Review

Predicting Student Performance Using Data Mining and Learning Analytics Techniques: A Systematic Literature Review

Abdallah Namoun * and **Abdullah Alshanqiti**

Faculty of Computer and Information Systems, Islamic University of Madinah, Madinah 42351, Saudi Arabia; amma@iu.edu.sa
* Correspondence: a.namoun@iu.edu.sa; Tel.: +966-598009417

Featured Application: **The herein survey is among the first research efforts to synthesize the intelligent models and paradigms applied in education to predict the attainment of student learning outcomes, which represent a proxy for student performance. The survey identifies several key challenges and provides recommendations for future research in the field of educational data mining.**

Abstract: The prediction of student academic performance has drawn considerable attention in education. However, although the learning outcomes are believed to improve learning and teaching, prognosticating the attainment of student outcomes remains underexplored. A decade of research work conducted between 2010 and November 2020 was surveyed to present a fundamental understanding of the intelligent techniques used for the prediction of student performance, where academic success is strictly measured using student learning outcomes. The electronic bibliographic databases searched include ACM, IEEE Xplore, Google Scholar, Science Direct, Scopus, Springer, and Web of Science. Eventually, we synthesized and analyzed a total of 62 relevant papers with a focus on three perspectives, (1) the forms in which the learning outcomes are predicted, (2) the predictive analytics models developed to forecast student learning, and (3) the dominant factors impacting student outcomes. The best practices for conducting systematic literature reviews, e.g., PICO and PRISMA, were applied to synthesize and report the main results. The attainment of learning outcomes was measured mainly as performance class standings (i.e., ranks) and achievement scores (i.e., grades). Regression and supervised machine learning models were frequently employed to classify student performance. Finally, student online learning activities, term assessment grades, and student academic emotions were the most evident predictors of learning outcomes. We conclude the survey by highlighting some major research challenges and suggesting a summary of significant recommendations to motivate future works in this field.

1. Introduction

Student academic performance in higher education (HE) is researched extensively to tackle academic underachievement, increased university dropout rates, graduation delays, among other tenacious challenges [1]. In simple terms, student performance refers to the extent of achieving short-term and long-term goals in education [2]. However, academicians measure student success from different perspectives, ranging from students' final grades, grade point average (GPA), to future job prospects [3]. The literature offers a wealth of computational efforts striving to improve student performance in schools and universities, most notably those driven by data mining and learning analytics techniques [4]. However, confusion still prevails regarding the effectiveness of the existing intelligent techniques and models.

The timely prediction of student performance enables the detection of low performing students, thus, empowering educators to intervene early during the learning process and implement the required interventions. Fruitful interventions include, but are not limited to, student advising, performance progress monitoring, intelligent tutoring systems development, and policymaking [5]. This endeavor is strongly boosted by computational advances in data mining and learning analytics [6]. A recent comprehensive survey highlights that approximately 70% of the reviewed work investigated student performance prediction using student grades and GPAs, while only 10% of the studies inspected the prediction of student achievement using learning outcomes [3]. This gap incited us to thoroughly investigate the work carried out where the learning outcomes are used as a proxy for student academic performance.

Outcome-based education is a paradigm of education that focuses on implementing and accomplishing the so-called learning outcomes [7]. In effect, student learning outcomes are goals that measure the extent to which students attain the intended competencies, specifically knowledge, skills, and values, at the end of a certain learning process. In our view, the student outcomes represent a more holistic metric for judging student academic achievements than mere assessment grades. This view concurs with the claim that the learning outcomes represent critical factors of student academic success [8]. Moreover, renowned HE accreditation organizations, such as ABET and ACBSP, use the learning outcomes as the building blocks for assessing the quality of educational programs [9]. Such importance calls for more research efforts to predict the attainment of learning outcomes, both at the course and program levels.

The lack of systematic surveys investigating the prediction of student performance using student outcomes has motivated us to pursue the objectives of this research. In a systematic literature review (i.e., SLR), a step-by-step protocol is executed to identify, select, and appraise the synthesized studies to answer specific research questions [10,11]. Our systematic survey aims to review the research works conducted in this field between 2010 and 2020 to:

- Deeply understand the intelligent approaches and techniques developed to forecast student learning outcomes, which represent the student academic performance.
- Compare the performance of existing models and techniques on different aspects, including their accuracy, strengths, and weaknesses.
- Specify the dominant predictors (e.g., factors and features) of student learning outcomes based on evidence from the synthesis.
- Identify the research challenges and limitations facing the current intelligent techniques for predicting academic performance using learning outcomes.
- Highlight future research areas to ameliorate the prediction of student performance using learning outcomes.

The remainder of this paper is organized into eight sections. Section 2 presents the foundational concepts of student performance prediction and highlights the surveys conducted in this field regarding their shortcomings. Section 3 outlines the systematic survey methodology that we adopted in this research, as well as the research questions and objectives that we intended to address. Section 4 details the answers to the research questions about the prediction of student performance using learning outcomes. Section 5 discusses the key findings and specifies the limitations. Section 6 proposes several recommendations, while Section 7 defines future research directions.

2. Background and Related Works

This section introduces the basic concepts of student outcomes and student performance, followed by identifying the research gaps in the literature concerning the prediction of student learning outcomes.

2.1. Student Outcomes

Outcome-based education (OBE) has emerged as a new school of thought in education and has recently enjoyed wide acceptability and adoption [12]. This educational paradigm shifts the focus of the teaching and learning process from the traditional teacher objectives to the so-called student outcomes. In simple terms, student outcomes refer to the knowledge, skills, and values to be attained by the students at the time of graduation or at end of a course [13]. The outcomes, representing the targeted competencies, might be defined and measured at the course level, i.e., course outcomes, or program level, i.e., program outcomes. Essentially, course outcomes enable the accomplishment of program outcomes, and their alignment (i.e., courses to program) is performed in a critical activity referred to as curriculum mapping. Computerized tools were developed to assist in realizing the OBE goals [14] and effectively document the educational assessment activities [15]. Their usefulness could be extended by incorporating intelligent models that can prognosticate the attainment of learning outcomes during academic terms.

Measuring student outcomes in higher education indisputably brings about various benefits, including the establishment of program expectations for students and course instructors, the practical assessment of the quality of courses and programs, and the provision of key success indicators of the program, among others [7,16]. Several quality assessment instruments, e.g., [17], and quality assurance frameworks, e.g., [18], were proposed to realize the outcome-based education philosophy and acquire program accreditations. Moreover, the ability to forecast the attainment of student outcomes adds further invaluable advantages, such as the ability to introduce corrective interventions to the learning processes. However, few works surveyed the intelligent prediction of student outcomes. Furthermore, the factors and attributes that impact educational outcomes are still vague. Studies suggest that these factors range between academic factors, e.g., teaching quality [19] and online engagement [20], and non-academic traits, e.g., family engagement [21] and student motivation [22]. In this work, we aim, through a systematic survey, to understand the landscape of student outcomes prediction using data mining and machine learning, identify the main challenges hindering the prediction of student outcomes, and propose relevant recommendations.

2.2. Student Performance

Albeit, due to the substantial educational shift in teaching and learning, i.e., OBE, student performance remains a significant concern in higher education [5], especially given the low grades and increasing dropout rates even at world-class universities [23]. Previous reviews showed that the cumulative GPA and course assessments are the most used predictors of student performance and success [24,25]. Indeed, several studies used next-term course grades as the main indicator of student performance, e.g., [26,27]. However, it is not uncommon to measure student performance in other forms, including dropout rate, student knowledge, post-course outcomes, among other indicators [28]. In our view, student academic performance should not be assessed using assessment grades only. Instead, it should be studied within a broader context, particularly using the student outcomes, which are now guiding the learning process by looking at the cohort performance. Moreover, recent research recommends exploring the prospect of predicting the attainment of student outcomes to infer student performance [29].

The intelligent techniques employed in learning analytics to forecast student achievements are generally categorized into supervised learning, unsupervised learning, data mining, and statistical approaches [3,28]. Each category incorporates a wealth of intelligent algorithms, such as Artificial Neural Networks, Support Vector Machine, K-Nearest Neighbor, and Random Forests. The attributes that predict student performance are surveyed extensively in the literature, leading to a mix of academic (e.g., pre-admission scores and entry qualifications) and non-academic factors (e.g., emotional intelligence and resilience) [30,31]. However, mystery still surrounds the factors that influence the attainment of course and program outcomes.

Measurable student outcomes are developed to improve the quality of learning processes and educational programs [13]. Effectively, these outcomes assess what students can perform with what they have learned. The attainment of learning outcomes, both at the course and program level, is performed using direct and indirect assessment methods at the end of the learning process. The direct assessment methods seek to find tangible evidence demonstrating student learning, while the indirect methods rely on the students' reflections on their learning experience. To calculate the attainment rate of outcomes, one should identify a priori the attainment targets and levels and then properly align student grades to the appropriate attainment level [13]. In our work, we examined the studies that predict the attainment of student outcomes, irrespective of their form.

2.3. Existing Student Performance Reviews and Literature Gaps

Our extensive review of previous surveys revealed that, to the best of our knowledge, no systematic literature survey was carried out focusing on the prediction of student academic performance from the learning outcomes perspective. Table 1 summarizes the prominent surveys carried out on the prediction of student performance and emphasizes their focus and weaknesses. Indeed, our search returned numerous surveys on the use of data mining techniques in education (i.e., EDM) to unravel student modelling activities and predict academic performance. These reviews suffered from several limitations, for they (1) were generally broad, (2) did not focus on using student outcomes as an indicator of student performance, (3) suffered from quality issues (e.g., methodologies not thoroughly defined), and (4) were not published in highly indexed venues. These weaknesses are highlighted in Table 1.

Table 1. Existing surveys on student performance prediction, their weaknesses and strengths.

Focus of Survey and Publication Venue	Type of Survey	Number of Bibliographic Databases Explored; Papers Reviewed	Metric of Student Performance Comparison	Models and Approaches Reviewed	Years Covered	Weakness		Strength
Prediction of student performance using data mining [24]; *Indexed Conference*	Systematic review	Four databases; 30 papers	Prediction accuracy (%)	Data mining techniques	(2002–Jan 2015)	− Did not forecast student outcomes − Reviewed a small number of papers − Did not discuss the limitations	+ +	Identified that attributes and methods that predict student performance Applied the PICO methodology
Prediction of student outcome using data mining [32]; *Symposium*	Systematic review	10 databases; 42 papers	Prediction accuracy (%)	Data mining techniques	Not indicated	− Focused on the Moodle virtual learning system only − Did not address student outcome from different perspectives − Did not discuss the limitations − Did not compare the predictive models	+ 	Formulated well defined questions
Techniques and algorithms used for student performance prediction [5]; *Indexed Conference*	Traditional literature review	One database; 88 papers/projects/reports	Not reported	Education analytics	(2013–2017)	− Did not assess the quality of the studies − Did not compare the models	+	Listed the techniques used in the learning analytics for predicting student performance
Data mining techniques to discover knowledge in education [33]; *Indexed Journal*	Traditional review	Databases not indicated; 240 papers	Not reported	Educational data mining	(2010–first quarter 2013)	− Did not focus on student outcomes Only 46 papers were reported about performance modelling	+	Covered different areas such as student performance modelling and assessment approaches
Performance prediction using data mining techniques [34]; *Unindexed Journal*	Systematic review	Six databases	Prediction accuracy (%)	Data mining techniques	(2007—July 2016)	− Did not survey student outcomes − Reported only five techniques − Did not discuss the limitations − Adopted a weak survey methodology	+	Discussed the factors predicting student performance
Performance prediction using machine learning [35]; *Unindexed Journal*	Literature survey	Not indicated	Prediction accuracy (%)	Machine learning models	Not indicated	− Did not survey student outcomes	+	Compared the performance of machine learning models
Features predicting student performance [3]; *Indexed Journal*	Systematic review	Three databases; 357 papers	Different measures of performance were considered.	Statistical approaches, data mining techniques, machine learning models	(2010–2018)	− Did not discuss student outcomes	+ + +	Adopted a robust methodology Highlighted the predictors of student performance Described the most used prediction methods
Preliminary results of predictive learning analytics [36]; *Indexed Conference*	Systematic review	Databases not indicated; 39 papers	Prediction accuracy (%)	Machine learning models	(2002–2016)	− Presented a preliminary study (2-page long) lacking crucial details − Did not publish the full results of the survey.	+	Reported the overall context of student outcomes

Other less relevant surveys published in the field focused on the effects of homework assignments on student performance [37], the impact of using interactive whiteboards on student achievement [38], the predictors of student success in the first year of study [39], and the factors of graduate success [40]. Unlike the above-mentioned surveys, our research opted to conduct a systematic review by implementing a comprehensive review process that allows synthesizing concrete answers to well-defined research questions, in the context of predicting student learning outcomes.

