Learner traits, including goals, preferences, knowledge level, learning styles, and academic motives, are often reflected in these profiles. The learner module delves into the imprints that students leave behind as they engage with the educational setting. After that, data pertaining to the learners' profiles are compiled, pre-processed, and kept in our database.

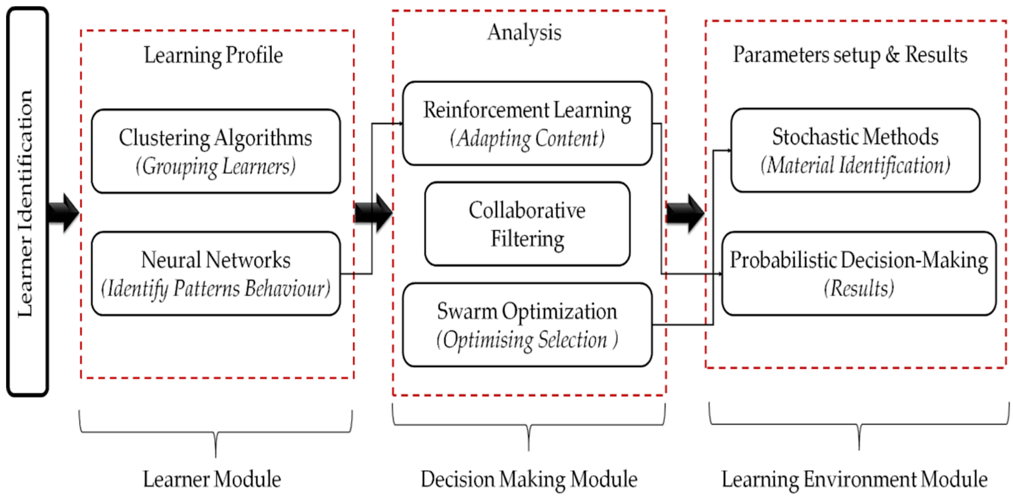

Figure 18. General adaptive learner architecture model.

The decision-making module is responsible for selecting the appropriate course materials that fit each learner's unique profile [108]. The module determines the sequence in which the learner will complete the tasks and assesses the student's readiness for the subsequent learning steps. The algorithm will automatically identify the suggested courses and the path that each learner should take based on their individual profiles. The profile, in this case, can play a crucial role by applying the meta-heuristics algorithm. In the research by Smaili et al., the authors tried to increase the effectiveness of online platforms and boost learners' academic performance by tailoring the MOOC materials to each individual learner [109]. The concept is intended to create a system that leverages learners' interactions with the learning environment to provide each learner with a personalised path that takes into consideration the diversity of their profiles. To achieve this, the authors propose building the best possible learning routes for the system using the PSO approach and perform a case study on online learning to demonstrate the efficacy of the suggested strategy.

4.5. Rationale and Review Objectives

This research aims to investigate recent developments in dropout prediction for MOOCs, as has been previously mentioned. The chosen papers show how predictive approaches have evolved and how they have been used in MOOCs. They were collected from prestigious journals and research databases between 2018 and 2024. Figure 19 shows the distribution of these studies based on the approaches used, highlighting a significant increase in the use of advanced algorithms, such as meta-heuristic methods and ML/DL techniques, for predicting student dropout.

The use of ML and DL has increased significantly over the years. While there were only a few studies on meta-heuristic methods between 2018 and 2020, research involving meta-heuristics has steadily grown in recent years. This indicates that MOOC dropout prediction studies are moving toward more advanced, data-driven models that integrate hybrid ML/DL/meta-heuristic approaches. This rise reflects growing awareness of the potential of these sophisticated techniques to improve prediction accuracy and adapt to diverse learning environments. These trends suggest that the field is rapidly evolving, with researchers employing more complex algorithms to enhance dropout prediction models. As a result, further advancements in dropout prediction methods are likely on the horizon, leading to more-effective strategies for retaining students in MOOCs and improving learning outcomes.

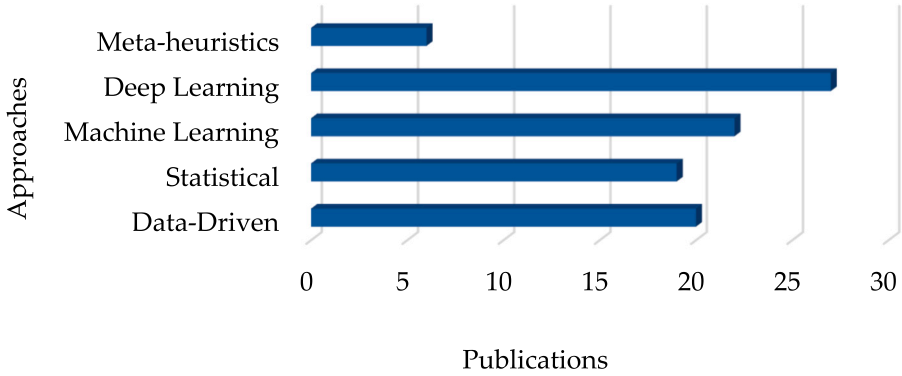

Figure 19. Distribution of related publications reviewed based on strategies used (statistical, data-driven, and intelligence approaches).

The percentage of articles gathered from various reliable and respectable sources for this study is highlighted by the distribution of publication sources. Several databases and publishers, including IEEE, MDPI, Elsevier, ACM, Springer, and Wiley, which were used as sources of collected papers. The distribution of various published sources as a proportion is shown in Figure 20. With 25% and 18% of the evaluated publications, respectively, IEEE and MDPI, according to the statistics, are the highest contributors to research on developing MOOCs dropout prediction models. Papers from these sources provide important new information on dropout prediction for MOOCs. ACM and IEEE digital libraries were the main sources of conference proceedings; other journals with respectable impact factors were also included such as Tylor & Francis, Nature, and Frontiers. Interestingly, IEEE and MDPI offer a good number of periodicals that are strongly recommended for scholars looking for MOOCs dropout prediction methods. This distribution highlights the important role that IEEE and MDPI have had in promoting this field of research.

