

Article

Predicting At-Risk Students in an Online Flipped Anatomy Course Using Learning Analytics

Alper Bayazit ^{1,*}, Nihal Apaydin ² and Ipek Gonullu ¹

¹ Department of Medical Education and Informatics, Faculty of Medicine, Ankara University, 06620 Ankara, Turkey

² Department of Anatomy, Faculty of Medicine, Ankara University, 06630 Ankara, Turkey

* Correspondence: abayazit@ankara.edu.tr

Abstract: When using the flipped classroom method, students are required to come to the lesson after having prepared the basic concepts. Thus, the effectiveness of the lecture depends on the students' preparation. With the ongoing COVID-19 pandemic, it has become difficult to examine student preparations and to predict student course failures with limiting variables. Learning analytics can overcome this limitation. In this study, we aimed to develop a predictive model for at-risk students who are at risk of failing their final exam in an introductory anatomy course. In a five-week online flipped anatomy course, students' weekly interaction metrics, quiz scores, and pretest scores were used to design a predictive model. We also compared the performances of different machine learning algorithms. According to the results, the Naïve Bayes algorithm showed the best performance for predicting student grades with an overall classification accuracy of 68% and with at-risk prediction accuracy of 71%. These results can be used as a traffic light project wherein the "at-risk" group will receive the red light, and thus, will require more effort to engage with the content and they might need to solve the quiz tests after an individual study period.



Citation: Bayazit, A.; Apaydin, N.; Gonullu, I. Predicting At-Risk Students in an Online Flipped Anatomy Course Using Learning Analytics. *Educ. Sci.* **2022**, *12*, 581. <https://doi.org/10.3390/educsci12090581>

Academic Editors: Maria Alessandra Sotgiu and Bernard John Moxham

Received: 9 May 2022

Accepted: 19 August 2022

Published: 24 August 2022

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



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Flipped classrooms (FC) are a teaching method that flips traditional teaching methods and makes it possible to spare time in class activities [1–3]. In traditional classrooms, students usually participate in classroom activities with no preparation. The teacher provides the basic concepts related to the lesson and the expectation is that students will understand and comprehend these concepts. Since the in-class activity time is limited, they devote most of the lesson time to these activities [4]. At the end of the lesson, the teacher assigns students homework that includes higher-level cognitive tasks like application, analysis, and synthesis [5–7]. However, since the students are out of the classroom, these tasks are very difficult to complete without a guide. The FC approach removes these limitations of traditional classrooms [8,9]. Teaching the basic concepts related to learning goals before the lesson by presenting different resources not only ensures that different student characteristics are considered [10] but requires students to reach these goals before the lesson [11]. At the beginning of the lesson, a short quiz is applied in which the lesson preparations of the learners are determined [12,13]. The quiz allows for the students' deficiencies to be determined as well as for feedback to be given to the student and tutor. In the remaining class time, activities that activate higher-level cognitive processes, such as classroom discussions, problem-solving, and case-based learning processes are implemented [14–16]. Within this paradigm, the teacher is a guide and student-centered learning processes are applied [17,18].

FC practices increase students' interactions with their content and enable more in-class activities [8], increase student motivation [19,20], enable the retention of the information [21,22], improve problem-solving skills [23,24], and increase learner satisfaction. With

the COVID-19 period, in particular, online flipped classroom applications have become widespread [25,26]. The synchronous participation of students in their courses may be limited in distance education, and this limited period of availability must be supported by application activities. Learning management system (LMS) interactions with content that includes course resources such as educational videos, online discussions, documents, and simulations have grown increasingly relevant and also support FC applications. While students who participate in a lecture following preparation can benefit more from online education, students who cannot make these preparations may not learn effectively and may fail the course. The decrease in student–instructor interactions and tracking, the difficulty of directly observing learners who are at risk of failing, as well as the difficulty of intervening before failure occurs, are among the most important limitations of flipped classroom practices. However, learning analytics can overcome these limitations.