3. Survey Methodology

This research performed a systematic review where the relevant academic works predicting student performance using learning outcomes were identified, selected, and critically evaluated using several criteria, as presented in the results section. To streamline our contributions, we formulated three key research questions as follows:

- RQ1-Learning Outcomes Prediction. How is student academic performance measured using learning outcomes?
- RQ2-Academic Performance Prediction Approaches. What intelligent models and techniques are devised to forecast student academic performance using learning outcomes?
- RQ3-Academic Performance Predictors. What dominant predictors of student performance using learning outcomes are reported?

The main objective of this survey was to create a comprehensive understanding of the landscape of academic performance prediction by focusing on the attainment of learning outcomes. To answer the above research questions accurately, we adopted the well-founded PICO model [41]. The PICO protocol emphasizes the definition of four key elements, namely population, intervention, comparison, and outcome. Concerning our research, population refers to the investigation of the learning outcomes prediction studies, the intervention refers to the intelligent approaches and factors used to predict the attainment of student outcomes, comparison refers to the performance prediction variability between the surveyed models, and outcome refers to the accuracy of these approaches as well as the predictors of learning outcomes. Table 2 details the PICO elements of our survey.

Table 2. PICO protocol adopted in our survey.

Population/Problem	Intervention	Comparison	Outcome
Studies predicting student performance using the learning outcomes	List of intelligent models and techniques	Comparison across the identified models and techniques	Quality and accuracy of the approaches Set of performance predictors of learning outcomes

Moreover, we applied the best practices for conducting useful systematic reviews [10]. As such, we identified and searched seven major online bibliographic databases, which contain engineering and science publications. These databases include the ACM digital library, IEEE Xplore, Google Scholar, Science Direct, Scopus, Springer, and Web of Science. These are the common databases searched by software engineering reviews and are expected to incorporate the studies investigating the predictive modeling of student outcomes. Other electronic databases, such as DBLP and CiteSeer, were excluded from the search since their results are inclusive within the previous seven databases. Furthermore, the databases that publish non-reviewed articles were ignored. Figure 1 summarizes the general steps of our full systematic review.
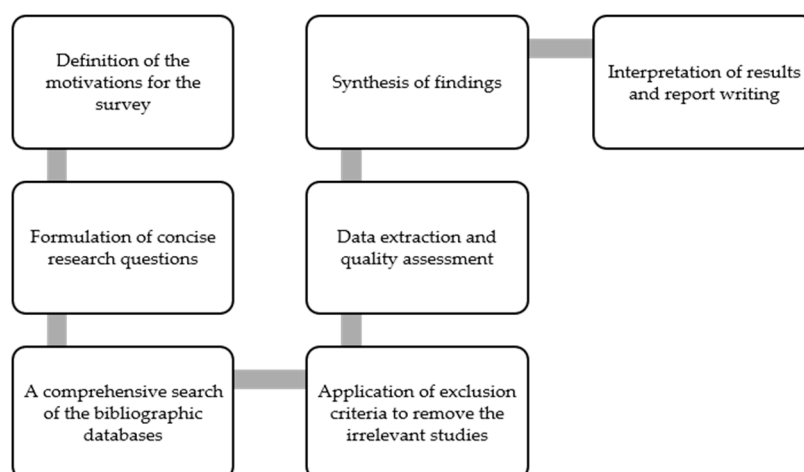
Figure 1. Major steps of our survey methodology.

We conducted the searches in November 2020 using the above databases with a focus on the studies published between 2010 and November 2020. The key terms that were devised to perform the searches were directly linked to the concepts of the research questions and PICO elements.

(predict* OR forecasting)
AND
("student learning outcomes" OR "student outcomes" OR "learning outcomes")
AND
("artificial intelligence" OR "machine learning" OR "data mining" OR "deep learning" OR "learning analytics")

It is worthwhile to note that the search string syntax was trialed multiple times and slightly modified in each database to obtain all relevant results, as recommended in [10]. When the searches were carried out on the full texts of our selected databases, thousands of irrelevant studies were fetched and returned. Therefore, we restricted our searches to the titles, abstracts, and keywords only, yielding a more reasonable pool of studies, as shown in Table 3.

Table 3. Inclusion criteria in our systematic literature review.

Inclusion Criteria	Description of Criteria
I1. Focus of study	Studies that explicitly predict student performance with a direct reference to the learning outcomes
I2. Empirical evidence of prediction	Studies that contain empirical evidence of the performance prediction
I3. Language of publication	Only articles written in English are considered
I4. Year of publication	Studies published between 2010 and 2020 (both years inclusive)
I5. Publication venue	Studies published in peer-reviewed scientific venues (e.g., conference or journal)
I6. Availability of text	Full text is accessible for analysis

3.1. Inclusion Criteria

Table 3 lists the inclusion criteria that were applied to shortlist the candidate articles for consideration in this review. In other words, the studies that did not satisfy the criteria listed below, were disregarded. For instance, non-refereed articles, such as technical reports, and non-English papers were excluded.

To ensure clarity and quality of our methodology, this systematic review followed the four stages advocated by the PRISMA statement and the reporting guidelines [42]. The first phase of PRISMA identifies the potential studies to investigate using automated and manual searches. The screening phase of the studies follows the identification phase to exclude duplicate and irrelevant studies. Next, the qualified articles are thoroughly read and assessed for eligibility, leading to the final set of studies to be included in our synthesis. In the screening and eligibility phases, we strictly applied the inclusion criteria listed in Table 3. Studies that did not directly refer to the prediction of learning outcomes (i.e., the outcome variable) were excluded from the synthesis. Moreover, it is not uncommon to see such a high drop in the number of papers that do not meet the inclusion criteria. Figure 2 shows the PRISMA flow diagram of our survey.
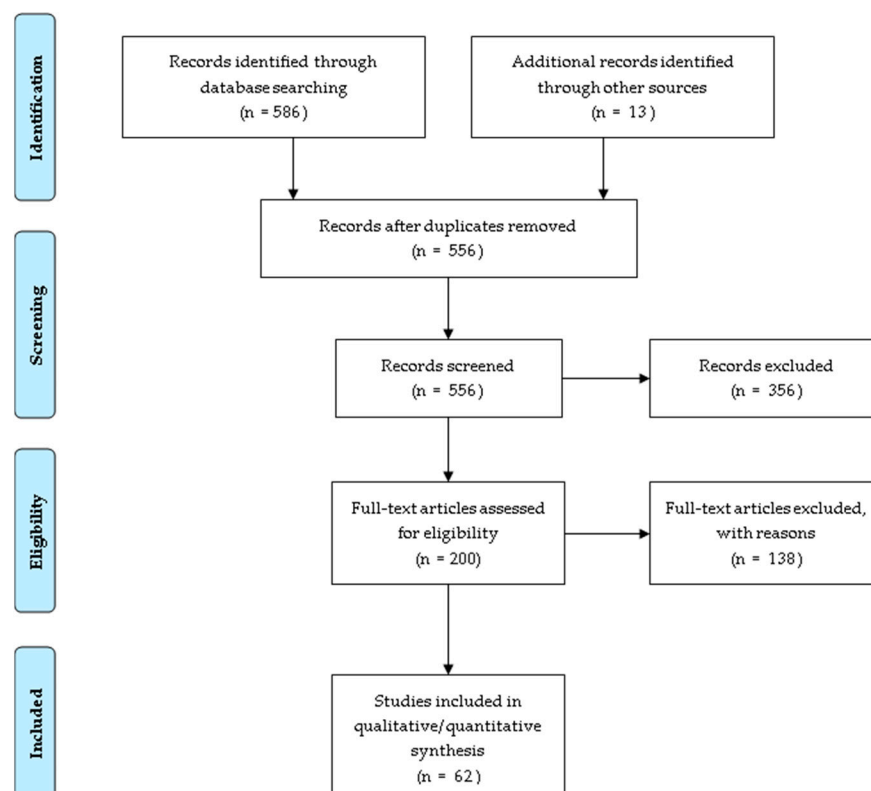

Figure 2. PRISMA flow diagram of our survey methodology.

The initial round of automated searches on the electronic databases gave a corpus containing a total of 586 articles, as listed in Table 4. After removing the duplicate publications and scanning the titles and abstracts, the number was reduced to 187 potentially relevant articles. Full scanning of the eligible articles reduced the search results to 51 relevant articles. Moreover, manual searches were executed by the authors to consider a further 11 primary articles.

To sum up, the automated search yielded 51 relevant articles. However, SLR guidelines suggest carrying out manual searches to overcome the threat of missing primary studies and improve the reliability of the survey [10]. To this end, we (1) hand searched different journal and conference publications and (2) looked into the reference lists of our candidate articles to identify new relevant articles. These manual search approaches gave an additional 11 primary articles. Hence, the final sample of articles judged to be relevant to the prediction of student outcomes using intelligent approaches, e.g., machine learning, amounted to 62 papers.

Table 4. Search results of several rounds from the electronic bibliographic databases.

	ACM	IEEE Xplore	Google Scholar	Science Direct	Scopus	Springer	Web of Science	Round Total
Round 1 (Initial Results)	64	152	65	91	63	115	36	586
Round 2 (Removing Duplicates)	64	148	65	91	28	114	33	543
Round 3 (Scanning the Title and Abstract)	12	63	10	13	53	9	27	187
Round 4 (Reading Full Text)	3	7	5	5	12	2	17	51
Round 5 (Manual Searches)	A further 11 articles were added through manual searches							62

3.2. Data Extraction

Upon applying the PRISMA approach [42], the final pool of selected studies was thoroughly analyzed to extract the data that assist in answering the research questions. The extracted data included:

- General information about the publication, for instance, publication year, venue type, country of publication, and number of authors;
- Educational dataset and context of prediction (e.g., students, courses, school, university, ... etc.);
- Input variables used for student outcomes prediction and the form in which they were predicted;
- Intelligent models and approaches used for the prediction of academic performance;
- Significant predictors of learning outcomes.

We applied thematic analysis to the extracted data to answer RQ1, RQ2, and RQ3. The data were grouped and categorized according to the themes reported in the results section. However, it was not feasible to carry out the meta-analysis of the selected studies, mainly because most educational datasets were either private or not possible to obtain. Below we detail the results of our synthesis analysis.

4. Survey Results

This section reports general information about the surveyed articles, the forms in which the student outcomes were forecasted, the intelligent models developed for performance prediction, and the predictors of student attainment of learning outcomes.

4.1. Publication Venues and Years

A total of 62 studies were analyzed to assist with answering the questions posited in our research. Figure 3 shows that these studies were published in peer-reviewed journal venues (35 studies, 56.45%) and conferences (27 studies, 43.55%). Overall, the publications appeared in four categories, most notably Computing and Engineering (22 studies, 35.48%) and Information Technology and Education (13 studies, 20.96%) venues. The student performance prediction papers also appeared in the Education (17 studies, 27.41%) and Psychology (8 studies, 12.90%) fields of study, as depicted in Figure 4.