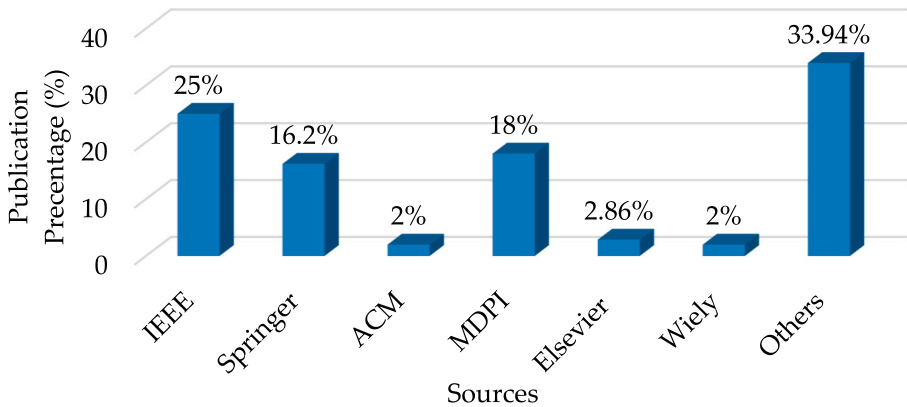

Figure 20. The percentage of all papers considered in this review based on publication venue.

5. Comprehensive Review Synthesis and Insights

Many techniques and methods have been utilised to predict dropout rates in MOOCs and increase student retention based on the earlier research, which is described in Section 4. A variety of dropout prediction strategies have been discussed in this review, with their efficacy and advantages and disadvantages noted. The review provides an informed analysis of existing prediction models and the viewpoints offered by scholars by examining the most recent research. In addition, it fills in the gaps in the literature that have not been thoroughly explored. This comprehensive analysis not only shows how far dropout prediction has come but also emphasises how much more work has yet to be done to close these gaps and improve prediction methods in order that they can be used more accurately and effectively in MOOCs.

It is evident from the discussed methodologies used in most MOOC dropout prediction studies that a range of predictive models and strategies can greatly enhance the accuracy of dropout forecasts. However, to maximise prediction performance, the right features and methods must be chosen. Numerous studies have concentrated on identifying pertinent features and optimising methods to increase model accuracy. Despite these developments, there are still a number of important variables that have a significant impact on prediction performance. These variables include learner engagement levels, the complexity of the course material, and user interaction patterns. Moreover, MOOCs dropout prediction is generally faced with a number of difficulties. These difficulties draw attention to important areas of research that could enhance learner memory and prediction accuracy. The evaluated research in this paper covers some of these problems, with many depending on advanced prediction models and ML/DL approaches [110]. To further improve forecast accuracy and efficacy in MOOCs, more thorough strategies that take into account a larger range of factors influencing dropout rates are still required.

Many of the reviewed studies often overlook the opportunity to optimise data exchange paths and select predictive models for dropout prediction based on key factors such as feature relevance, learner engagement, and model performance. Moreover, many of these studies also overlook the challenges of integrating diverse datasets and learner behaviours into predictive models, which can impact the accuracy and effectiveness of dropout predictions [110,111]. To address these limitations, future research could focus on implementing advanced algorithms and combining deep learning with meta-heuristic techniques. By refining feature selection, data integration, and model tuning, the accuracy of dropout predictions and the overall reliability of MOOCs can be improved. In addition, incorporating factors such as student engagement patterns and course interactions could further enhance prediction models and help reduce dropout rates.

Previous research has explored various techniques for feature selection and model training, but these methods often involve complex, resource-intensive processes that may not fully capture the nuances of learner behaviour. Future hybrid approaches could enable more efficient data processing and sophisticated routing mechanisms in predictive models, enhancing both energy efficiency and model performance [111].

6. Research Gaps in Dropout Prediction Optimisation Perspective

Reducing dropout rates in the MOOC space is still a major challenge. Due to the varied and dynamic nature of learner behaviour, one of the main problems found in recent studies is the difficulty faced in accurately predicting dropout rates. Conventional approaches frequently use static models that might not take into consideration changes in learner engagement and progress that occur in real time, resulting in forecasts and interventions that are not ideal. New developments in data analytics and ML present viable ways to improve dropout prediction models [112]. A variety of features, including student interaction patterns, engagement levels, and past performance data, are incorporated into these techniques. Many existing studies have not fully explored the potential of integrating real-time data with adaptive learning strategies. This gap indicates that additional study should be undertaken on creating models that can more accurately anticipate the future by dynamically adapting to the behaviour of learners. In addition, even while clustering techniques and other predictive models have demonstrated promise, they frequently run into issues like managing massive datasets and making sure the models hold up well across various course types and learner demographics [113]. For example, although clustering might be useful in identifying trends and putting comparable learners in one group, it might not always be able to convey the subtleties of each learner's experience. More advanced algorithms are required in order to handle a variety of datasets and offer insights into the reasons contributing to dropout.

Furthermore, the significance of combining intervention techniques with dropout prediction models is frequently disregarded in current research. It is not enough to just anticipate dropout; in order to improve student retention, customised interventions based on these predictions must be implemented. Future research should examine the efficient integration of prediction models with personalised learning tactics to support students who are considered to be at-risk and improve overall completion rates of courses [113,114]. Therefore, even if a lot has been accomplished in terms of comprehending and forecasting MOOC dropout, more sophisticated, real-time, and adaptable methods are still required. To tackle these obstacles, a blend of enhanced prediction models, enhanced integration with intervention tactics, and a more profound comprehension of student conduct and course dynamics will be necessary. To improve prediction accuracy and intervention tactics, recent developments in MOOC dropout prediction have increasingly used clever techniques. Algorithms such as DT and SVM are two popular ML algorithms that are used to detect at-risk students and forecast dropout rates based on different learning indicators and engagement patterns [114]. Predictive models, for example, can foresee dropout risks and recommend early interventions based on analysis of student interaction data, assignment completion rates, and forum participation. But even with these developments, there are still a number of important research gaps (RGs), as shown in the following sections.

A. RG1: Optimising Feature Selection to Enhance Prediction Model Accuracy

Dropout prediction models often take into account a wide range of factors, such as learning behaviour, engagement indicators, and demographic data. Model performance is greatly impacted by the choice of these features and the preparation of the data. Even though dimensionality reduction and feature selection are used to increase prediction accuracy, the efficacy of these strategies differs greatly throughout MOOC platforms and course types. Issues with data quality and feature relevance plague current models frequently, which negatively affects overall prediction performance. In order to increase the robustness and accuracy of dropout prediction models, future research should concentrate on enhancing feature selection procedures and data pre-processing techniques.

B. RG2: Developing Adaptive Prediction Models for Diverse MOOC Environments

To properly customise interventions, students with comparable traits and learning styles are grouped together using clustering and segmentation techniques. Many models now in use, however, fall short in addressing the context-specific elements and diversity of MOOC contexts, including learner background, course difficulty, and content type. Advanced models that can adapt to different contexts and learning environments, such as ensemble learning and deep learning approaches, show promise, but they often need significant computational resources and ongoing fine-tuning. In order to provide more precise forecasts and successful dropout prevention tactics, research should focus on the creation of adaptable and context-aware algorithms that can dynamically adapt to various MOOC venues and learner profiles.