Learning analytics consists of four elements: collecting data on learners and the learning environment, pre-processing the data collected, applying machine learning algorithms and other statistical methods to determine the patterns of the students, and interventions to increase success and prevent failure [27–29]. These components also enable the creation of educational early warning systems. An early warning system aims to predict academic performance or course achievement at an earlier stage [27]. Identifying students at risk of academic failure and providing them with the help they need may help these students succeed. “At-risk” refers to students at risk of failing either their final exam or failing to progress [30]. Identifying students who are at risk of failing or not completing courses can benefit higher education institutions [31]. For this purpose, researchers or instructors can apply learning analytics to create predictive models to enable the early detection of students with failure risk and provide them with the appropriate feedback and intervention. One of the most widely used techniques in early warning systems is supervised learning. These techniques are efficient in helping to solve non-linear real-time problems as classification plays a vital role in machine-based learning algorithms [32]. Supervised machine learning is a technique in which machines are trained using data that is labeled with attributes, and based on this data, the output is predicted in the previous experience. Classification is a supervised machine learning concept, where a class label is predicted for a sample of input data. In previous studies, predictive models were created based on classification methods; researchers usually compared five classifiers to forecast accuracy and identify the best performing classification algorithm. These classifiers are k-nearest neighbors (kNN) [33], decision trees (DT) [34], naïve Bayes (NB) [35], random forest (RF) [36], and support vector machines (SVM) [37]. In summary, to create an early warning system in an online course that is flipped: (1) students’ LMS interactions must be transformed into a weekly interaction metric and merged with their quiz scores for each week; (2) data pre-processing procedures must be applied for an appropriate predictive model; (3) a model must be created to predict the students’ final grades with two-class (pass/fail or safe/at-risk) classification and the best performing algorithm must be determined based on the model performance metrics; (4) the model must be used as an early warning system for at-risk students before the final exam. This “warning message” represents the most important stage of learning analytics, “interventions”.

Hung, Shelton [38] applied machine learning techniques to develop a general model for higher education and K-12. They used several variables including gender, grade, LMS interactions, content views, participation in online discussions, number of replies to comments, and student grade checks. For classification predictions, they applied three commonly used machine learning techniques: SVM, deep neural network (DNN), and RF. According to the results, DNN and RF showed accuracy rates in predicting at-risk students of between 82.6% and 85% for the K-12 data and between 95% and 97% for higher education data. Hung, Wang [39] applied time-series clustering to early identify at-risk online students. Their study aimed to deliver more accurate predictions than those of traditional analysis and to identify the at-risk patterns of the students based on the model, which comprise interaction and personal student data. Students’ LMS interaction

data, including the number of course materials accessed, number of forums read, number of discussions posted, as well as the number of replies from these environments, were collected. In addition, other variables such as the final grade (which was also used as a supervised variable), gender, ethnicity, and admission status were included in the analysis. The results showed that a time-series approach was able to identify at-risk students during the semester for early intervention. In addition, the decision tree was the best performing and most stable model. Akçapinar, Altun [27] employed learning analytics to design and develop an early warning system for at-risk students. Their study aimed to compare the classification accuracy of machine learning algorithms in identifying at-risk students and to investigate the effect of different pre-process applications on prediction performance. In their study, they used LMS interaction features including several question ratings, unique session days, unique posting days, the total number of posts, the number of total sessions, the number of tags created, session duration, and the number of responses to the discussion forums. They included student grades in the analyses as a supervised learning variable for different classification techniques, such as classification tree, CN2 rules, NB, neural network, k-NN, RF, and SVM. The k-NN algorithm accurately predicted at-risk students at the end of the term with a rate of 89%. In addition, students who were at-risk at the end of term were predicted to a rate of 74% in 3 weeks. Saqr, Fors [29] applied learning analytics to a medical course for the early detection of under-achieving students in a blended learning environment. In their predictive study, they analyzed the LMS interactions of 133 medical students in order to develop a model for final grades. The collected data included Moodle LMS login counts, number of course views, number of forum views, total time spent in LMS, and formative assessment scores. According to their results, the predictive model showed 63.5% accuracy and 53.9% of at-risk student prediction. They also analyzed the predictive importance of the linear modeling and found login counts and views of course information were the most important measures for the predictive model.

Purdue University used the “course signals” project as an early warning system in their LMS [40]. Students are unaware of how they are progressing within their courses, for example, if they only take a few assessments during their courses and do not see their final grades until the end of the term. Signals remedy this problem. The components, including students’ performance, effort, prior academic achievements, and student characteristics, are weighted and fed into an algorithm. This creates a traffic signal based on the probability of success or failure [40,41]. The red light shows a high likelihood of failure, yellow shows potential problems, and green shows a high likelihood of success. Motivational feedback typically includes remarks like “Well done” or “Fantastic”, whereas informative feedback provides details on the learner’s progress, and how it compares to where they were previously or to others. A green traffic signal represents positive feedback, which the instructor can reinforce with a message of encouragement [42]. An example of this kind of intervention is a study by Foster and Siddle [30] who used an LMS that generated “no-engagement” alerts if students did not engage with any resources for 14 consecutive days. They found that learning analytics-based interventions had a positive effect on student engagement and participation.