It can be seen that the number of studies endeavoring to forecast learning outcomes as an indicator of student success is on a constant rise. Figure 5 shows that the interest in student outcomes prediction models was increasing since 2017, which coincides with the global educational shift towards outcome-based assessment and accreditation efforts. Our search of the databases retrieved articles published until the start of November 2020, which might explain the slight decrease in the number of published articles in the year 2020.
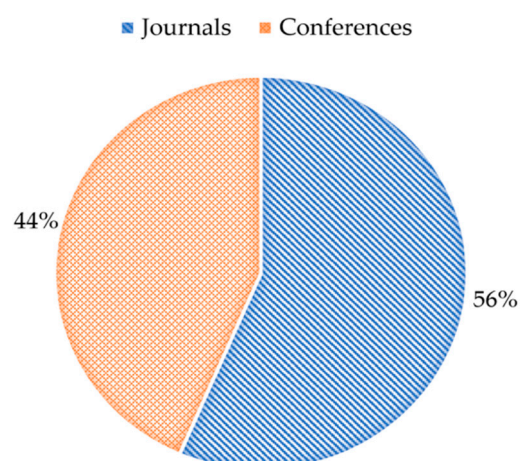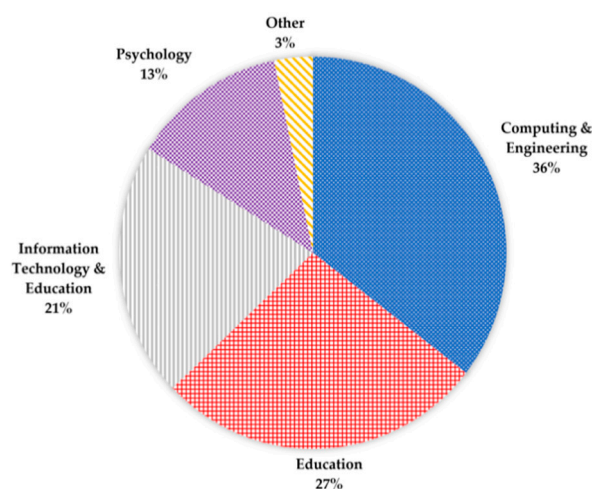
Figure 3. Types of articles reviewed.

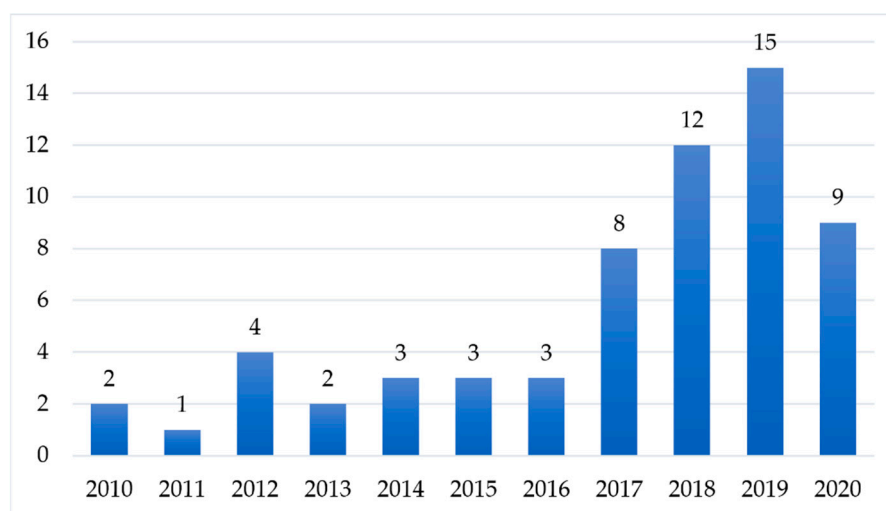**Figure 4.** Category of publication venues.

Figure 5. Number of articles distributed per year of publication, from 2010 to November 2020.

Figure 6 shows that approximately 48.38% (30) of the published predictive models were produced by the efforts of more than three authors. Single authors produced only 8% (5) of the studies.
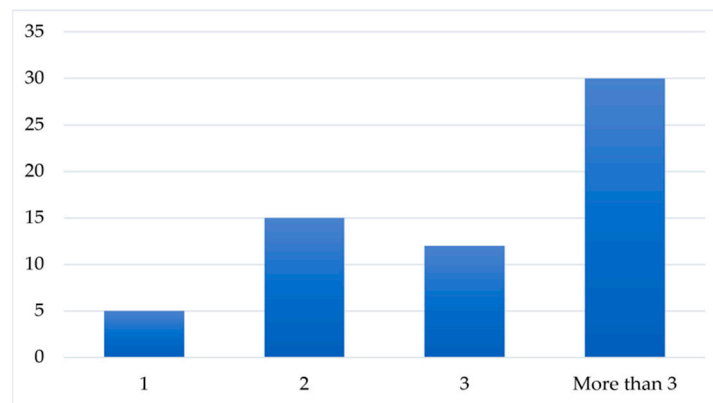
Figure 6. Frequency of articles based on the number of authors in each publication.

4.2. Experimental Datasets and the Context of Performance Prediction

All selected studies reported using at least one educational dataset to test its prediction model or understand the factors influencing the attainment of student outcomes. Thirty-two studies (51.6%) reported collecting performance data from traditional classroom learning, 21 studies (33.8%) from virtual learning environments, and nine studies (14.5%) from blended learning environments (e.g., a mix of online and face-to-face learning activities). The performance prediction models were applied to the university (72.58%), school (25.81%), and kindergarten data (1.61%). When we explored the type of degrees pursued by the students whose performance was being predicted, 43 (69.35%) studies examined the performance of undergraduate university students, and seven (11.29%) studies investigated the performance of high school students. Only two studies (3.22%) reported examining the performance of postgraduate students [43,44].

When we looked at the context of prediction, the datasets and prediction models were applied mainly to courses in the natural sciences field (i.e., STEM) (33 studies, 53.22%). Figure 7 shows less emphasis on the courses belonging to the social sciences field (8 studies, 12.90%). More precisely, learning outcomes predictions were developed for Computer Science (13 studies, 20.96%), Mathematics (5 studies, 8.06%), and Engineering majors (4 studies, 6.45%), as depicted in Figure 8.
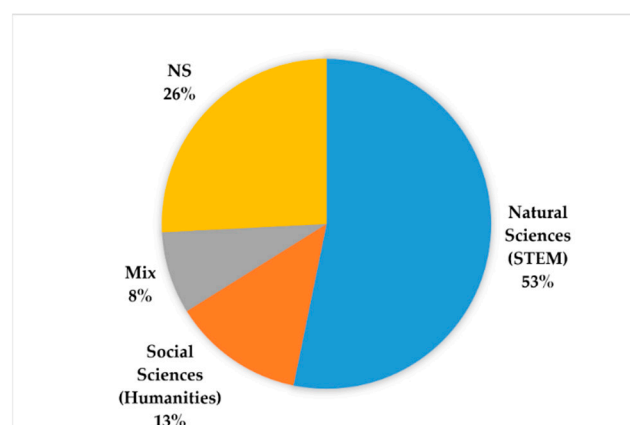

Figure 7. Discipline of academic performance predictions using learning outcomes.
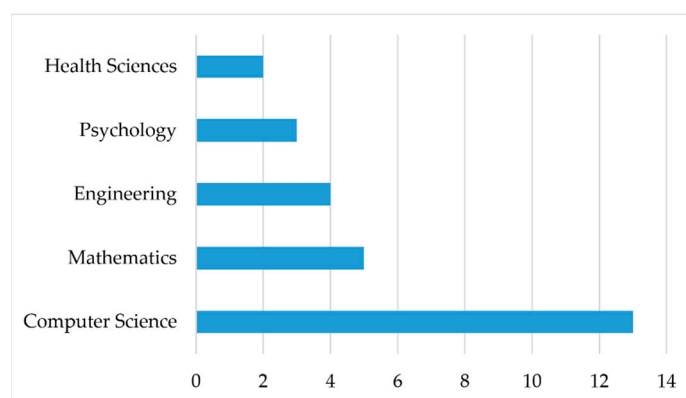
Figure 8. Frequency of prediction studies per type of major.

Since the studies were performed in 23 different countries, we chose to cluster them per continent, as shown in Figure 9. Twenty-five (40%) studies took place in the USA alone, followed by Europe (18 studies, 29%) and Asia (13 studies, 20%). Moreover, the training datasets for 59 studies were collected from one country only. However, two studies collected their data from students enrolled in more than one country [45,46].
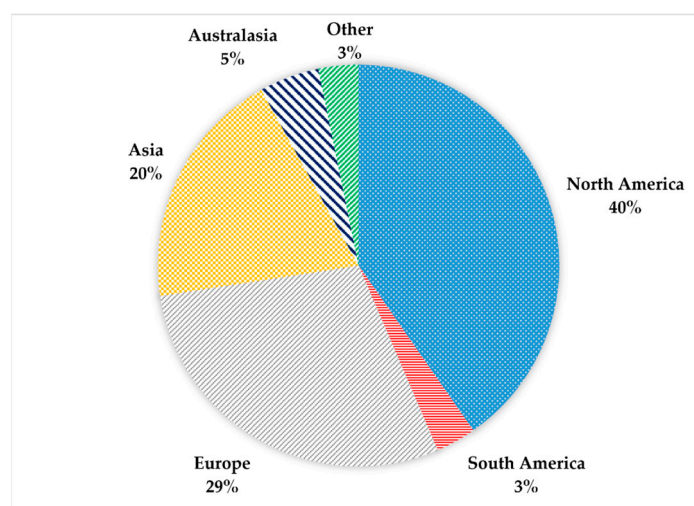

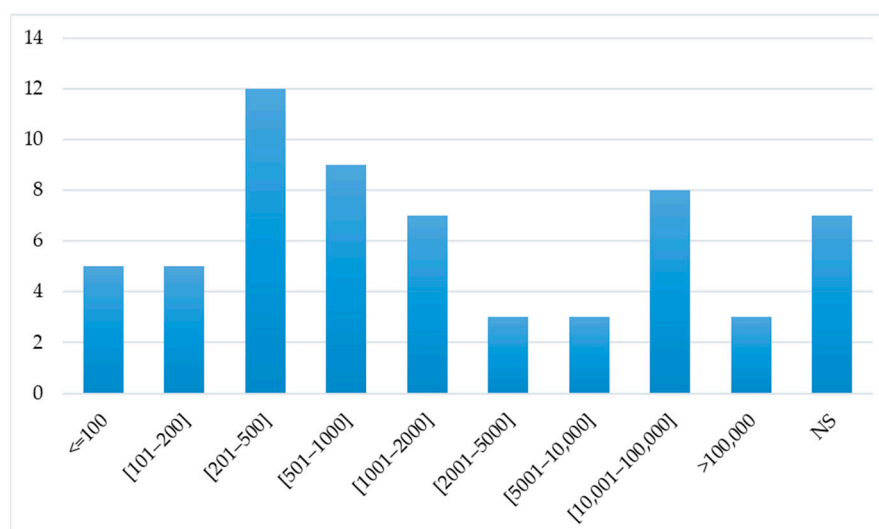**Figure 9.** Distribution of studies across the continents.

All models attempted to predict academic outcomes except for one [47], which predicted academic and non-academic outcomes. The non-academic outcomes were measured using students' self-reports of self-esteem, satisfaction with life, and sense of meaning. Thirty-seven (59.67%) experimental datasets were collected from the same environment (i.e., a single school or university). However, there are a few studies that expanded their data collection activities to multiple schools or universities within the same district, e.g., ref. [48] collected student data from 750 schools, [49] from 113 schools, and [45] from 5 universities. Generally, it can be observed that the studies investigating student performance in schools collected their educational data from multiple schools, as reported in Table 5. In contrast, studies investigating academic performance in higher education employed data from a single university (36 articles, 58.06%). However, 12 articles did not specify the number of schools or universities involved in the data collection process.

Table 5. Source of educational dataset in the studies.

Source	Number	Number of Studies (Percentage of Occurrence)	Studies
School	One	1 (1.61%)	[50]
	Multiple	11 (17.74%)	[47–49,51–58]
University	One	36 (58.06%)	[42,44,59–92]
	Multiple	2 (3.22%)	[45,93]
Not Specified		12 (19.35%)	[46,94–104]

When we inspected the number of courses from which the experimental data were drawn, we discovered that ten studies used a single course/subject, eight studies used two courses, and three studies used four courses, amounting to 35.48% of the surveyed articles. Moreover, 18 (29.03%) studies used between four courses (i.e., [80]) and 270 courses (i.e., [70]) to test the correctness of their predictive models. Nonetheless, it was unclear how many courses were used in the remaining 22 (35.48%) studies.