Improving the accuracy and computing efficiency of prediction algorithms is necessary to optimise dropout prediction in MOOCs, given the dynamic and heterogeneous character of online learning environments. The capacity of dropout prediction algorithms to handle and evaluate large volumes of learner data in real-time, allowing for targeted and timely interventions, is a critical component of their efficacy. The development of sophisticated computational methods that can handle big datasets effectively and retain high prediction accuracy is necessary for this optimisation. The effectiveness of computational algorithms is critical for real-time applications because quick identification of at-risk pupils is essential for putting retention tactics into action. To increase dropout prediction systems' responsiveness, methods including dimensionality reduction, effective data processing, and algorithmic optimisation are crucial. Furthermore, the accuracy of predictions and intervention tactics will be improved by integrating adaptive learning algorithms that can dynamically adapt to changing learner behaviours and course circumstances.

7. Study Limitations and Future Directions

This comprehensive review provides valuable insights into MOOC dropout prediction models, but it has several limitations. The scope was confined to specific databases and MOOC-focused studies, potentially overlooking relevant research from other sources. The variability in methodologies, datasets, and models complicates consistent conclusions. While emphasising the importance of feature selection, this study identifies challenges such as integrating real-time data, managing diverse learner behaviours, and ensuring computational efficiency for large-scale MOOCs. In addition, it struggles to connect predictions to hybrid approaches and address scalability issues in advanced models, such as deep learning.

With MOOCs growing in complexity and size, it has become clear that the dropout prediction optimisation aspect is important and has a lot of potential to grow in the future. The issue comes in developing prediction algorithms to maximise accuracy and efficiency, ensuring early interventions that can improve learner retention and success. By offering flexible answers to the ever-changing problems MOOCs encounter, recent advances in ML have demonstrated promise in this field. Large-scale learner data can be analysed using ML/DL algorithms to find trends and more accurately forecast the likelihood of dropout [115]. These developments make it possible to create intelligent models that adjust to changing student behaviours and course dynamics, which enhances prediction accuracy and intervention techniques. A number of important considerations need to be made while creating dropout prediction models:

- Accuracy and efficiency: To manage massive amounts of data effectively, the model must be able to predict dropout risks with accuracy while consuming the least amount of computational resources.
- Adaptability: The model's adaptability is crucial as it needs to accommodate a range of learning contexts and student profiles. It must also account for differences in course content, learner engagement, and interaction patterns.
- Real-time capabilities: If dropouts need to be avoided before they happen, prompt prediction and action are essential.
- Effectiveness: Over time, the strategy should help maintain gains in student achievement and retention.

Even with this advancement, a lot of research has not adequately addressed all the factors, like varied student engagement levels, course complexity, and the effects of diverse teaching strategies, which influence the prediction of dropout rates. There is currently little research being conducted on advanced algorithms such as swarm intelligence and bioinspired techniques [116]. While these methods have the potential to yield more intelligent answers, more research is necessary to adapt them to a variety of MOOC scenarios.

Meta-heuristic techniques like swarm intelligence (SI) and bio-inspired (BI) techniques offer potential directions for future research, being intelligent means of improving dropout prediction in MOOCs. These approaches, which have their roots in collective behaviours and natural systems, have special benefits for dealing with the complexity and dynamic nature of online learning environments [116,117]. The effective contributions of several BI and SI approaches to dropout prediction optimisation are reviewed in the following paragraphs. As an example, Algorithm 1 describes the processes related to optimising the

MOOC dropout prediction based on optimised feature selection using hybrid ML/DL and meta-heuristic approaches.

Algorithm 1: MOOC dropout prediction optimisation using ML/DL and meta-heuristic approaches.

1. ***initialise processes***
2. ***data collection from dataset***
3. collect (user_interactions, course_progress, demographic_data, engagement_metrics, etc.)
4. ***data preprocessing***
5. clean_Data(), remove_noisy_values()
6. handle_class_imbalance() # use imbalanced data learning techniques
7. encode_categorical_variables(), normalise_or_scale_features()
8. ***feature extraction***
9. extract_key_features(), identify_temporal_patterns()
10. ***feature selection optimisation***
11. initialise_metaheuristic_algorithms()
12. select_algorithm (PSO, ACO, GA, FA, etc.)
13. optimise_dropout_features()
14. ensemble_metaheuristics(select_and combine_algorithms())
15. ***dropout prediction model using ML/DL approaches***
16. train_approaches() # e.g., RF, SVM, NN, etc.
17. fine_tune_Model_via_ensemble()
18. **repeat** step **15**
19. apply_metaheuristics())
20. apply_LSTM()
21. capture_temporal_dependencies()
22. predict_dropout_based_on_sequence()
23. ***model evaluation***
24. evaluate_model(accuracy, F1-score, precision, recall, . . . etc.)
25. feedback_loop() # Incorporate feedback from dropout/persistent users
26. ***End***

Bio-inspired (BI) approaches, such as, ACO, GA, and PSO algorithms, offer effective solutions for dropout prediction. ACO-based models use ant foraging behaviour to efficiently analyse learner paths and inform intervention strategies for dropout risk. GA-based models optimise dropout prediction through evolutionary processes, while PSO models draw inspiration from the social behaviours of fish and birds to enhance prediction accuracy. Future research can explore how these algorithms can be modified to consider MOOC participant behaviours and changing course material [117].

The BI approaches with ML/DL models incorporate different prediction elements and parameters. By optimising parameters and more thoroughly examining the solution space, these hybrid methodology can be used to improve dropout prediction models' accuracy. Other SI techniques, like ABC and Firefly, can optimise feature selection for dropout prediction models. Future studies could explore how these algorithms can adapt to large-scale data in MOOCs and merge them with deep learning methodologies to tackle complex learner data correlations.

Hybrid approaches combining SI and BI methodologies with ML/DL techniques can enhance dropout prediction. These models are more resilient and flexible, allowing for real-time data management and adaptability to student behaviour. Future research should focus on creating BI and SI methodologies that can adapt to new information and expand to large student populations. Thus, more in-depth studies of these methods should be conducted in the future, with an emphasis on how they integrate with DL/ML, how flexible they are in real time, and how scalable they are [118]. By leveraging the natural principles underlying these approaches, researchers can develop innovative solutions to improve learner retention and success in MOOCs.