Studies in the literature frequently used the flipped classroom method in anatomy lessons [21,43–46]. This is because the basic concepts in anatomy are suitable for the level of knowledge and comprehension and can be given to students as homework. In-class activities may be administered via the means of problem-based scenarios, case-based discussions, and question-answer techniques. Students may not use higher-level cognitive activities in out-of-class activities. A well-designed flipped anatomy course can provide a deeper learning process via in-class activities. However, the prediction of course success based on student interactions and quiz scores in flipped classrooms is a subject that requires further investigation as it has the potential to aid in the prevention of student failure. Thus, this study aims to create a predictive model to identify at-risk students in an online flipped anatomy course. The study aims to answer two research questions: (1) Which supervised learning technique can predict at-risk students in an online flipped anatomy

course? (2) What is the classification accuracy of the best algorithm for predicting at-risk students in an online flipped anatomy course?

2. Materials and Methods

2.1. Participants

This study aimed to predict at-risk students in a flipped classroom. Thus, the study included pharmacy students enrolled in an anatomy course during the 2020–2021 academic year; 69 out of 75 first-year pharmacy students took part in the study. Every participant in this study volunteered to participate and their log data were collected in the Moodle LMS. The researchers then combined exam scores with the pre-processed log data.

2.2. Flipped Anatomy Course

We implemented this study as a part of a compulsory anatomy course (PHM101) for undergraduate pharmacy students who are studying at a public university in Asia. Before the application, an exam was applied to determine the students' prior knowledge within the scope of the course. This exam aimed to measure the basic anatomy concepts given at the beginning of the course as well as the knowledge regarding related subjects before their enrollment in the course. The data collection procedure was completed within a six-week period of the course. Details of the course schedule and topics have been provided in the table.

The course was delivered fully online via the flipped classroom method during the COVID-19 lockdown period. The instructor shared educational videos and resources related to course topics online on LMS. Students were expected to prepare for the relevant resources during the period between the sharing date of the resources and the date of the lesson. There are time intervals in the "student preparation" column in Table 1. At the beginning of the lesson, quizzes were given to measure student preparation. Students' misconceptions and errors in the quizzes were discussed; discussions and question-answer techniques were used during the in-class (synchronous virtual classroom lessons) activities.

Table 1. Data collection weeks, exams, in-class activities, and variable names.

Weeks	Course Topics	Student Preparation	In-Class Exam	In-Class Activities	Variables Names
-	Prior Knowledge Exam: 28 November 2020			Prior	
1	Urinary System (US)	December 1–7	Quiz US	Case Studies and Discussions	w1_eng & exam1
2	Reproductive System (RS)	December 8–14	Quiz RS	Case Studies and Discussions	w2_eng & exam2
3	Nervous System (NS)	December 15–22	Quiz NS	Case Studies and Discussions	w3_eng & exam3
4	Spinal Cord and Spinal Plexuses (SCSP)	December 23–29	Quiz SCSP	Case Studies and Discussions	w4_eng & exam4
5	Cranial Nerves and Autonomic Nervous System (CNAN)	December 30–January 9	Quiz CNAN	Case Studies and Discussions	w5_eng & exam5
6	Final Exam: 10 January 2021			final (target variable)	

2.3. Variables

In this study, we exported Moodle raw data as an Excel file comprising 119,517 lines with 9 columns: time, username, affected user, event context, component, event name, description, origin, and IP address. We used the MoodleMiner tool (<http://bote2.hacettepe.edu.tr/moodleminer/>; accessed on 25 January 2022) to transform the raw dataset into a transactional format in which the students were placed as unique rows and their interaction

variables were placed as columns to be analyzed [47]. The raw data comprised 34,277 lines, including every students' interactions with the course contents. After pre-processing, seven variables were exported; these are given in Table 2.

Table 2. Pre-processed Moodle interaction variables and their descriptions.

No	Interaction Variable	Description
1	n_session:	The number of sessions by the student
2	n_ShortSession	The number of short sessions by the student
3	d_Time:	The total time the student has spent on the Moodle LMS
4	n_UIQUEDay	The number of unique days logged in by the student
5	n_TotalAction	The number of total activities
6	n_CourseView	The number of course (Anatomy) views
7	n_ResourceView	The number of course resource views

MoodleMiner calculates the “engagement score” by using these seven variables. As a pre-process, the tool calculates percentile rank by averaging seven variables. Thus, the calculated value is called the “engagement score”, and it also provides percentile values. For instance, if a student's engagement score is 90, this also means that their activity is higher than the 90% percent of other students.