The experimental datasets included performance data about as little as less than 1000 students (50% of the surveyed studies). Figure 10 shows that the number of studies using datasets including 1001 to 10,000 students, amounts to 13 (20.96%) articles. Overall, the studies that included data points of more than 10,000 students amounted to 11 (17.74%), three of which used a sample size greater than 100,000 students (i.e., [78,85,102]). The remaining seven (11.29%) studies did not specify the sample size of their student dataset. When we inspected the models' prediction accuracy based on the size of the dataset, we found varied results. For example, datasets containing less than 100 students gave weak predictions (e.g., 83 students resulted in an accuracy = (48–100%) [69]; 134 students resulted in an accuracy = 81.3% [64]) and acceptable predictions (e.g., 100 students resulted in an accuracy = 90%, recall = 90%, and precision = 74% [59]). Similarly, datasets of more than 100,000 students gave mixed findings. For instance, a pool of 597,692 students gave an impressive accuracy = 98.81%, AUC = 99.73%, sensitivity = 98.46%, and specificity = 99.20% [85]. However, a sample of 130,000 students gave an accuracy = 48–55%, RMSE = 8.65–10.00, and MAE = 6.09–7.74 [78]. Similarly, a sample of 142,438 students gave an RMSE = 0.34 and AUC = 0.81 [102].

Figure 10. Frequency of studies based on the number of students in the experimental datasets.

It is worth pointing out here that classifying the training datasets into two pools, small or sufficient sample size, to draw meaningful conclusions about the significance

of the results is no simple matter. This is because such a division is influenced by several intertwined factors, including the diversity of input features impacting the outcome variable, the tolerance for errors, and the type of prediction (e.g., statistical analysis or learning) implemented. Moreover, comparing the performance of models that were trained on differing datasets (i.e., characteristics and sizes) might not be conclusive. There is no disagreement that larger the sample size we have to train the predictive models, the stronger is the predictions we obtain. However, this was not evident from our analysis of the synthesis.

4.3. Learning Outcomes as Indicators of Student Performance

As stated earlier, we considered only the research articles that predicted student outcomes as a representative of student performance and success. It is worth noting that the articles that define the learning outcomes as their outcome variable were considered in our analysis, irrespective of the form of the learning outcomes. However, other studies that referred to student academic achievements using in-class assessment metrics, such as GPA or grades, without any reference to the learning outcomes were excluded from the survey. Overall, 56 (90.32%) studies attempted to forecast course outcomes, while three studies looked at the feasibility of predicting program outcomes. Only two studies, (i.e., [103,104]), calculated student performance at both the course and program levels. Furthermore, most of the predictive models estimated the learning attainment of students individually (55 studies, 88.70%) rather than collectively (i.e., cohorts of students) (4 studies, 6.45%). However, three studies (i.e., [52,77,94]) predicted the performance of individual students as well as student cohorts.

The prediction of student performance was achieved in two ways, formative and summative. In the formative prediction of learning outcomes, student features are considered throughout different points of the academic semester, in a bid to inform the instructors about the expected achievements of their students. This formative prediction empowers instructors to implement the necessary interventions early enough in the course. However, in the summative prediction, the learning outcomes are predicted at the end of the semester. Thirty-eight (61.29%) models provided summative predictions, while 19 (30.64%) models provided formative predictions (e.g., weekly or monthly) of student performance. Only five studies calculated both formative and summative predictions of student performance, i.e., [46,68,76,84,88].

Typically, the attainment of student outcomes, whether at the course or program level, might be assessed and measured via direct or indirect methods. The direct methods use various types of course-level assessments, such as assignments and examinations, to obtain insights about student achievements. However, the indirect methods of assessment depend mainly on student opinions and feedback about their learning experiences. In our survey, most studies (50 studies, 80.64%) predicted student learning using direct measures. Nine (14.51%) studies used self-reports of students on their learning experience (i.e., indirect measures) to predict their performance. However, three notable studies [56,57,97] provided learning outcomes predictions using direct and indirect assessment.

Furthermore, we inspected the form in which the learning outcomes were forecasted in the surveyed studies. Table 6 shows the results of the thematic analysis, revealing six distinctive types. The learning outcomes were predicted mostly in the form of performance classes (34 occurrences), achievement scores (20 occurrences), perceived competence (5 occurrences), self-reports of educational aspects (3 occurrences), and failure/graduation rates (3 occurrences).

Figure 11 depicts that 80% of the models predicting academic performance standings classified the outcomes into two to four classes. The remaining 20% of the models forecasted more than 4 class labels of learning outcome performance. Examples of binary (dichotomous) classes are 'pass' and 'fail' [86,87], 'certification' and 'no certification' [85], and 'on-time graduation' and 'not on-time graduation' [60]. A 4-class outcome example predicted students with variable risks [101], e.g., high risk (HR), medium risk (MR), low

risk (LR), and no risk (NR). Ordinal performance ranks were also predicted; for instance, student outcomes were classified into five performance ranks, specifically fail, satisfactory, good, very good, and excellent [93].

Table 6. Distribution of studies based on the type of learning outcomes predicted.

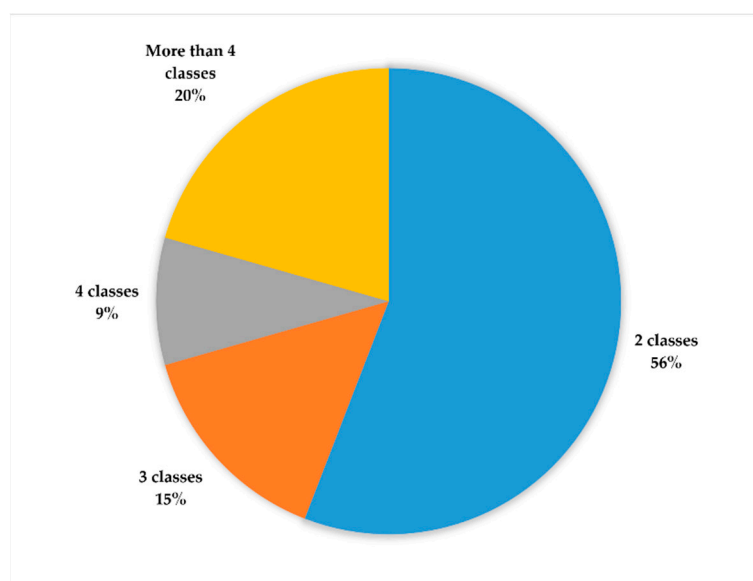| Learning Outcome Type	Number of Occurrences	Studies
Performance classes (Categorical; Binary, nominal and ordinal)	34	[45,46,55,59–70,73–77,79,80,83–89,91,93,94,100,101,103]
Achievement/grade scores (Continuous; Interval)	20	[43,44,50–57,78,81,82,84,90,92,96–99,102]
Perceived competence and achievements (Continuous; Interval scale)	5	[47,57,71,72,97]
Self-reports about educational aspects (Continuous; Interval scale)	3	[56,81,97]
Failure/dropout/graduation rates (Continuous; Ratio)	3	[48,49,95]
Other (e.g., college enrollment, careers, time to graduate, attendance, . . . etc.)	6	[45,48,49,56,58,66]
NS	1	[104]

Figure 11. Distribution of student performance class labels predicted in the models predicting learning outcome as standings (i.e., ranks).

4.4. Predictive Models of Learning Outcomes

In learning analytics, predictive modeling focuses primarily on improving the accuracy of student performance predictions. In contrast, explanatory modeling focuses on identifying and explaining the factors that lead to the predicted achievements of students [105]. The intelligent models suggested for predicting learning outcomes were mainly predictive in nature (52 studies, 87.09%), with only ten models (16.12%) trying to explain the predictions by linking them to the exact features leading to the observed performance, i.e., [46–48,62,77,82,85,87,95,102].

Fifty-four (87.70%) studies employed single intelligent models for predicting the attainment of learning outcomes. Remarkably, only eight studies (i.e., [60,65,66,80,84,93,96,101]) explored the use of hybrid intelligent models to improve the accuracy of academic perfor-

mance predictions. Hybrid or ensemble classifiers involve the integration of heterogeneous learning techniques to boost the predictive performance [106].

Table 7 categorizes the 62 articles according to the intelligent learning genre they implemented to predict academic performance. Overall, five types of predictive analytics emerged with statistical models appearing the most (28 studies, 45.61%), followed by supervised learning models (25 studies, 40.32%). The use of unsupervised learning alone appeared only in one study [67].

Table 7. Distribution of student outcomes predictive models per learning type.

Learning Type	Number of Studies (Percentage of Occurrence (%))	Studies
Statistical analysis	28 (45.16%)	[43–45,47–50,52–54,56–58,64,71–73,75,76,81–83,90–92,97–99]
Supervised machine learning	25 (40.32%)	[46,51,55,59–62,65,68–70,74,78,80,84–89,95,96,101–103]
Data mining	5 (8.06%)	[63,77,94,100,104]
Supervised and unsupervised learning	3 (4.83%)	[66,79,93]
Unsupervised machine learning	1 (1.61%)	[67]

We delved into the types of intelligent methods and algorithms that are used to forecast the attainment of student outcomes, and clustered the proposed models into six categories. Table 8 shows that regression analysis was the most frequently used (51.61%) prediction techniques. Artificial neural networks and tree-based models came into the second position, making together a total of 29.02%. Bayesian approaches made only 8% of the predictive models. Notably, support vector machines were employed in two studies (i.e., [59,60]).

Table 8. Distribution of intelligent predictive algorithms per category.

Learning Model	Number of Studies (Percentage of Occurrence (%))	Studies
Statistical models (Correlation and Regression)	32 (51.61%)	[43–47,49,50,52–54,56–58,64,67,69,71–73,76,81–83,88–92,97–99]
Neural networks	9 (14.51%)	[51,68,70,74,78,86,95,96,102]
Tree-based models (Decision trees)	9 (14.51%)	[55,63,65,66,75,77,85,101,104]
Bayesian-based models	5 (8.06%)	[61,62,79,93,94]
Support Vector Machines	2 (3.22%)	[59,60]
Instance-based models	1 (1.62%)	[103]
Other	4 (6.45%)	[48,80,84,100]

The literature postulates several metrics to evaluate the proposed machine learning models [2,3,5]. Accuracy is probably the most popular metric employed to judge the effectiveness of a predictive model. Accuracy refers to the ratio between the correctly predicted outputs over the total predictions. Other evaluation metrics include precision, recall, ROC-AUC, R Square, F1-score, among others. The surveyed papers used multiple performance metrics to evaluate the quality of their student performance predictions. Twenty-one (33.87%) studies used a single metric to evaluate the prognosis of learning outcomes. Seven studies (11.29%) used two metrics, while seven other studies (11.29%) used three performance metrics. Notably, two studies used five evaluation metrics, including accuracy, kappa, AUC, sensitivity, specificity, relative absolute error (RAE), relative

squared error (RSE), and R-squared [66,85]. However, 21 (33.87%) studies did not specify the metrics they used, so the quality of their predictions was inconclusive.

When we counted the frequency of performance metrics utilized in the studies, we found that 28 (45.16%) intelligent models used 'accuracy' to measure the prediction quality, followed by the root mean square error (RMSE) (10 studies, 16.12%), ROC-AUC (8 studies, 12.90%), R square (8 studies, 12.90%), and mean absolute error (MAE) (7 studies, 11.29%). Table 9 summarizes the top and worst-performing prediction models of learning outcomes. Accordingly, the hybrid random forest [101] demonstrated the best classification accuracy, while the linear regression gave the worst predictions [88].

Table 9. Summary of top and worst prediction models based on the accuracy of predictions.