8. Conclusions

Predicting student dropout rates effectively is becoming more and more crucial as the area of education develops, especially in online learning contexts. To effectively customise interventions, educators and institutions use a range of variables, including academic achievement, engagement measures, and demographic data. Modern intelligence methods, especially data analytics and machine learning, have greatly improved dropout prediction models' accuracy and dependability. We can use these technologies to analyse large volumes of data and find trends and patterns that can point to a student's likelihood of disengaging. Furthermore, integrating both swarm intelligence and bio-inspired methodologies can lead to more-sophisticated predictive models. This combination not only enhances the accuracy of predictions but also allows for more targeted and timely interventions, ultimately supporting student success.

Despite the advancements in recent research, there are still difficulties in smoothly combining different predictors and modifying models to account for real-time variations in learner behaviour. Future research must concentrate on creating hybrid models that achieve a compromise between complexity, scalability, and adaptability because online education is dynamic and requires models that can change as student interactions do. Furthermore, continuous assessments and evaluations are essential to improve MOOC dropout prediction techniques. By enabling persistent methodological improvement, these assessments will guarantee that interventions are both successful and sensitive to the needs of a varied group of students. We can greatly improve student achievement and retention by utilising cutting-edge technologies and creative strategies, opening the door to a more effective and interesting online learning environment.

Author Contributions: Conceptualisation, S.A. and B.S.; methodology, S.A.; software, S.A. and B.S.; validation, S.A., B.S. and A.L.; formal analysis, S.A.; investigation, S.A. and B.S.; resources, S.A., B.S. and A.L.; data curation, S.A.; writing—original draft preparation, S.A.; writing—review and editing, B.S. and A.L.; visualisation, S.A., B.S. and A.L.; supervision, B.S. and A.L.; project administration, B.S. and A.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

Abbreviations

The following abbreviations are used in this manuscript:

AdaBoost	Adaptive Boosting Learning
CART	Classification and Regression Tree Algorithm
CGDC	Combining Group Convolution and Dilated Causal Convolution
DFS	Deep Feature Synthesis
ETC	Extra Tree Classifier
Extra	Extremely Randomised Trees Algorithm
FIAR	Feature Importance Association Rule
FLPADPM	Federated Learning Pattern Aware Dropout Prediction Model
GBC	Gradient Boosting Classifier
GBL	Gradient Boosting Learning
GBDT	Gradient Boosting Decision Tree
GNB	Gaussian Naïve Bayes
GTB	Gradient Tree Boosting
IQPSO	Intelligent Quantum Particle Swarm Optimisation
KNN	k-nearest Neighbours Technique
LDA	Linear Discriminant Analysis

LightGBM	Light Gradient Boosting Decision Tree Implementation
LSTM	Long Short-Term Memory Network
MCC	Matthews Correlation Coefficient
MLP	Multilayer Perceptron
OULAD	Open University Learning Analytics Dataset
PCA	Principal Component Analysis
PLSTM	Pulsed Long Short-Term Memory
Ridge	Ridge Classification Method
RS	Rough Sets
SGD	Stochastic Gradient Descent Algorithm
SMOTE	Synthetic Minority Over-sampling Technique
SVR	Support Vector Regression
TCN	Temporal Convolutional Network
XGBoost	Extreme Gradient Boosting
3WD	Three-Way Decisions

References

1. Wang, W.; Zhao, Y.; Wu, Y.J.; Goh, M. Factors of dropout from MOOCs: A bibliometric review. *Libr. Hi Tech* **2023**, *41*, 432–453. [CrossRef]
2. Feng, W.; Tang, J.; Liu, T.X. Understanding Dropouts in MOOCs. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; Volume 33, pp. 517–524.
3. Goopio, J.; Cheung, C. The MOOC dropout phenomenon and retention strategies. *J. Teach. Travel Tour.* **2020**, *21*, 177–197. [CrossRef]
4. Borrella, I.; Caballero-Caballero, S.; Ponce-Cueto, E. Taking action to reduce dropout in MOOCs: Tested interventions. *Comput. Educ.* **2021**, *179*, 104412. [CrossRef]
5. Vitiello, M.; Walk, S.; Chang, V.; Hernandez, R.; Helic, D.; Guetl, C. MOOC Dropouts: A Multi-system Classifier. In *Data Driven Approaches in Digital Education. EC-TEL 2017*; Lavoué, É., Drachsler, H., Verbert, K., Broisin, J., Pérez-Sanagustín, M., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2017; Volume 10474.
6. Li, J.; Li, S.; Zhao, Y.; Guo, L.; Hao, F.; Ren, M.; Li, K. Predicting Dropouts Before Enrollments in MOOCs: An Explainable and Self-Supervised Model. *IEEE Trans. Serv. Comput.* **2023**, *16*, 4154–4167. [CrossRef]
7. Mrhar, K.; Douimi, O.; Abik, M. A Dropout Predictor System in MOOCs Based on Neural Networks. *J. Autom. Mob. Robot. Intell. Syst.* **2021**, *14*, 72–80. [CrossRef]
8. Aldowah, H.; Al-Samarraie, H.; Alzahrani, A.I.; Alalwan, N. Factors affecting student dropout in MOOCs: A cause and effect decision-making model. *J. Comput. High. Educ.* **2020**, *32*, 429–454. [CrossRef]
9. Althibyani, H.A. Predicting student success in MOOCs: A comprehensive analysis using machine learning models. *PeerJ Comput. Sci.* **2024**, *10*, e2221. [CrossRef]
10. Li, Y. The factors influencing the continuance intention of MOOCs: The perspective of socio-technical approach. *Educ. Inf. Technol.* **2024**. [CrossRef]
11. Li, J. Construction of Modern Educational Technology MOOC Platform Based on Courseware Resource Storage System. *Int. J. Emerg. Technol. Learn. (iJET)* **2017**, *12*, 105. [CrossRef]
12. Goel, Y.; Goyal, R. On the Effectiveness of Self-Training in MOOC Dropout Prediction. *Open Comput. Sci.* **2020**, *10*, 246–258. [CrossRef]
13. Hafshejani, F.F.; Zeinalipour, H.; Ebner, M.; Fini, A.A.S.; Vajargah, K.F.; Jafari, E. Identifying the Essentials of Corporate MOOCs' Application to the Leading Organizations. *Int. J. Emerg. Technol. Learn. (iJET)* **2023**, *18*, 143–159. [CrossRef]
14. Loh, H.S.; Martins van Jaarsveld, G.; Mesutoglu, C.; Baars, M. Supporting social interactions to improve MOOC participants' learning outcomes: A literature review. *Front. Educ.* **2024**, *9*, 1345205. [CrossRef]
15. Dalipi, F.; Zdravkova, K.; Ahlgren, F. Sentiment Analysis of Students' Feedback in MOOCs: A Systematic Literature Review. *Front. Artif. Intell.* **2021**, *4*, 728708. [CrossRef]
16. Alexiou, A.T.; Sarup, P.; Kumar, A.; Gupta, G.K. MOOCs Theories, Trends, Critics, and Life Sciences Applications: Updates on MOOCs. In *Library and Information Science in the Age of MOOCs*; Kaushik, A., Ed.; IGI Global: Hershey, PA, USA, 2018; pp. 240–251.