In this study, we calculated the “engagement score” for each week of data collection and merged them with students' preparations (in-class quiz scores) for each week (Table 1, variables column). We aimed to create a predictive model for the final score (target variable/dependent variable) by using eleven independent variables (prior, exam1, exam2, exam3, exam4, exam5, w1_eng, w2_eng, w3_eng, w4_eng, and w5_eng).

2.4. Data Analysis

To answer the research questions, we applied the knowledge discovery in databases process (KDD). The KDD is the development of good measures of interesting and discovered patterns [48]. According to Reference [49], data mining, defined as the extraction of interesting (non-trivial, implicit, previously unknown, and potentially useful) patterns or knowledge, plays an important role in the KDD process. Al Shalabi, Shaaban [50] emphasized that data mining, which seeks to discover unrecognized associations between data items in an existing database, is extracting valid, previously unseen or unknown, and comprehensible information from large databases. KDD consists of five steps: data selection, data pre-processing, data mining, pattern evaluation, and knowledge discovery [49].

We used the Orange data mining (DM) tool for each step of the KDD process (Figure 1). The tool is useful for visual programming and explorative data analysis and can be written in Python. Orange has multiple components known as widgets [51] which are dragged-dropped and connected to each other for the KDD procedures. In our predictive model, we applied the steps of the KDD process. These steps are explained in the following section.

2.4.1. Data Selection

We only combined students' Moodle interaction and exam scores. The log data usernames field was used to collect the interaction features of the students. This information was merged with students' exam scores, final scores, prior knowledge scores, and in-class quiz scores.

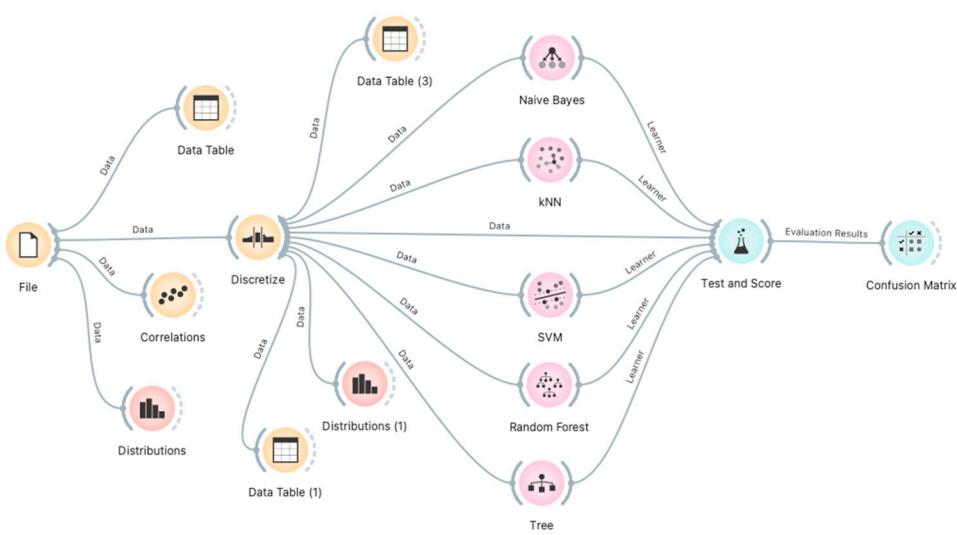


Figure 1. Data analyses were designed with the Orange DM tool.

2.4.2. Data Preprocessing

The discretization process is one of the most important data preprocessing tasks in data mining. Presently, many discretization methods, including Boolean reasoning, equal frequency binning, and entropy, are used [52]. All our continuous variables were transformed into two categories, which were divided based on cut-off scores. The cut-off scores for the engagement scores of each (w1_eng—w5_eng) week were decided at 50 points. This is because calculating these values between 0 and 100 based on the percentile rank and the median value (50) of this score can classify students with high engagement (1) and low engagement (0). Cut-off scores for exam results (prior and exam1–exam5) were decided as 60 points, which is also a passing grade for the courses at our university. For the exam, scores above 60 points were passed (1), while all other scores and the students who did not take the exams were coded as failed (0). The cut-off score for the final examination was also determined to be 60 points. Therefore, at-risk students were coded as 0 (below 60) and we coded the others as safe (1). Discretization makes it possible to apply supervised learning techniques to detect at-risk students via classification techniques [27,53].