Top 5 Performing Prediction Models (Accuracy %)	Worst 5 Performing Prediction Models (Accuracy %)
Hybrid Random Forest [101]: 99.25–99.98%	Linear Regression [88]: 50%
Feedforward 3-L Neural Networks [74]: 98.81%	Bagging [78]: 48–55%
Random Forest [85]: 98%	Mixed-effects Logistic Regression [76]: 69%
Naive Bayes [93]: 96.87%	Discriminant Function Analysis [45]: 64–73%
Artificial Neural Network [86]: 95.16–97.30%	Logistic Regression [89]: 76.2%

Figure 12 shows that 38 (61.29%) studies did not benchmark the performance of their intelligent models against any baseline competitors. Fifteen (24.19%) studies compared their models with one to three competitor models. The remaining studies (14.51%) made a performance comparison with four or more baseline classifiers. The most compared against techniques included the Decision Tree (9 times), K-Nearest Neighbor (9 times), Support Vector Machines (8 times), Naïve Bayes (8 times), and Random Forest (6 times).
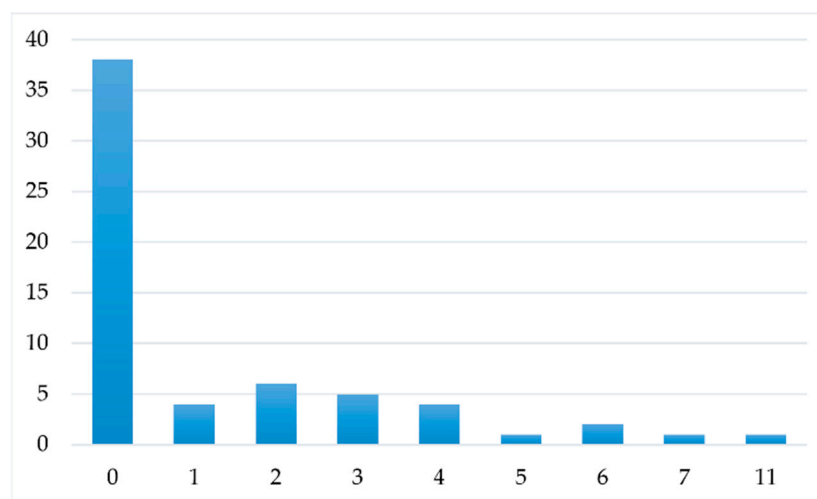

Figure 12. Number of intelligent methods against which each predictive model is benchmarked.

Only 5 (8.06%) studies reported using multiple datasets to verify the performance of their predictive models to check the consistency and validity of the learning outcomes predictions [65,80,86,96,102]. The remaining studies (91.93%) used only one dataset. With respect to the software used to analyze the datasets, statistical tools (e.g., SPSS, R, and Mplus) appeared in 14 studies, followed by data mining tools (e.g., WEKA), and machine learning frameworks (e.g., Keras, TensorFlow, and Scikit-learn). Other tools used included numerical computation (e.g., Octave) and in-house developed software (6 studies). It is worth noting that 29 articles did not specify the software tools they used to develop their predictive models.

4.5. Dominant Factors Predicting Student Learning Outcomes

Our survey revealed that 23 (37.09%) studies explored the impact of one to three factors on the attainment of student outcomes. However, 32 (51.61%) studies used more than three features to forecast student performance. The range of features varied between 4 (e.g., [53]) to 263 (e.g., [55]). Seven studies did not indicate the number of factors used to forecast student success. However, the dominant factors that were demonstrated to influence the attainment of student outcomes were substantially fewer. The strength of the evidence was grouped into three classes, namely strong, medium, and weak. Thirty-one (50%) studies reported finding strong evidence (i.e., statistical evidence or high prediction accuracy) about the predictive power of their factors. Six models showed medium evidence of the effects of the factors, while seven models reported a weak significance of the predictive factors they inspected. However, 18 studies were inconclusive about the strength of their findings.

We coded the factors (100 occurrences with 14 studies not reporting the influential factors) that were found to impact the performance of students into themes. Overall, six primary themes emerged from our qualitative analysis. Figure 13 shows that online learning activities and patterns (19 times) were the key predictors of student learning outcomes. This was mainly relevant to virtual or blended learning studies, where all or part of the student learning occurs online. Examples of online learning behavior included resource access time [84], site engagement [62], and time and number of online sessions [76]. The next prominent predictor of student performance was the assessment data during the semester (17 times), such as assignment [102] and quiz scores [44,82], and exam grades [69]. A surprising dominant factor of student achievements that prevailed was student academic emotions, which refer to student interests and enthusiasm [83], intrinsic motivations [92], and professor–student rapport [71]. The next influential features were grouped under previous academic achievements [45,46,48,66] and the teaching environment and style [81,98,99].
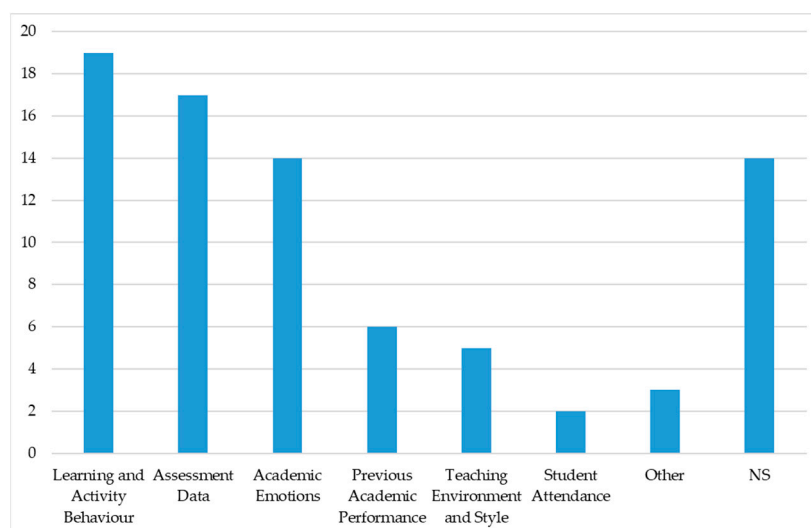

Figure 13. Dominant factors influencing the attainment of learning outcomes.; NS = not specified.

4.6. Quality Assessment of Reviewed Models

To assess the quality of the synthesized studies, we applied eight guidelines suggested in [3]. These guidelines were developed to evaluate the data analytics models. The guidelines assessed the clarity of the research questions and thoroughness of the methodology, and the use of a second dataset for validating the performance prediction models, among other vital aspects. Moreover, we took the liberty to add two quality assessment criteria, specifically (1) the practical implications of the student performance prediction model and (2) limitations of the model. Table 10 shows the overall quality assessment results of our 62 articles. Each study was carefully inspected and rated as to whether it satisfied each of

the ten guidelines. Studies that did not report information about a specific guideline were assumed to not fulfill the criterion.

Table 10. Quality assessment of surveyed studies; Yes = Condition satisfied, No = Condition not satisfied.

Assessment Criterion	Yes (%)	No (%)
1. Verification of predictive model with a second dataset	8.06%	91.94%
2. Threats to validity reported	12.90%	87.10%
3. Research implications and recommendations	20.96%	79.04%
4. Well-defined research questions	33.87%	66.13%
5. Use of separate training and testing datasets	35.48%	64.52%
6. Research limitations and challenges	37.09%	62.91%
7. Results detailed sufficiently	56.45%	43.56%
8. Predictor variables clearly described	77.42%	22.58%
9. Predictions being made are clear	82.25%	17.75%
10. Data collection instruments stated	82.25%	17.75%
11. Sound research methodology	83.87%	16.13%
12. Clear research contributions	90.32%	9.68%

Strikingly, only 21 studies (33.87%) posited clear research questions to motivate the learning analytics model development. While many studies described their contributions and research methodology clearly, they suffered from serious drawbacks. Only five (8.06%) studies stated verifying their predictive models using a second dataset. The majority of studies (87.10%) did not discuss the threats to the validity of their student performance predictions. Moreover, 49 articles (79.04%) did not draw any practical implications of their research findings, which considerably restricted the usefulness of the results for learning analytics and higher education. Finally, the models that discussed their limitations and challenges were limited to only 23 (37.09%).

5. Discussion

5.1. Key Findings

Outcome-based education has become pivotal for higher education leaders and accreditation organizations [12]. Moreover, learning analytics has gained tremendous momentum in the past decade to overcome the barriers hindering student learning [107]. Learning analytics and educational data mining (i.e., EDM) are proclaimed to improve the attainment of student learning outcomes [108]. There are also several calls for automating the assessment of student outcomes, which represent a proxy for student performance and success [9,29]. However, it is unclear how student outcomes are modeled and predicted at the course and program level using data mining and machine learning models. The current survey was carried out as an attempt to bridge this research gap.

The choice for surveying the last decade was motivated by the recent technological advances in artificial intelligence and data mining, coupled with the prominence of the outcome-based theory in education. The closest study to ours was the survey reported in [36], which explored 39 studies predicting learning outcomes in learning analytics between 2002 and 2016. Although the survey tried to summarize the main techniques used for predicting the learning outcomes, it failed to detail the results of the predictions. Our findings confirmed some previous observations. For instance, there seemed to be a growing interest in understanding student performance in learning management systems (LMS). The sample size of the datasets remained small to train the predictive models sufficiently. The predicted variable (i.e., learning outcome) evolved from a binary term to take a multi-

rank form; however, student grades are still used to refer to the learning outcomes. Lastly, the accuracy of the supervised learning models improved to reach unprecedented levels.

Our survey showed that developing models that forecast student learning outcomes is on the rise since 2017, with a significant portion of the articles published in computing and IT venues. Approximately half of the surveyed studies predicted learning outcomes of traditional classroom learning, while the other half focused on online and blended learning, due to its ever-increasing importance. More emphasis is directed toward undergraduate university courses and the STEM specialties (i.e., science, technology, engineering, and mathematics). The developed countries (e.g., the USA and Europe) are taking the lead in researching learning analytics of student outcomes. Below, we revisit each research question separately and highlight the main findings.

- RQ1-Learning Outcomes Prediction. How is student academic performance measured using learning outcomes?

  Here our analysis focused on understanding the forms in which the learning outcomes were measured in the selected studies. Our first observation was that the synthesized literature used the term 'student outcomes' or 'learning outcomes' incautiously, without adopting or linking them to any formal definition. The rather vague definition of the predicted variable (i.e., learning outcomes) by the predictive models was considered a major weakness that raises concerns about the usefulness and validity of the learning analytics results. Therefore, it is important for the researchers to clearly define the student learning outcomes variable that their intelligent models would estimate.

  Our next observation was that most experimental datasets came from a single educational entity, and 35% of studies predicted the learning outcomes for no more than four courses. The datasets used to train the predictive models were relatively small in many studies, with a sample size less than 1000 students. Notably, most surveyed models attempted to predict outcomes at the course level (90%). The academic performance was mostly measured for individual students instead of cohorts. Only a few studies modeled program-level outcomes. The predictions of educational outcomes were made both during and at the end of the semester. However, the projections focused on the direct measures of student performance more than the student perceptions of the learning process.

  Generally, the developed models analyzed student data to predict the learning outcomes in their variant forms, including student achievements, dropout and at-risk rates, and feedback and recommendation. Approximately 34 studies forecasted student performance in the form of academic standing classes (majority from two to four academic classes). Education program assessment is the epicenter activity undertaken to achieve a myriad of strategic goals [17], such as the improvement of program quality and realization of outcome-based education. Usually, student performance, whether assessed directly (e.g., examinations) or indirectly (e.g., student self-reports) is measured using rubrics. Rubrics can be considered the equivalent of academic performance classes to evaluate whether the learning outcomes fulfill certain thresholds or attainment levels.

- RQ2-Academic Performance Prediction Approaches. What intelligent approaches and techniques are devised to forecast student academic performance using learning outcomes?