17. Mehra, A.; Kant, P. A Comparative Evaluation of MOOCs and Classroom Learning in Engineering and Science in India: A MOOC Policy Assessment. In Proceedings of the European Conference on Education 2024: Official Conference Proceedings, Online, 11–15 July 2024; pp. 719–738.
18. Ardchir, S.; Rachik, A.; Ouassit, Y.; Moulouki, R.; Azzouazi, M. An Efficient Feature Selection Approach for MOOCs Dropout Prediction. In *Advanced Intelligent Systems for Sustainable Development (AI2SD'2020). AI2SD 2020*; Kacprzyk, J., Balas, V.E., Ezziyyani, M., Eds.; Advances in Intelligent Systems and Computing; Springer: Cham, Switzerland, 2022; Volume 1417.
19. Youssef, M.; Mohammed, S.; Hamada, E.K.; Wafaa, B.F. A predictive approach based on efficient feature selection and learning algorithms' competition: Case of learners' dropout in MOOCs. *Educ. Inf. Technol.* **2019**, *24*, 3591–3618. [CrossRef]
20. Mansouri, N.; Soui, M.; Abed, M. SFS feature selection with decision tree classifier for massive open online courses (MOOCs) recommendation. *J. Comput. Educ.* **2023**, *11*, 1089–1110. [CrossRef]
21. Bahi, H.A.A.; Boussaha, K.; Laboudi, Z. Meta-features Based Architecture for the Automatic Selection of Prediction Models for MOOCs. In *Novel & Intelligent Digital Systems: Proceedings of the 3rd International Conference (NiDS 2023). NiDS 2023*; Kabassi, K., Mylonas, P., Caro, J., Eds.; Lecture Notes in Networks and Systems; Springer: Cham, Switzerland, 2023; Volume 784.
22. Ige, O.P.; Gan, K.H. Ensemble filter-wrapper text feature selection methods for text classification. *Comput. Model. Eng. Sci.* **2024**, *141*, 1847–1865. [CrossRef]
23. Liu, L.; Wan, L. Innovative models for enhanced student adaptability and performance in educational environments. *PLoS ONE* **2024**, *19*, e0307221. [CrossRef]
24. Ardchir, S.; Ouassit, Y.; Ounacer, S.; Jihal, H.; Goumari, Y.E.L.; Azouazi, M. Improving Prediction of MOOCs Student Dropout Using a Feature Engineering Approach. In *Advanced Intelligent Systems for Sustainable Development (AI2SD'2019). AI2SD 2019*; Ezziyyani, M., Ed.; Advances in Intelligent Systems and Computing; Springer: Cham, Switzerland, 2020; Volume 1102.
25. Gligorea, I.; Cioca, M.; Oancea, R.; Gorski, A.-T.; Gorski, H.; Tudorache, P. Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review. *Educ. Sci.* **2023**, *13*, 1216. [CrossRef]
26. Chi, Z.; Zhang, S.; Shi, L. Analysis and Prediction of MOOC Learners' Dropout Behaviour. *Appl. Sci.* **2023**, *13*, 1068. [CrossRef]
27. Márquez-Vera, C.; Cano, A.; Romero, C.; Noaman, A.Y.M.; Mousa Fardoun, H.; Ventura, S. Early dropout prediction using data mining: A case study with high school students. *Expert Syst.* **2016**, *33*, 107–124. [CrossRef]
28. Papadimitriou, A. Adaptive and Intelligent MOOCs: How They Contribute to the Improvement of the MOOCs' Effectiveness. *Int. J. Res. e-Learn.* **2023**, *9*, 1–21. [CrossRef]
29. Wu, F.; Zhang, J.; Shi, Y.; Yang, X.; Song, W.; Peng, Z. Predicting MOOCs Dropout with a Deep Model. In *Web Information Systems Engineering—WISE 2020*; Huang, Z., Beek, W., Wang, H., Zhou, R., Zhang, Y., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2020; Volume 12343.
30. Deeva, G.; De Smedt, J.; De Weerdt, J. Educational Sequence Mining for Dropout Prediction in MOOCs: Model Building, Evaluation, and Benchmarking. *IEEE Trans. Learn. Technol.* **2022**, *15*, 720–735. [CrossRef]
31. Yang, F.; Xu, Z.; Wang, H.; Sun, L.; Zhai, M.; Zhang, J. A hybrid feature selection algorithm combining information gain and grouping particle swarm optimization for cancer diagnosis. *PLoS ONE* **2024**, *19*, e0290332. [CrossRef]
32. Asli, M.F.; Hamzah, M.; Ibrahim, A.A.A.; Ayub, E. Problem characterization for visual analytics in MOOC learner's support monitoring: A case of Malaysian MOOC. *Heliyon* **2020**, *6*, e05733. [CrossRef] [PubMed]
33. Khalid, A.; Lundqvist, K.; Yates, A.; Ghzanfar, M.A. Novel online Recommendation algorithm for Massive Open Online Courses (NoR-MOOCs). *PLoS ONE* **2021**, *16*, e0245485. [CrossRef]
34. Chen, J.; Fang, B.; Zhang, H.; Xue, X. A systematic review for MOOC dropout prediction from the perspective of machine learning. *Interact. Learn. Environ.* **2022**, *32*, 1642–1655. [CrossRef]
35. Dalipi, F.; Imran, A.S.; Kastrati, Z. MOOC dropout prediction using machine learning techniques: Review and research challenges. In Proceedings of the 2018 IEEE Global Engineering Education Conference (EDUCON), Santa Cruz de Tenerife, Spain, 17–20 April 2018; pp. 1007–1014.
36. Huang, H.; Jew, L.; Qi, D. Take a MOOC and then drop: A systematic review of MOOC engagement pattern and dropout factor. *Heliyon* **2023**, *9*, e15220. [CrossRef]
37. Moreno-Marcos, P.M.; Alario-Hoyos, C.; Munoz-Merino, P.J.; Kloos, C.D. Prediction in MOOCs: A Review and Future Research Directions. *IEEE Trans. Learn. Technol.* **2018**, *12*, 384–401. [CrossRef]
38. Liyanage, N.G.S.; Sandamali, K.A.D.; Fernando, W.M.A.P.S. Systematic review on high dropout rates in moocs—Reasons and solutions. In *Proceeding of the Open University Research Sessions (OURS 2023)*; The Open University of Sri Lanka: Colombo, Sri Lanka, 2023; ISSN 2012-9912.
39. Mduma, N.; Kalegele, K.; Machuve, D. A Survey of Machine Learning Approaches and Techniques for Student Dropout Prediction. *Data Sci. J.* **2019**, *18*, 14. [CrossRef]
40. Sharma, A. A Review on Various Developments in Moocs For Prediction of Dropouts To Form A Basis For Prescriptive Framework To Reduce Dropout Rate. *Think India J.* **2019**, *22*, 372–384.