2.4.3. Data Mining

This study aimed to create a predictive model for pharmacy students' final scores based on their Moodle interaction data and flipped classroom preparations. For this purpose, we applied five commonly used classification algorithms: k-nearest neighbors (kNN), decision trees (DT), naïve Bayes (NB), random forest (RF), and support vector machines (SVM).

The k-nearest neighbors (kNN) algorithm is also called the lazy learning algorithm and has been successfully applied in real applications with big data [33]. The algorithm uses several distance metrics, such as Euclidian and Manhattan distances, to solve the classification problem of a categorical dependent variable.

Decision trees (DT) are also used for classification problems for categorical dependent variables. Researchers use DT-based learning algorithms in prediction or rule induction solutions for educational problems. Investigating the effect of demographic characteristics such as regional, socio-economic, and educational levels, age, gender, and disability status on academic outcomes in the online learning environment [54] is an example of the use of decision trees in education.

Naïve Bayes (NB) classification aims to determine the class of dependent variable and is calculated by the probabilities of the classes [55]. Predicting students' graduation rates via their various achievement indicators [35] is an example of the usage of the NB algorithm in an educational setting.

The random forest (RF) algorithm is a statistical framework with a very high generalization accuracy and quick training times for classification tasks [56]. RF uses tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees, which comprise randomly selected dependent and independent variables in the forest [57]. The prediction of university students' academic success via course examination results [36] is an example of random forest classification approach.

Support vector machines (SVM) aim to identify a hyperplane which can meet classification requirements [58]. SVM can identify students' literacy skills (low, high) based on eye movement metrics during text reading [59] or classify learners in multi-class (high, average, or low) categories.

2.4.4. Performance Evaluation

Since our study aimed to classify students as either at-risk or safe and compare machine learning classification methods based on the accuracy of prediction models, we used performance metrics and confusion matrices for model evaluation. The confusion matrix is provided in Table 3, divided into four sections representing at-risk and safe situations in a binary classification problem.

Table 3. Structure of confusion matrix.

		Predicted Values	
		At-Risk	Safe
Actual Values	At-Risk	TP	FP
	Safe	FN	TN

TP (True Positive): The number of unsuccessful students that were classified as "at-risk".

TN (True Negative): The number of successful students that were classified as "safe".

FP (False Positive): The number of unsuccessful students that were classified as "safe".

FN (False Negative): The number of successful students that were classified "at-risk".

Classification Accuracy (CA): The ratio of the total number of correct predictions and the total number of predictions (1).

Recall: The percentage of total results correctly classified by the algorithm (2).

Precision: The ratio between the True Positives and all the Positives (3).

F-Measure: The Harmonic mean of the Precision and Recall (4).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$F - \text{measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

As the purpose of this study was to predict the students who would fail the final examination (at-risk), TP (True Positive) classification values will be considered. Thus, for the model evaluation, we focused on the models with a higher TP value. We also used 10-fold cross-validation during testing and scoring of the models created.

3. Results

3.1. Research Question 1: Which Supervised Learning Technique Can Predict At-Risk Students in an Online Flipped Anatomy Course?

We used independent variables to predict the target variable (final score). Model performance metrics have been provided in Table 4.

Table 4. Evaluation results of the algorithms.

Model	AUC	CA	F	Precision	Recall
RF	0.795	0.696	0.533	0.571	0.500
DT	0.794	0.696	0.588	0.556	0.625
NB	0.703	0.681	0.607	0.531	0.708
SVM	0.690	0.681	0.476	0.556	0.417
kNN	0.689	0.667	0.489	0.524	0.458

The target variable (final score) was classified into two categories with a cut-off score of 60. According to the classification algorithms model performances, the naïve Bayes algorithm performed the best when predicting the final scores in a flipped anatomy classroom model with a classification accuracy of 68% and a recall value of 71%.

3.2. What Is the Classification Accuracy of the Best Algorithm for Predicting At-Risk Students in an Online Flipped Anatomy Course?

In our model, the naïve Bayes algorithm performed well when detecting at-risk students. The confusion matrix of the algorithm is given in Table 5.

Table 5. Confusion matrix of naïve Bayes algorithm.

		Predicted		
		At-Risk	Safe	Total
Actual	At-Risk	17	7	24
	Safe	15	30	45
Total		32	37	69

According to Table 5, the naïve Bayes algorithm correctly classified 17 out of 24 (71%) unsuccessful students (TP value) before the final examination in a flipped anatomy course. The algorithm's interpretation of 7 students led to mistakes, as the model classified some successful students as "at-risk" by mistake (i.e., registering as a False Positive—FP).