  Although the number of publications in the field of educational data mining is growing yearly [109,110], the research efforts focused on developing models that can estimate learning outcomes are still unsatisfying. For instance, many outcome assessment tools lack sufficient intelligence to predict student performance [15]. In our survey, we found that the performance prediction models were developed, in most cases, as stand-alone modules and not part of a program assessment software. About 87.70% of the devised models relied on a single intelligent technique, even though the ensemble techniques are well-known to boost the prediction accuracy [3]. Moreover, fewer models were augmented to explain and justify the prediction of learning outcomes, despite their importance [111].

  Nearly 86% of the synthesized models fall within the statistical modeling and super machine learning. Only a few models tried to forecast student outcomes using unsu-

pervised learning techniques. We used the taxonomy presented in [109] to classify the predictive techniques emerging from our synthesis. Regression, neural network, and tree-based models were the most used classification techniques for predicting the attainment of student learning outcomes. Accuracy was the most calculated metric for evaluating the performance of the predictive models. Other evaluation metrics reported included RMSE, ROC-AUC, R square, and MAE. The best performing predictive models were the Hybrid Random Forest, Feedforward 3-L Neural Network, and Naïve Bayes, while the worst-performing models were the Linear Regression and Mixed-effects Logistic Regression. Remarkably, 61% of the proposed models did not benchmark their performance against other baseline classifiers. Finally, five studies re-examined the validity of their models on multiple datasets.

- RQ3-Academic Performance Predictors. What dominant predictors of student performance using learning outcomes are reported?

Learning analytics insights about student outcomes in the domain of education necessitate the investigation of features impacting academic performance [107]. Such understanding empowers the implementation of personal recommendations by the concerned education stakeholders [112]. However, our systematic survey demonstrated a lack of explanatory models that go beyond predicting the student performance to pinpointing the features that genuinely impact the attainment of course and program-level outcomes.

Approximately a third of the studies listed no more than three dominant factors to be influencing the accuracy of the academic outcome predictions. Similarly, nearly 30% of the studies were inconclusive about the effects of the features they explored. The thematic analysis revealed that student online learning patterns, term assessment scores, and student academic emotions are the top three predictors of learning outcomes.

The sample size of the synthesized studies differed significantly, as well as the number of courses used for understanding the impact of some features on the student learning outcomes. What works for one course might not work for another, and what works for one batch of students might behave adversely for another. In fact, student performance predictive models are known to work well, particularly for the datasets on which they were trained (i.e., model overfitting) and therefore have limited generalizability to new students and disciplines [111].

5.2. Challenges and Weaknesses of Existing Predictive Models

In our survey quest, we were enlightened about several challenges and underexplored areas prevailing in the existing learning outcomes prediction models. Future studies implementing machine learning models to prognosticate the attainment of student learning outcomes should pay close attention to the research challenges below and take necessary actions to mitigate them.

- Research challenge one: The prediction of academic performance of student cohorts to assist in the automation of course and program-level outcomes assessment.
- Research challenge two: The use and availability of multiple datasets from various disciplines to strengthen the validity of the predictive model. The datasets should comprise a large sample size of students to draw any meaningful conclusions.
- Research challenge three: The inspection of the effects of different features on the attainment of student outcomes to contribute to academic corrective interventions in higher education, i.e., the shift from predictive analytics to explanatory analytics.
- Research challenge four: The use of multiple performance evaluation metrics to assess the quality of the learning outcomes predictions.
- Research challenge five: The lack of unsupervised learning techniques devised to forecast student attainment of the learning outcomes.
- Research challenge six: The application of automated machine learning (i.e., AutoML) to the problem of student outcomes prediction was rarely conducted, except in [84]. Addressing this challenge would enable the development of ML models that automate the machine learning pipeline tasks, making the tasks of featurization, classification,

and forecasting efficient and accessible to the non-technical audience (e.g., education leaders and course instructors) in different disciplines.

5.3. Threats to Validity

In software engineering, validity assessment incorporates four types, namely internal, external, construct, and conclusion validity [113]. In this survey, we followed the recommended protocols to reduce the threats to validity and improve the quality of our conclusions. As such, we have:

- Defined the methodology, including search key terms and phrases, publication venues . . . etc., to enable the replicability of the survey.
- Used the manual search to incorporate any missing articles in the synthesis.
- Applied the appropriate inclusion and exclusion criteria to focus on student performance modeling using learning outcomes. These constituted the selection criteria of the survey.
- Selected all studies that meet the inclusion criteria irrespective of the researchers' background or nationality to eliminate any culture bias.
- Ensured that the primary studies are not repeated in the synthesis by removing the duplicates.
- Defined the quality assessment criteria based on previous surveys and recommendations [3].

However, the validity of our findings was largely influenced by the quality of the models we synthesized. We noticed that most studies focused on highlighting the models and factors that succeeded in forecasting student performance, thus, introducing a publication bias. Negative results were seldom published in the selected articles, which might have affected the results of our review. Indeed, this limits the practicality of the implications and recommendations.

Among the critical threats that hinder conducting valid surveys is missing any primary studies during the search process. To minimize this risk, we followed the best practices for conducting survey literature reviews in software engineering [10,11]. We also varied the critical search phrases for each electronic bibliographic database to retrieve as many relevant papers as possible. To reduce any subjective interpretation, we reviewed the extracted data and classification.

Concerning external validity, it is dangerous to assume the same observations for different disciplines (e.g., economics, history, . . . etc.), since most surveyed studies modeled student performance in a single discipline. Furthermore, the results ought to be treated with caution, especially with respect to generalization to other educational systems worldwide. Around 70% of studies were conducted in the USA and Europe alone, restricting the applicability of the results to the developing countries.

5.4. Survey Limitations

This work suffers from several qualifying limitations that are worthwhile to acknowledge herein. As with all types of reviews, there is a probability that we missed some works predicting student learning outcomes because of our selected search keywords and phrases, or during the screening process. Moreover, it was not possible to perform a meta-analysis of the previous findings to confirm the statistical significance of the synthesized predictive models, due to the unavailability of the datasets and the diverse techniques used to forecast student outcomes. We deliberately restricted our search of the intelligent predictive models of learning outcomes to the last decade only (i.e., 2010–2020), which witnessed a significant boost in machine learning on the one hand, and outcome-based education on the other hand. Therefore, we might have missed some critical works published before 2010. It was also observed that some studies did not report all experimental and prediction details, e.g., dataset characteristics, type of predictive models, and the factors influencing student academic success. For instance, 21 studies did not specify the performance metrics of the predictive models they devised. This eventually affected the quality of our synthe-

sis analysis. Unfortunately, many studies did not follow a detailed methodology, which made the assessment more challenging. Our survey was motivated by three research questions, which could have framed the review process, and thereby, the conclusions we reached. Other research questions might be asked and answered differently, leading to different results. Our search was restricted to peer-reviewed journals and conference articles, which could have overlooked valuable studies reported in dissertations, as well as in the unpublished literature.

6. Practical Implications and Recommendations

Based on the above challenges and limitations, we suggest the subsequent recommendations for research exploring the predictive learning analytics of student outcomes.

- Recommendation one: Formalize a clear definition of the variable 'learning outcomes' before embarking on the development of predictive models that measure the attainment of learning outcomes.
- Recommendation two: Build predictive models for non-technical majors, e.g., humanities, and for supporting teaching and learning in developing countries. These educational settings and contexts have different characteristics and features; therefore, specialized analytics models ought to be developed to work correctly in these settings.
- Recommendation three: Produce and share educational datasets for other researchers to explore and use after anonymizing any sensitive student data.
- Recommendation four: Build intelligent models that predict program-level outcomes as well as cohort academic performance. This would assist educational leaders in undertaking the activities of assessment and improve the quality of their programs.
- Recommendation five: Devise machine learning models that endeavor to explain and justify the attainment levels of student outcomes and explore the effectiveness of hybrid models in improving the accuracy of student outcomes predictions.

7. Future Directions

We strongly encourage the research community to conduct further work in the area of modeling the attainment of student outcomes, which is evidently still in its infancy, especially at the program level. The accuracy of the existing models ought to be improved and tested on multiple datasets to judge their validity and generalizability. More efforts should be dedicated toward understanding the impact of various factors on student performance and how these factors profoundly drive decision making at the course and program-level outcomes. In other words, the new efforts should work on developing explanatory predictions rather than models that merely forecast student performance. There is an overarching need to explain the relationship between possibly significant predictors and the observed attainment of learning outcomes, i.e., defining causal relationships and explanations that serve the learning analytics. Moreover, predicting learning outcomes should extend to other majors, such as humanities. Future work should consider reporting results regarding the intelligent models and factors that do not forecast student learning outcomes, i.e., negative results, besides publishing the positive results.

8. Conclusions

This systematic survey applied the SLR research recommendations to investigate the prediction of student outcomes, which is considered a proxy for student performance, using data mining and machine learning models. In particular, we applied the PRISMA protocol and SLR guidelines to produce the review. The exhaustive search of seven bibliographic databases yielded a synthesis of 62 primary articles. These articles presented intelligent models to forecast student performance using learning outcomes. The predictive models were published in peer-reviewed venues, spanning from 2010 till November 2020. To the best of our knowledge, this was the first published work that summarized the outstanding efforts of other researchers who studied the attainment of student outcomes. The prominent challenges included the prediction of academic performance at the program level and

student cohorts, the lack of explanatory analytics of the learning outcomes, the validation of performance prediction models to minimize the inherent underspecification problem of intelligent models, and the automation of the learning analytics tasks. We call upon the research community to implement the recommendations concerning (1) the prediction of program-level outcomes and (2) validation of the predictive models using multiple datasets from different majors and disciplines.

References

1.	Daniel, B. Big data and analytics in higher education: Opportunities and challenges. *Br. J. Educ. Technol.* **2015**, *46*, 904–920. [CrossRef]
2.	Zohair, L.M.A. Prediction of student's performance by modelling small dataset size. *Int. J. Educ. Technol. High. Educ.* **2019**, *16*, 27. [CrossRef]
3.	Hellas, A.; Ihantola, P.; Petersen, A.; Ajanovski, V.V.; Gutica, M.; Hynninen, T.; Knutas, A.; Leinonen, J.; Messom, C.; Liao, S.N. Predicting academic performance: A systematic literature review. In Proceedings of the Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education, Larnaca, Cyprus, 2–4 July 2018; pp. 175–199.
4.	Baradwaj, B.K.; Pal, S. Mining educational data to analyze students' performance. *Int. J. Adv. Comput. Sci. Appl.* **2012**, *2*, 63–69. [CrossRef]
5.	Zhang, L.; Li, K.F. Education analytics: Challenges and approaches. In Proceedings of the 2018 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA), Krakow, Poland, 16–18 May 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 193–198. [CrossRef]
6.	Daud, A.; Aljohani, N.R.; Abbasi, R.A.; Lytras, M.D.; Abbas, F.; Alowibdi, J.S. Predicting student performance using advanced learning analytics. In Proceedings of the 26th International Conference on World Wide Web Companion, Perth, Australia, 3–7 April 2017; pp. 415–421.
7.	Macayan, J.V. Implementing outcome-based education (OBE) framework: Implications for assessment of students' performance. *Educ. Meas. Eval. Rev.* **2017**, *8*, 1–10.
8.	Yassine, S.; Kadry, S.; Sicilia, M.A. A framework for learning analytics in moodle for assessing course outcomes. In Proceedings of the 2016 IEEE Global Engineering Education Conference (EDUCON), Abu Dhabi, UAE, 10–13 April 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 261–266.
9.	Rajak, A.; Shrivastava, A.K.; Shrivastava, D.P. Automating outcome based education for the attainment of course and program outcomes. In Proceedings of the 2018 Fifth HCT Information Technology Trends (ITT), Dubai, UAE, 28–29 November 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 373–376.
10.	Kitchenham, B.; Charters, S. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*; EBSE: Keele, UK, 2007; pp. 1–65.
11.	Okoli, C.; Schabram, K. A guide to conducting a systematic literature review of information systems research. *Ssrn Eletronic J.* **2010**, *10*. [CrossRef]
12.	Kaliannan, M.; Chandran, S.D. Empowering students through outcome-based education (OBE). *Res. Educ.* **2012**, *87*, 50–63. [CrossRef]
13.	Premalatha, K. Course and program outcomes assessment methods in outcome-based education: A review. *J. Educ.* **2019**, *199*, 111–127. [CrossRef]
14.	Kanmani, B.; Babu, K.M. Leveraging technology in outcome-based education. In Proceedings of the International Conference on Transformations in Engineering Education, New Delhi, India, 5–8 January 2015; Natarajan, R., Ed.; Springer: New Delhi, India, 2015; pp. 415–421.