41. Lin, Y.; Chen, H.; Xia, W.; Lin, F.; Wang, Z.; Liu, Y. A Comprehensive Survey on Deep Learning Techniques in Educational Data Mining. *arXiv* **2024**, arXiv:2309.04761v4.
42. Alhothali, A.; Albsisi, M.; Assalahi, H.; Aldosemani, T. Predicting Student Outcomes in Online Courses Using Machine Learning Techniques: A Review. *Sustainability* **2022**, *14*, 6199. [CrossRef]
43. Guo, T.; Bai, X.; Tian, X.; Firmin, S.; Xia, F. Educational Anomaly Analytics: Features, Methods, and Challenges. *Front. Big Data* **2022**, *4*, 811840. [CrossRef] [PubMed]
44. Khan, L.H.; Haque, M.A.; Ibrahim, E.A.; Mostakim, M.; Hossain, M.I. Interpretable MOOC Dropout Prediction using Different Ensemble Methods and XAI. In Proceedings of the 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 4–6 May 2023; pp. 636–643.
45. Syed Mustapha, S.M.F.D. Predictive Analysis of Students' Learning Performance Using Data Mining Techniques: A Comparative Study of Feature Selection Methods. *Appl. Syst. Innov.* **2023**, *6*, 86. [CrossRef]
46. Del Bonifro, F.; Gabbrielli, M.; Lisanti, G.; Zingaro, S.P. Student Dropout Prediction. In Proceedings of the Artificial Intelligence in Education: 21st International Conference, AIED 2020, Ifrane, Morocco, 6–10 July 2020; Proceedings, Part I; Volume 12163, pp. 129–140.
47. Rabelo, A.M.; Zárate, L.E. A Model for Predicting Dropout of Higher Education Students. *Data Sci. Manag.* **2024**, 1–33. [CrossRef]
48. Mduma, N. Data Balancing Techniques for Predicting Student Dropout Using Machine Learning. *Data* **2023**, *8*, 49. [CrossRef]
49. Ujkani, B.; Minkovska, D.; Stoyanova, L. Application of Logistic Regression Technique for Predicting Student Dropout. In Proceedings of the 2022 XXXI International Scientific Conference Electronics (ET), Sozopol, Bulgaria, 13–15 September 2022; pp. 1–4.
50. Zhang, Y.; Zhang, Q.; Liu, X. Dropout Predictions of Ideological and Political MOOC Learners Based on Big Data. In Proceedings of the WHICEB 2020 Proceedings, Wuhan, China, 5 July 2020.
51. Xing, W.; Du, D. Dropout Prediction in MOOCs: Using Deep Learning for Personalized Intervention. *J. Educ. Comput. Res.* **2019**, *57*, 547–570. [CrossRef]
52. Wang, W.; Yu, H.; Mia, C. Deep Model for Dropout Prediction in MOOCs. In Proceedings of the 2nd International Conference on Crowd Science and Engineering (ICCSE'17), Beijing China, 6–9 July 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 26–32.
53. Kabathova, J.; Drlik, M. Towards Predicting Student's Dropout in University Courses Using Different Machine Learning Techniques. *Appl. Sci.* **2021**, *11*, 3130. [CrossRef]
54. El-Rashidy, M.A.; Farouk, A.; El-Fishawy, N.A.; Aslan, H.K.; Khodeir, N.A. New weighted BERT features and multi-CNN models to enhance the performance of MOOC posts classification. *Neural Comput. Appl.* **2023**, *35*, 18019–18033. [CrossRef]
55. Duru, I.; Sunar, A.S.; White, S.; Diri, B. Deep Learning for Discussion-Based Cross-Domain Performance Prediction of MOOC Learners Grouped by Language on FutureLearn. *Arab. J. Sci. Eng.* **2021**, *46*, 3613–3629. [CrossRef] [PubMed]
56. Altaf, S.; Asad, R.; Ahmad, S.; Ahmed, I.; Abdollahian, M.; Zaindin, M. A Hybrid Framework of Deep Learning Techniques to Predict Online Performance of Learners during COVID-19 Pandemic. *Sustainability* **2023**, *15*, 11731. [CrossRef]
57. Alghamdi, A. Evaluating Factors Influencing Learner Satisfaction in Massive Open Online Course Selection: A Data-Driven Approach Using Machine Learning. *Arab. J. Sci. Eng.* **2024**. [CrossRef]
58. Mandal, L.; Kar, A. Feature Engineering for Predicting Student Dropout in Massive Open Online Courses. *Tuijin Jishu/J. Propuls. Technol.* **2024**, *45*, 1–9.
59. Blundo, C.; Loia, V.; Orciuoli, F. A Time-Aware Approach for MOOC Dropout Prediction Based on Rule Induction and Sequential Three-Way Decisions. *IEEE Access* **2023**, *11*, 113189–113198. [CrossRef]
60. Psathas, G.; Chatzidaki, T.K.; Demetriadis, S.N. Predictive Modeling of Student Dropout in MOOCs and Self-Regulated Learning. *Computers* **2023**, *12*, 194. [CrossRef]
61. Cheng, Y.; Pereira Nunes, B.; Manrique, R. Not Another Hardcoded Solution to the Student Dropout Prediction Problem: A Novel Approach Using Genetic Algorithms for Feature Selection. In *Intelligent Tutoring Systems. ITS 2022*; Crossley, S., Popescu, E., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2020; Volume 13284.
62. Poudyal, S.; Mohammadi-Aragh, M.J.; Ball, J.E. Hybrid Feature Extraction Model to Categorize Student Attention Pattern and Its Relationship with Learning. *Electronics* **2022**, *11*, 1476. [CrossRef]
63. Dass, S.; Gary, K.; Cunningham, J. Predicting Student Dropout in Self-Paced MOOC Course Using Random Forest Model. *Information* **2021**, *12*, 476. [CrossRef]
64. Jin, C. MOOC student dropout prediction model based on learning behaviour features and parameter optimization. *Interact. Learn. Environ.* **2020**, *31*, 714–732. [CrossRef]
65. Mohamad, N.; Ahmad, N.B.; Jawawi, D.N.A.; Hashim, S.Z.M. Feature Engineering for Predicting MOOC Performance. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2020.
66. 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]