4. Discussion

In flipped classrooms, students' pre-class activities, content–student interactions, and in-class quiz phases deliver more effective lectures. Tracking these activities via learning analytics enables the early prediction of at-risk students. This study aimed to create a predictive algorithm for final scores based on their interaction and preparation for a flipped anatomy course.

The results showed that the created model achieved at-risk prediction accuracy of 71%. This finding suggests that major determinants of student performance in flipped classrooms are students' prior knowledge scores, weekly engagement scores, and weekly quiz scores. We also found that the NB algorithm performed well when detecting at-risk students. In order to attain the best performance, it is often necessary to combine different data sources like user profiles and sequentially dynamic data from sources like interaction logs [60]. In our study, we only merged weekly engagement scores with weekly quiz scores of the students, which might explain why different studies resulted in better scores of different machine learning algorithms. On the other hand, in our study, the RF model also showed a high accuracy for the general prediction model. Another study

also found RF models with high prediction accuracy for both K-12 education and higher education [38]. Hung, Shelton [38] found that DNN and RF can accurately predict at-risk students for both K-12 and higher education in the range of 82.6–97%. As seen in the literature, combining many datasets can boost the prediction model's accuracy [27,38,61]. The RF technique works based on the decision trees [57]. Therefore, DT models can also show high performance for the prediction of at-risk students. Similarly, our study also showed high accuracies with DT models. Akçapınar, Altun [27] investigated the impact of various pre-processing applications on prediction performance by adjusting the cut-off scores. This is another important issue for prediction studies. In this study, we selected the cut-off scores based on the decision of pass grade for university, however, changing the cut-off scores can also affect the model's performance. Selecting norm-based criteria can help researchers generalize their models. They also found that the k-NN algorithm accurately predicted at-risk students at the end of the term with a rate of 89%. In addition, students who were at-risk at the end of term were predicted to a rate of 74% in 3 weeks. In our study, k-NN, which is also called “lazy learning”, showed sub-standard performance. In the literature, k-NN is usually used with continuous independent variables. That might explain why k-NN did not show high accuracy for prediction performance. Our study can predict “at-risk” students just before the final exam and this is the limitation of our study. Application of equal-frequency discretization can be the reason for higher accuracies in their predictive model for identifying at-risk students. The balanced distribution of the dependent variable might be a reason for the success of these methods. In addition, their study also included several metrics including participation in forums, writing posts, and creating tags. These actions might lead to higher classification performance. Setting a cut-off score with the passing grade of the university may change the performance of the algorithms. Previous studies mostly used median values or equal-frequency discretization for categorizing the continuous variables. That may lead to a differentiation in performing different machine learning algorithms. Saqr, Fors [29] used conventional statistical approaches to a medical course for the early detection of under-achieving students. The predictive model showed 63.5% accuracy and 53.9% of at-risk student prediction. According to our results, machine learning algorithms performed better when compared to statistical models such as logistic regression.

The sudden outbreak of the COVID-19 pandemic forced many institutions throughout the world to transition to emergency remote teaching [62,63]. Management of this crisis has depended on the participation of students and instructors in online learning processes and the effective use of online instructional materials and activities [64,65]. The biggest problem that emerged in these processes was the necessity of tracking students' interactions with these instructional materials, tasks, assignments, and their active participation by the instructors [66,67]. Although educators can directly observe the general condition and progress of the students through formative assessments and in-class activities, it becomes difficult to follow up, give feedback, and prevent student failure in online learning environments. After the COVID-19 pandemic, educational institutions carried out hybrid (blended) learning guided by the curriculum, and the predictive model performance can increase by combining different data sets in face-to-face and online learning environments to prevent student failure by identifying students at risk. We can also use this model as an early warning system in similar crises and lockdown situations in the future.

There are some contradictions in the literature in terms of combining students' demographics with interaction data. For example, Hung, Wang [39] stated that none of the students' demographics were selected as important predictors for identifying at-risk patterns. On the other hand, there is also significant correlation between online learners' performance and their demographic characteristics [54]. Combination of different datasets can increase the model's performance; however, we could not collect multiple data sources because of the lockdown process. Another limitation of this study is prevention of downloading the online course materials. We used scripts to prevent the students who are trying to download the video lecture assignments and online contents. However, some

students may have downloaded the videos or files via browser extensions or other plug-ins, resulting in poor model performance. According to Wilson, Watson [68], it is still unclear if learning analytics are useful for measuring, predicting, and improving student performance. Because learning analytics examines interaction data based on clickstream behavior, which assumes that online navigation is directly linked to learning processes, however, this assumption ignores some situations like disorientation or searching. Therefore, we received verbal feedback from the participants for the convenience of the instructional methods and materials, but it is still another limitation of our study. When using learning analytics in the field of health education, students' behavioral, emotional, and cognitive engagement should also be considered.