15. Namoun, A.; Taleb, A.; Benaida, M. An expert comparison of accreditation support tools for the undergraduate computing programs. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* **2018**, *9*, 371–384. [CrossRef]

16. Mahajan, M.; Singh, M.K.S. Importance and benefits of learning outcomes. *IOSR J. Humanit. Soc. Sci.* **2017**, *22*, 65–67. [CrossRef]

17. Namoun, A.; Taleb, A.; Al-Shargabi, M.; Benaida, M. A learning outcome inspired survey instrument for assessing the quality of continuous improvement cycle. *Int. J. Inf. Commun. Technol. Educ. (IJICTE)* **2019**, *15*, 108–129. [CrossRef]

18. Taleb, A.; Namoun, A.; Benaida, M. A holistic quality assurance framework to acquire national and international. *J. Eng. Appl. Sci.* **2019**, *14*, 6685–6698. [CrossRef]

19. Singh, R.; Sarkar, S. *Teaching Quality Counts: How Student Outcomes Relate to Quality of Teaching in Private and Public Schools in India*; Young Lives: Oxford, UK, 2012; pp. 1–48.

20. Philip, K.; Lee, A. Online public health education for low and middle-income countries: Factors influencing successful student outcomes. *Int. J. Emerg. Technol. Learn. (IJET)* **2011**, *6*, 65–69. [CrossRef]

21. Garbacz, S.A.; Herman, K.C.; Thompson, A.M.; Reinke, W.M. Family engagement in education and intervention: Implementation and evaluation to maximize family, school, and student outcomes. *J. Sch. Psychol.* **2017**, *62*, 1–10. [CrossRef] [PubMed]

22. Nonis, S.A.; Fenner, G.H. An exploratory study of student motivations for taking online courses and learning outcomes. *J. Instr. Pedagog.* **2012**, *7*, 2–13.

23. Polyzou, A.; Karypis, G. Feature extraction for next-term prediction of poor student performance. *IEEE Trans. Learn. Technol.* **2019**, *12*, 237–248. [CrossRef]

24. Shahiri, A.M.; Husain, W.; Abdul Rashid, N. A review on predicting student's performance using data mining techniques. *Procedia Comput. Sci.* **2015**, *72*, 414–422. [CrossRef]

25. Tatar, A.E.; Düştegör, D. Prediction of academic performance at undergraduate graduation: Course grades or grade point average? *Appl. Sci.* **2020**, *10*, 4967. [CrossRef]

26. Elbadrawy, A.; Polyzou, A.; Ren, Z.; Sweeney, M.; Karypis, G.; Rangwala, H. Predicting student performance using personalized analytics. *Computer* **2016**, *49*, 61–69. [CrossRef]

27. Cui, Y.; Chen, F.; Shiri, A.; Fan, Y. Predictive analytic models of student success in higher education: A review of methodology. *Inf. Learn. Sci.* **2019**, *120*, 208–227. [CrossRef]

28. Rastrollo-Guerrero, J.L.; Gómez-Pulido, J.A.; Durán-Domínguez, A. Analyzing and predicting students' performance by means of machine learning: A review. *Appl. Sci.* **2020**, *10*, 1042. [CrossRef]

29. Alshanqiti, A.; Namoun, A. Predicting student performance and its influential factors using hybrid regression and multi-label classification. *IEEE Access* **2020**, *8*, 203827–203844. [CrossRef]

30. Mthimunye, K.; Daniels, F.M. Predictors of academic performance, success and retention amongst undergraduate nursing students: A systematic review. *S. Afr. J. High. Educ.* **2019**, *33*, 200–220. [CrossRef]

31. Dixson, D.D.; Worrell, F.C.; Olszewski-Kubilius, P.; Subotnik, R.F. Beyond perceived ability: The contribution of psychosocial factors to academic performance. *Ann. N. Y. Acad. Sci.* **2016**, *1377*, 67–77. [CrossRef] [PubMed]

32. Felix, I.; Ambrósio, A.P.; Lima, P.D.S.; Brancher, J.D. Data mining for student outcome prediction on moodle: A systematic mapping. In Proceedings of the Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE), Fortaleza, Brazil, 29 October–1 November 2018; Volume 29, p. 1393. [CrossRef]

33. Peña-Ayala, A. Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Syst. Appl.* **2014**, *41*, 1432–1462. [CrossRef]

34. Kumar, M.; Singh, A.J.; Handa, D. Literature survey on student's performance prediction in education using data mining techniques. *Int. J. Educ. Manag. Eng.* **2017**, *7*, 40–49. [CrossRef]

35. Ofori, F.; Maina, E.; Gitonga, R. Using machine learning algorithms to predict students' performance and improve learning outcome: A literature based review. *J. Inf. Technol.* **2020**, *4*, 33–55.

36. Hu, X.; Cheong, C.W.; Ding, W.; Woo, M. A systematic review of studies on predicting student learning outcomes using learning analytics. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference, Vancouver, BC, Canada, 13–17 March 2017; pp. 528–529. [CrossRef]

37. Magalhães, P.; Ferreira, D.; Cunha, J.; Rosário, P. Online vs traditional homework: A systematic review on the benefits to students' performance. *Comput. Educ.* **2020**, *152*, 103869. [CrossRef]

38. Digregorio, P.; Sobel-Lojeski, K. The effects of interactive whiteboards (IWBs) on student performance and learning: A literature review. *J. Educ. Technol. Syst.* **2010**, *38*, 255–312. [CrossRef]

39. van der Zanden, P.J.; Denessen, E.; Cillessen, A.H.; Meijer, P.C. Domains and predictors of first-year student success: A systematic review. *Educ. Res. Rev.* **2018**, *23*, 57–77. [CrossRef]

40. Bain, S.; Fedynich, L.; Knight, M. The successful graduate student: A review of the factors for success. *J. Acad. Bus. Ethics* **2011**, *3*, 1.

41. Petersen, K.; Vakkalanka, S.; Kuzniarz, L. Guidelines for conducting systematic mapping studies in software engineering: An update. *Inf. Softw. Technol.* **2015**, *64*, 1–18. [CrossRef]

42. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Prisma Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ* **2009**, *6*, 1–8. [CrossRef]

43. Ming, N.C.; Ming, V.L. Predicting student outcomes from unstructured data. In *UMAP Workshops*; CEUR Workshop Proceedings: Aachen, Germany, 2012.

44. Heise, N.; Meyer, C.A.; Garbe, B.A.; Hall, H.A.; Clapp, T.R. Table quizzes as an assessment tool in the gross anatomy laboratory. *J. Med. Educ. Curric. Dev.* **2020**, *7*. [CrossRef]

45. Shulruf, B.; Bagg, W.; Begun, M.; Hay, M.; Lichtwark, I.; Turnock, A.; Warnecke, E.; Wilkinson, T.J.; Poole, P.J. The efficacy of medical student selection tools in Australia and New Zealand. *Med. J. Aust.* **2018**, *208*, 214–218. [CrossRef] [PubMed]

46. Moreno-Marcos, P.M.; Pong, T.C.; Muñoz-Merino, P.J.; Kloos, C.D. Analysis of the factors influencing learners' performance prediction with learning analytics. *IEEE Access* **2020**, *8*, 5264–5282. [CrossRef]

47. Martin, A.J.; Nejad, H.G.; Colmar, S.; Liem, G.A.D. Adaptability: How students' responses to uncertainty and novelty predict their academic and non-academic outcomes. *J. Educ. Psychol.* **2013**, *105*, 728. [CrossRef]

48. Bowers, A.J.; Zhou, X. Receiver operating characteristic (ROC) area under the curve (AUC): A diagnostic measure for evaluating the accuracy of predictors of education outcomes. *J. Educ. Stud. Placed Risk (JESPAR)* **2019**, *24*, 20–46. [CrossRef]

49. Palmer, L.E.; Erford, B.T. Predicting student outcome measures using the ASCA national model program audit. *Prof. Couns.* **2012**, *2*, 152–159. [CrossRef]

50. Fauth, B.; Decristan, J.; Rieser, S.; Klieme, E.; Büttner, G. Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learn. Instr.* **2014**, *29*, 1–9. [CrossRef]

51. Harred, R.; Cody, C.; Maniktala, M.; Shabrina, P.; Barnes, T.; Lynch, C. How Long is Enough? Predicting Student Outcomes with Same-Day Gameplay Data in an Educational Math Game. In Proceedings of the Educational Data Mining (Workshops), Montréal, QC, Canada, 2–5 July 2019; pp. 60–68.

52. Aldrup, K.; Klusmann, U.; Lüdtke, O.; Göllner, R.; Trautwein, U. Social support and classroom management are related to secondary students' general school adjustment: A multilevel structural equation model using student and teacher ratings. *J. Educ. Psychol.* **2018**, *110*, 1066. [CrossRef]

53. Van Ryzin, M. Secondary school advisors as mentors and secondary attachment figures. *J. Community Psychol.* **2010**, *38*, 131–154. [CrossRef]

54. Porayska-Pomsta, K.; Mavrikis, M.; Cukurova, M.; Margeti, M.; Samani, T. Leveraging non-cognitive student self-reports to predict learning outcomes. In Proceedings of the International Conference on Artificial Intelligence in Education, London, UK, 27–30 June 2018; Springer: Cham, Switzerland, 2018; pp. 458–462.

55. Kórösi, G.; Esztelecki, P.; Farkas, R.; Tóth, K. Clickstream-based outcome prediction in short video MOOCs. In Proceedings of the 2018 International Conference on Computer, Information and Telecommunication Systems (CITS), Colmar, France, 11–13 July 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–5.

56. Brinkworth, M.E.; McIntyre, J.; Juraschek, A.D.; Gehlbach, H. Teacher-student relationships: The positives and negatives of assessing both perspectives. *J. Appl. Dev. Psychol.* **2018**, *55*, 24–38. [CrossRef]

57. Mantzicopoulos, P.; Patrick, H.; Strati, A.; Watson, J.S. Predicting kindergarteners' achievement and motivation from observational measures of teaching effectiveness. *J. Exp. Educ.* **2018**, *86*, 214–232. [CrossRef]

58. Aelterman, N.; Vansteenkiste, M.; Haerens, L. Correlates of students' internalization and defiance of classroom rules: A self-determination theory perspective. *Br. J. Educ. Psychol.* **2019**, *89*, 22–40. [CrossRef]

59. Simjanoska, M.; Gusev, M.; Ristov, S.; Bogdanova, A.M. Intelligent student profiling for predicting e-assessment outcomes. In Proceedings of the 2014 IEEE Global Engineering Education Conference (EDUCON), Istanbul, Turkey, 3–5 April 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 616–622.

60. Pang, Y.; Judd, N.; O'Brien, J.; Ben-Avie, M. Predicting students' graduation outcomes through support vector machines. In Proceedings of the 2017 IEEE Frontiers in Education Conference (FIE), Indianapolis, IN, USA, 18–21 October 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–8.

61. Liu, K.F.R.; Chen, J.S. Prediction and assessment of student learning outcomes in calculus a decision support of integrating data mining and Bayesian belief networks. In Proceedings of the 2011 3rd International Conference on Computer Research and Development, Shanghai, China, 11–13 March 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 299–303.

62. Smith, V.C.; Lange, A.; Huston, D.R. Predictive modeling to forecast student outcomes and drive effective interventions in online community college courses. *J. Asynchronous Learn. Netw.* **2012**, *16*, 51–61. [CrossRef]

63. Pavani, M.; Teja, A.R.; Neelima, A.; Bhavishya, G.; Sukrutha, D.S. Prediction of student outcome in educational sector by using decision tree. *Int. J. Technol. Res. Eng.* **2017**, *4*, 2347–4718.