67. Mohamed Nafuri, A.F.; Sani, N.S.; Zainudin, N.F.A.; Rahman, A.H.A.; Aliff, M. Clustering Analysis for Classifying Student Academic Performance in Higher Education. *Appl. Sci.* **2022**, *12*, 9467. [CrossRef]
68. Yağcı, M. Educational data mining: Prediction of students' academic performance using machine learning algorithms. *Smart Learn. Environ.* **2022**, *9*, 11. [CrossRef]
69. Xia, X.; Qi, W. Driving STEM learning effectiveness: Dropout prediction and intervention in MOOCs based on one novel behavioural data analysis approach. *Humanit. Soc. Sci. Commun.* **2024**, *11*, 430. [CrossRef]
70. Xia, X.; Qi, W. Dropout prediction and decision feedback supported by multi temporal sequences of learning behaviour in MOOCs. *Int. J. Educ. Technol. High. Educ.* **2023**, *20*, 32. [CrossRef]
71. Xu, C.; Zhu, G.; Ye, J.; Shu, J. Educational Data Mining: Dropout Prediction in XuetangX MOOCs. *Neural Process. Lett.* **2022**, *54*, 2885–2900. [CrossRef]
72. Zhu, X.; Ye, Y.; Zhao, L.; Shen, C. MOOC Behavior Analysis and Academic Performance Prediction Based on Entropy. *Sensors* **2021**, *21*, 6629. [CrossRef] [PubMed]
73. Yin, S.; Lei, L.; Wang, H.; Chen, W. Power of Attention in MOOC Dropout Prediction. *IEEE Access* **2020**, *8*, 202993–203002. [CrossRef]
74. 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; IEEE: New York, NY, USA, 2019.
75. Liao, J.; Tang, J.; Zhao, X. Course Drop-out Prediction on MOOC Platform via Clustering and Tensor Completion. *Tsinghua Sci. Technol.* **2019**, *24*, 412–422. [CrossRef]
76. Basnet, R.B.; Johnson, C.; Doleck, T. Dropout prediction in Moocs using deep learning and machine learning. *Educ. Inf. Technol.* **2022**, *27*, 11499–11513. [CrossRef]
77. Sraidi, S.; Smaili, E.M.; Azzouzi, S.; Charaf, M.E.H. A Neural Network-Based System to Predict Early MOOC Dropout. *Int. J. Eng. Pedagog. (iJEP)* **2022**, *12*, 86–101. [CrossRef]
78. Talebi, K.; Torabi, Z.; Daneshpour, N. Ensemble models based on CNN and LSTM for dropout prediction in MOOC. *Expert Syst. Appl.* **2023**, *235*, 121187. [CrossRef]
79. Liu, H.; Chen, X.; Zhao, F. Learning behaviour feature fused deep learning network model for MOOC dropout prediction. *Educ. Inf. Technol.* **2024**, *29*, 3257–3278. [CrossRef]
80. Patel, K.K.; Amin, K. Predictive modeling of dropout in MOOCs using machine learning techniques. *Sci. Temper* **2024**, *15*, 2199–2206. [CrossRef]
81. Zhang, X.; Wang, X.; Zhao, J.; Zhang, B.; Zhang, F. IC-BTCN: A Deep Learning Model for Dropout Prediction of MOOCs Students. *IEEE Trans. Educ.* **2024**, *67*, 974–982. [CrossRef]
82. Pan, F.; Zhang, H.; Li, X.; Zhang, M.; Ji, Y. Achieving optimal trade-off for student dropout prediction with multi-objective reinforcement learning. *PeerJ Comput. Sci.* **2024**, *10*, e2034. [CrossRef] [PubMed]
83. Doss, A.N.; Krishnan, R.; Karuppasamy, A.D.; Sam, B. Learning Analytics Model for Predictive Analysis of Learners Behaviour for an Indigenous MOOC Platform (Tadakhul System) in Oman. *Int. J. Inf. Educ. Technol.* **2024**, *14*, 961–967.
84. Smaili, E.M.; Daoudi, M.; Oumaira, I.; Azzouzi, S.; Charaf, M.E.H. Towards an Adaptive Learning Model using Optimal Learning Paths to Prevent MOOC Dropout. *Int. J. Eng. Pedagog. (iJEP)* **2023**, *13*, 128–144. [CrossRef]
85. Yujiao, Z.; Ling Weay, A.; Shaomin, S.; Palaniappan, S. Dropout Prediction Model for College Students in MOOCs Based on Weighted Multi-feature and SVM. *J. Inform. Web Eng.* **2023**, *2*, 29–42. [CrossRef]
86. Kumar, G.; Singh, A.; Sharma, A. Ensemble Deep Learning Network Model for Dropout Prediction in MOOCs. *Int. J. Electr. Comput. Eng. Syst.* **2023**, *14*, 2. [CrossRef]
87. Zhang, T.; Liu, H.; Tao, J.; Wang, Y.; Yu, M.; Chen, H.; Yu, G. Enhancing Dropout Prediction in Distributed Educational Data Using Learning Pattern Awareness: A Federated Learning Approach. *Mathematics* **2023**, *11*, 4977. [CrossRef]
88. Shou, Z.; Chen, P.; Wen, H.; Liu, J.; Zhang, H. MOOC Dropout Prediction Based on Multidimensional Time-Series Data. *Math. Probl. Eng.* **2022**, *2022*, 2213292. [CrossRef]
89. Anjali, C.A.; Bai, V.R. An Early Prediction of Dropouts for At-risk Scholars in MOOCs using Deep Learning. In Proceedings of the 2022 International Conference on Computing, Communication, Security and Intelligent Systems (IC3SIS), Kochi, India, 23–25 June 2022; pp. 1–6.
90. 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]
91. 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]
92. Jin, C. Dropout prediction model in MOOC based on clickstream data and student sample weight. *Soft Comput.* **2021**, *25*, 8971–8988. [CrossRef]