5. Conclusions

The flipped classroom approach is effective in terms of student participation in the classroom and providing higher-order thinking skills [8,12,16]. However, the most important limitation of the method is the difficulty in ensuring that students are prepared for class activities. A predictive method, which also has a higher prediction performance on summative exam scores, can be used for the early warning system. For this purpose, learning analytics can benefit instructors and researchers in identifying at-risk students based on their interactions with the learning management systems. Applying several pre-processing techniques and comparing the effect of different machine learning algorithms on classification performance can benefit the models created.

For further studies, the results of this study can be used as a traffic light project [69]. Misclassifications (FN and FP) can be represented with a yellow light, so the students need to be careful about the final examination. The green light can be shown to the students in the TN group, and motivational feedback can be provided. The TP group needs to receive a red light and must put in the effort to engage with the content, and they might also need to solve the quiz test after an individual study period. In addition, further studies also can combine “no-engagement” alerts during the online anatomy course applications. These kinds of interventions may also prevent the failure of the students.

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

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Ethics Committee of Ankara University Faculty of Medicine (protocol code: 2021000092-1, date of approval: 7 July 2021) for studies involving humans.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, A.B., upon reasonable request.

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

References

1. Abeysekera, L.; Dawson, P. Motivation and cognitive load in the flipped classroom: Definition, rationale and a call for research. *High. Educ. Res. Dev.* **2014**, *34*, 1–14. [[CrossRef](#)]
2. Hawks, S.J. The flipped classroom: Now or never? *AANA J.* **2014**, *82*, 264–269. [[PubMed](#)]
3. Tucker, B. The flipped classroom. *Educ. Next* **2012**, *12*, 82–83.
4. Rotellar, C.; Cain, J. Research, perspectives and recommendations on implementing the flipped classroom. *Am. J. Pharm. Educ.* **2016**, *80*, 34. [[CrossRef](#)]
5. Limniou, M.; Schermbrucker, I.; Lyons, M. Traditional and flipped classroom approaches delivered by two different teachers: The student perspective. *Educ. Inf. Technol.* **2018**, *23*, 797–817. [[CrossRef](#)]