64. Zacharis, N.Z. A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *Internet High. Educ.* **2015**, *27*, 44–53. [CrossRef]

65. Gray, C.C.; Perkins, D. Utilizing early engagement and machine learning to predict student outcomes. *Comput. Educ.* **2019**, *131*, 22–32. [CrossRef]

66. Iatrellis, O.; Savvas, I.K.; Fitsilis, P.; Gerogiannis, V.C. A two-phase machine learning approach for predicting student outcomes. *Educ. Inf. Technol.* **2020**, 1–20. [CrossRef]

67. Kuzilek, J.; Vaclavek, J.; Zdrahal, Z.; Fuglik, V. Analysing Student VLE Behaviour Intensity and Performance. In Proceedings of the European Conference on Technology Enhanced Learning, Delft, The Netherlands, 16–19 September 2019; Springer: Cham, Switzerland, 2019; pp. 587–590.

68. Raga, R.; Raga, J. Early Prediction of Student Performance in Blended Learning Courses Using Deep Neural Networks. In Proceedings of the 2019 International Symposium on Educational Technology (ISET), Hradec Kralove, Czech Republic, 2–4 July 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 39–43.

69. Walsh, K.R.; Mahesh, S. Exploratory study using machine learning to make early predictions of student outcomes. In Proceedings of the Twenty-third Americas Conference on Information Systems, Data Science and Analytics for Decision Support (SIGDSA), Boston, MA, USA, 10–12 August 2017; AIS: Atlanta, GA, USA, 2017; pp. 1–5.

70. Olama, M.M.; Thakur, G.; McNair, A.W.; Sukumar, S.R. Predicting student success using analytics in course learning management systems. In *Next-Generation Analyst II*; International Society for Optics and Photonics: Washington, DC, USA, 2014; p. 91220M.

71. Wilson, J.H.; Ryan, R.G. Professor–student rapport scale: Six items predict student outcomes. *Teach. Psychol.* **2013**, *40*, 130–133. [CrossRef]

72. Wilson, J.H.; Ryan, R.G.; Pugh, J.L. Professor–student rapport scale predicts student outcomes. *Teach. Psychol.* **2010**, *37*, 246–251. [CrossRef]

73. Kuzilek, J.; Vaclavek, J.; Fuglik, V.; Zdrahal, Z. Student Drop-out Modelling Using Virtual Learning Environment Behaviour Data. In Proceedings of the European Conference on Technology Enhanced Learning, Leeds, UK, 3–5 September 2018; Springer: Cham, Switzerland, 2018; pp. 166–171.

74. Zaporozhko, V.V.; Parfenov, D.I.; Shardakov, V.M. Development Approach of Formation of Individual Educational Trajectories Based on Neural Network Prediction of Student Learning Outcomes. In Proceedings of the International Conference of Artificial Intelligence, Medical Engineering, Education, Moscow, Russia, 3–4 October 2019; Springer: Cham, Switzerland, 2019; pp. 305–314.

75. Ruiz, S.; Urretavizcaya, M.; Rodríguez, C.; Fernández-Castro, I. Predicting students' outcomes from emotional response in the classroom and attendance. *Interact. Learn. Environ.* **2020**, *28*, 107–129. [CrossRef]

76. Eagle, M.; Carmichael, T.; Stokes, J.; Blink, M.J.; Stamper, J.C.; Levin, J. Predictive Student Modeling for Interventions in Online Classes. In Proceedings of the 11th International Conference on Educational Data Mining EDM, Buffalo, NY, USA, 15–18 July 2018; pp. 619–624.

77. Alonso, J.M.; Casalino, G. Explainable Artificial Intelligence for Human-Centric Data Analysis in Virtual Learning Environments. In Proceedings of the International Workshop on Higher Education Learning Methodologies and Technologies Online, Novedrate, Italy, 6–7 June 2019; Springer: Cham, Switzerland, 2019; pp. 125–138.

78. Kőrösi, G.; Farkas, R. MOOC Performance Prediction by Deep Learning from Raw Clickstream Data. In Proceedings of the International Conference on Advances in Computing and Data Sciences, Maharashtra, India, 23–24 April 2020; Springer: Singapore, 2020; pp. 474–485.

79. Culligan, N.; Quille, K.; Bergin, S. Veap: A visualization engine and analyzer for press#. In Proceedings of the 16th Koli Calling International Conference on Computing Education Research, Koli, Finland, 24–27 November 2016; pp. 130–134.

80. Umer, R.; Mathrani, A.; Susnjak, T.; Lim, S. Mining Activity Log Data to Predict Student's Outcome in a Course. In Proceedings of the 2019 International Conference on Big Data and Education, London, UK, 27–29 March 2019; pp. 52–58.

81. Yadav, A.; Alexander, V.; Mehta, S. Case-based Instruction in Undergraduate Engineering: Does Student Confidence Predict Learning. *Int. J. Eng. Educ.* **2019**, *35*, 25–34.

82. Strang, K.D. Beyond engagement analytics: Which online mixed-data factors predict student learning outcomes? *Educ. Inf. Technol.* **2017**, *22*, 917–937. [CrossRef]

83. Ketonen, E.; Lonka, K. Do situational academic emotions predict academic outcomes in a lecture course? *Procedia Soc. Behav. Sci.* **2012**, *69*, 1901–1910. [CrossRef]

84. Tsiakmaki, M.; Kostopoulos, G.; Kotsiantis, S.; Ragos, O. Implementing AutoML in educational data mining for prediction tasks. *Appl. Sci.* **2020**, *10*, 90. [CrossRef]

85. Al-Shabandar, R.; Hussain, A.; Laws, A.; Keight, R.; Lunn, J.; Radi, N. Machine learning approaches to predict learning outcomes in Massive open online courses. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 713–720.

86. Yu, C.H.; Wu, J.; Liu, A.C. Predicting learning outcomes with MOOC clickstreams. *Educ. Sci.* **2019**, *9*, 104. [CrossRef]

87. Zabriskie, C.; Yang, J.; DeVore, S.; Stewart, J. Using machine learning to predict physics course outcomes. *Phys. Rev. Phys. Educ. Res.* **2019**, *15*, 020120. [CrossRef]

88. Nguyen, V.A.; Nguyen, Q.B.; Nguyen, V.T. A model to forecast learning outcomes for students in blended learning courses based on learning analytics. In Proceedings of the 2nd International Conference on E-Society, E-Education and E-Technology, Taipei, Taiwan, 13–15 August 2018; pp. 35–41.

89. Guo, S.; Wu, W. Modeling student learning outcomes in MOOCs. In Proceedings of the 4th International Conference on Teaching, Assessment, and Learning for Engineering, Zhuhai, China, 10–12 December 2015; pp. 1305–1313.

90. Foung, D.; Chen, J. A learning analytics approach to the evaluation of an online learning package in a Hong Kong University. *Electron. J. E Learn.* **2019**, *17*, 11–24.

91. Akhtar, S.; Warburton, S.; Xu, W. The use of an online learning and teaching system for monitoring computer aided design student participation and predicting student success. *Int. J. Technol. Des. Educ.* **2017**, *27*, 251–270. [CrossRef]

92. Gratiano, S.M.; Palm, W.J. Can a five minute, three question survey foretell first-year engineering student performance and retention? In Proceedings of the 123rd ASEE Annual Conference & Exposition, New Orleans, LA, USA, 26–29 June 2016.

93. Vasić, D.; Kundid, M.; Pinjuh, A.; Šerić, L. Predicting student's learning outcome from Learning Management system logs. In Proceedings of the 2015 23rd International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Bol (Island of Brac), Croatia, 16–18 September 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 210–214.

94. Felix, I.; Ambrosio, A.; Duilio, J.; Simões, E. Predicting student outcome in moodle. In Proceedings of the Conference: Academic Success in Higher Education, Porto, Portugal, 14–15 February 2019; pp. 1–2.

95. Alkoot, F.M. Using classifiers to predict student outcome at HITN-PAAET. In Proceedings of the 18th International Conference on Machine Learning and Data Analysis, Tokyo, Japan, 22–24 May 2016.

96. Wang, X.; Mei, X.; Huang, Q.; Han, Z.; Huang, C. Fine-grained learning performance prediction via adaptive sparse self-attention networks. *Inf. Sci.* **2020**, *545*, 223–240. [CrossRef]

97. Pianta, R.C.; Ansari, A. Does attendance in private schools predict student outcomes at age 15? Evidence from a longitudinal study. *Educ. Res.* **2018**, *47*, 419–434. [CrossRef]

98. Hill, H.C.; Charalambous, C.Y.; Chin, M.J. Teacher characteristics and student learning in mathematics: A comprehensive assessment. *Educ. Policy* **2019**, *33*, 1103–1134. [CrossRef]

99. Anderson, K.A. A national study of the differential impact of novice teacher certification on teacher traits and race-based mathematics achievement. *J. Teach. Educ.* **2020**, *71*, 247–260. [CrossRef]

100. Lima, P.D.S.N.; Ambrósio, A.P.L.; Félix, I.M.; Brancher, J.D.; Ferreira, D.J. Content Analysis of Student Assessment Exams. In Proceedings of the 2018 IEEE Frontiers in Education Conference (FIE), San Jose, CA, USA, 3–6 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–9.

101. Sokkhey, P.; Okazaki, T. Developing web-based support systems for predicting poor-performing students using educational data mining techniques. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 23–32. [CrossRef]

102. Sales, A.; Botelho, A.F.; Patikorn, T.; Heffernan, N.T. Using big data to sharpen design-based inference in A/B tests. In Proceedings of the Eleventh International Conference on Educational Data Mining, Buffalo, NY, USA, 15–18 January 2018.

103. Bhatia, J.; Girdhar, A.; Singh, I. An Automated Survey Designing Tool for Indirect Assessment in Outcome Based Education Using Data Mining. In Proceedings of the 2017 5th IEEE International Conference on MOOCs, Innovation and Technology in Education (MITE), Bangalore, India, 27–28 October 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 95–100.

104. Bindra, S.K.; Girdhar, A.; Bamrah, I.S. Outcome based predictive analysis of automatic question paper using data mining. In Proceedings of the 2017 2nd International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 19–20 October 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 629–634.

105. Joksimović, S.; Kovanović, V.; Dawson, S. The journey of learning analytics. *Herdsa Rev. High. Educ.* **2019**, *6*, 27–63.

106. Kumari, P.; Jain, P.K.; Pamula, R. An efficient use of ensemble methods to predict students academic performance. In Proceedings of the 2018 4th International Conference on Recent Advances in Information Technology (RAIT), Dhanbad, India, 15–17 March 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.

107. Arroway, P.; Morgan, G.; O'Keefe, M.; Yanosky, R. *Learning Analytics in Higher Education*; Research report; ECAR: Louisville, CO, USA, 2016; p. 17.

108. Viberg, O.; Hatakka, M.; Bälter, O.; Mavroudi, A. The current landscape of learning analytics in higher education. *Comput. Hum. Behav.* **2018**, *89*, 98–110. [CrossRef]

109. Manjarres, A.V.; Sandoval, L.G.M.; Suárez, M.S. Data mining techniques applied in educational environments: Literature review. *Digit. Educ. Rev.* **2018**, *33*, 235–266.

110. Romero, C.; Ventura, S. Educational data mining and learning analytics: An updated survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2020**, *10*, e1355. [CrossRef]

111. Shmueli, G. To explain or to predict? *Stat. Sci.* **2010**, *25*, 289–310. [CrossRef]

112. Ranjeeth, S.; Latchoumi, T.P.; Paul, P.V. A survey on predictive models of learning analytics. *Procedia Comput. Sci.* **2020**, *167*, 37–46. [CrossRef]

113. Zhou, X.; Jin, Y.; Zhang, H.; Li, S.; Huang, X. A map of threats to validity of systematic literature reviews in software engineering. In Proceedings of the 2016 23rd Asia-Pacific Software Engineering Conference (APSEC) IEEE, Hamilton, New Zealand, 6–9 December 2016; pp. 153–160.