93. 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]
94. 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]
95. Qiu, L.; Liu, Y.; Hu, Q.; Liu, Y. Student dropout prediction in massive open online courses by convolutional neural networks. *Soft Comput.* **2019**, *23*, 10287–10301. [CrossRef]
96. 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.
97. Nie, H.; Wen, Y.; Cao, B.; Liang, B. MOOC Dropout Prediction Using Learning Process Model and LightGBM Algorithm. In *Computer Supported Cooperative Work and Social Computing. Chinese CSCW 2023*; Sun, Y., Lu, T., Wang, T., Fan, H., Liu, D., Du, B., Eds.; Communications in Computer and Information Science; Springer: Singapore, 2024; Volume 2012.
98. El Aouifi, H.; El Hajji, M.; Es-Saady, Y. A hybrid approach for early-identification of at-risk dropout students using LSTM-DNN networks. *Educ. Inf. Technol.* **2024**, *29*, 18839–18857. [CrossRef]
99. Jatain, D.; Niranjanamurthy, M.; Dayananda, P. A Hybrid Bio-inspired Fuzzy Feature Selection Approach for Opinion Mining of Learner Comments. *SN Comput. Sci.* **2024**, *5*, 135. [CrossRef]
100. Zhou, Y.; Niu, K.; Lv, H.; Lu, G.; Pan, Y. CGDC-LSTM: A novel hybrid neural network model for MOOC dropout prediction. In Proceedings of the 2023 International Joint Conference on Neural Networks (IJCNN), Gold Coast, Australia, 18–23 June 2023; pp. 1–8.
101. Chen, X.; Liu, H.; Zhao, F. Prediction of MOOCs Dropout Based on IQPSO-PLSTM Model. In Proceedings of the 2023 8th International Conference on Mathematics and Artificial Intelligence (ICMAI '23), Chongqing, China, 7–9 April 2023; Association for Computing Machinery: New York, NY, USA, 2023; pp. 7–12.
102. Zakaria, A.L.J.; Anas, B.; Oucamah, C.M.M. Intelligent System for Personalised Interventions and Early Drop-out Prediction in MOOCs. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* **2022**, *13*, 700–710. [CrossRef]
103. Nithya, S.; Umarani, S. MOOC Dropout Prediction using FIAR-ANN Model based on Learner Behavioural Features. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* **2022**, *13*, 607–617.
104. Liu, H.; Zhang, W. A Hybrid Deep Learning Model for MOOCs Dropout Prediction. In Proceedings of the 2022 4th International Conference on Computer Science and Technologies in Education (CSTE), Xi'an, China, 6–8 May 2022; pp. 178–183.
105. Zhang, Y.; Chang, L.; Liu, T. MooCs dropout prediction based hybrid deep neural network. In Proceedings of the International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), Chongqing, China, 29–30 October 2020.
106. Chen, J.; Feng, J.; Sun, X.; Wu, N.; Yang, Z.; Chen, S. MOOC Dropout Prediction Using a Hybrid Algorithm Based on Decision Tree and Extreme Learning Machine. *Math. Probl. Eng.* **2019**, *2019*, 8404653. [CrossRef]
107. Al-Zawqari, A.; Peumans, D.; Vandersteen, G. A flexible feature selection approach for predicting students' academic performance in online courses. *Comput. Educ. Artif. Intell.* **2022**, *3*, 100103. [CrossRef]
108. El-Sabagh, H.A. Adaptive e-learning environment based on learning styles and its impact on development students' engagement. *Int. J. Educ. Technol. High. Educ.* **2021**, *18*, 53. [CrossRef]
109. Smaili, E.M.; Khoudda, C.; Sraidi, S.; Charaf, M.E.H. An Innovative Approach to Prevent Learners' Dropout from MOOCs using Optimal Personalized Learning Paths: An Online Learning Case Study. *Stat. Optim. Inf. Comput.* **2022**, *10*, 45–58. [CrossRef]
110. Findeisen, S.; Brodsky, A.; Michaelis, C.; Schimmelpenningh, B.; Seifried, J. Dropout intention: A valid predictor of actual dropout? *Empir. Res. Vocat. Educ. Train.* **2024**, *16*, 10. [CrossRef]
111. Imran, A.S.; 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 (ICCAI '19), Bali, Indonesia, 19–22 April 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 190–195.
112. Rahimi, A.R. A tri-phenomenon perspective to mitigate MOOCs' high dropout rates: The role of technical, pedagogical, and contextual factors on language learners' L2 motivational selves, and learning approaches to MOOC. *Smart Learn. Environ.* **2024**, *11*, 11. [CrossRef]
113. Matz, S.C.; Bukow, C.S.; Peters, H.; Deacons, C.; Dinu, A.; Stachl, C. Using machine learning to predict student retention from socio-demographic characteristics and app-based engagement metrics. *Sci. Rep.* **2023**, *13*, 5705.
114. Ahmed, E. Student Performance Prediction Using Machine Learning Algorithms. *Appl. Comput. Intell. Soft Comput.* **2024**, *2024*, 4067721. [CrossRef]
115. Mbunge, E.; Batani, J.; Mafumbate, R.; Gurajena, C.; Fashoto, S.; Rugube, T.; Akinnuwesi, B.; Metfula, A. Predicting Student Dropout in Massive Open Online Courses Using Deep Learning Models—A Systematic Review. In *Cybernetics Perspectives in Systems. CSOC 2022*; Silhavy, R., Ed.; Lecture Notes in Networks and Systems; Springer: Cham, Switzerland, 2022; Volume 503.
116. Li, J.; Xue, E. Dynamic Interaction between Student Learning Behaviour and Learning Environment: Meta-Analysis of Student Engagement and Its Influencing Factors. *Behav. Sci.* **2023**, *13*, 59. [CrossRef] [PubMed]

117. Potharlanka, J.L.; Nirupama Bhat, M. Feature importance feedback with Deep Q process in ensemble-based metaheuristic feature selection algorithms. *Sci. Rep.* **2024**, *14*, 2923. [CrossRef] [PubMed]
118. Goran, R.; Jovanovic, L.; Bacanin, N.; Stankovic, M.; Simic, V.; Antonijevic, M.; Zivkovic, M. Identifying and Understanding Student Dropouts Using Metaheuristic Optimised Classifiers and Explainable Artificial Intelligence Techniques. *IEEE Access* **2024**, *12*, 122377–122400. [CrossRef]