6. Østerlie, O. Flipped learning in physical education: Why and how? In *Physical Education and New Technologies*; Croatian Kinesiology Association: Zagreb, Croatia, 2016; pp. 166–176.
7. Unal, Z.; Unal, A. Comparison of student performance, student perception and teacher satisfaction with traditional versus flipped classroom models. *Int. J. Instr.* **2017**, *10*, 145–164. [[CrossRef](#)]
8. Akçayır, G.; Akçayır, M. The flipped classroom: A review of its advantages and challenges. *Comput. Educ.* **2018**, *126*, 334–345. [[CrossRef](#)]
9. Brewer, R.; Movahedazarhoughligh, S. Successful stories and conflicts: A literature review on the effectiveness of flipped learning in higher education. *J. Comput. Assist. Learn.* **2018**, *34*, 409–416. [[CrossRef](#)]
10. Namaziandost, E.; Çakmak, F. An account of EFL learners' self-efficacy and gender in the Flipped Classroom Model. *Educ. Inf. Technol.* **2020**, *25*, 4041–4055. [[CrossRef](#)]
11. Cresap, L. Preparing university students for flipped learning. In *Blended Learning: Concepts, Methodologies, Tools, and Applications*; IGI Global: Hershey, PA, USA, 2017; pp. 1510–1531.
12. DeLozier, S.J.; Rhodes, M.G. Flipped classrooms: A review of key ideas and recommendations for practice. *Educ. Psychol. Rev.* **2017**, *29*, 141–151. [[CrossRef](#)]
13. Sailer, M.; Sailer, M. Gamification of in-class activities in flipped classroom lectures. *Br. J. Educ. Technol.* **2021**, *52*, 75–90. [[CrossRef](#)]
14. McWhirter, N.; Shealy, T. Case-based flipped classroom approach to teach sustainable infrastructure and decision-making. *Int. J. Constr. Educ. Res.* **2020**, *16*, 3–23. [[CrossRef](#)]
15. Oliván-Blázquez, B.; Aguilar-Latorre, A.; Gascón-Santos, S.; Gómez-Poyato, M.J.; Valero-Errazu, D.; Magallón-Botaya, R.; Heah, R.; Porroche-Escudero, A. Comparing the use of flipped classroom in combination with problem-based learning or with case-based learning for improving academic performance and satisfaction. *Act. Learn. High. Educ.* **2022**. [[CrossRef](#)]
16. Tawfik, A.A.; Lilly, C. Using a flipped classroom approach to support problem-based learning. *Technol. Knowl. Learn.* **2015**, *20*, 299–315. [[CrossRef](#)]
17. Bohaty, B.S.; Redford, G.J.; Gadbury-Amyot, C.C. Flipping the classroom: Assessment of strategies to promote student-centered, self-directed learning in a dental school course in pediatric dentistry. *J. Dent. Educ.* **2016**, *80*, 1319–1327. [[CrossRef](#)]
18. Estrada, A.C.M.; Vera, J.G.; Ruiz, G.R.; Arrebola, I.A. Flipped classroom to improve university student centered learning and academic performance. *Soc. Sci.* **2019**, *8*, 315. [[CrossRef](#)]
19. Awidi, I.T.; Paynter, M. The impact of a flipped classroom approach on student learning experience. *Comput. Educ.* **2019**, *128*, 269–283. [[CrossRef](#)]
20. Zainuddin, Z. Students' learning performance and perceived motivation in gamified flipped-class instruction. *Comput. Educ.* **2018**, *126*, 75–88. [[CrossRef](#)]
21. Day, L.J. A gross anatomy flipped classroom effects performance, retention, higher-level thinking in lower performing students. *Anat. Sci. Educ.* **2018**, *11*, 565–574. [[CrossRef](#)]
22. Shatto, B.; L'Ecuyer, K.; Quinn, J. Retention of Content Utilizing a Flipped Classroom Approach. *Nurs. Educ. Perspect.* **2017**, *38*, 206–208. [[CrossRef](#)]
23. Alias, M.; Iksan, Z.H.; Karim, A.A.; Nawawi, A.M.H.M.; Nawawi, S.R.M. A novel approach in problem-solving skills using flipped classroom technique. *Creat. Educ.* **2020**, *11*, 38. [[CrossRef](#)]
24. Wen, A.S.; Zaid, N.M.; Harun, J. Enhancing students' ICT problem solving skills using flipped classroom model. In Proceedings of the 2016 IEEE 8th International Conference on Engineering Education (ICEED), Kuala Lumpur, Malaysia, 7–8 December 2016; IEEE: Piscataway, NJ, USA; pp. 187–192.
25. Latorre-Coscolluela, C.; Suárez, C.; Quiroga, S.; Sobradiel-Sierra, N.; Lozano-Blasco, R.; Rodríguez-Martínez, A. Flipped Classroom model before and during COVID-19: Using technology to develop 21st century skills. *Interact. Technol. Smart Educ.* **2021**, *18*, 189–204. [[CrossRef](#)]
26. Tang, T.; Abuhmaid, A.M.; Olaimat, M.; Oudat, D.M.; Aldhaeebi, M.; Bamanger, E. Efficiency of flipped classroom with online-based teaching under COVID-19. *Interact. Learn. Environ.* **2020**, *1*–12. [[CrossRef](#)]
27. Akçapınar, G.; Altun, A.; Aşkar, P. Using learning analytics to develop early-warning system for at-risk students. *Int. J. Educ. Technol. High. Educ.* **2019**, *16*, 40. [[CrossRef](#)]
28. Rubio-Fernández, A.; Muñoz-Merino, P.J.; Kloos, C.D. A learning analytics tool for the support of the flipped classroom. *Comput. Appl. Eng. Educ.* **2019**, *27*, 1168–1185. [[CrossRef](#)]
29. Saqr, M.; Fors, U.; Tedre, M. How learning analytics can early predict under-achieving students in a blended medical education course. *Med. Teach.* **2017**, *39*, 757–767. [[CrossRef](#)]
30. Foster, E.; Siddle, R. The effectiveness of learning analytics for identifying at-risk students in higher education. *Assess. Eval. High. Educ.* **2019**, *45*, 842–854. [[CrossRef](#)]
31. Scholes, V. The ethics of using learning analytics to categorize students on risk. *Educ. Technol. Res. Dev.* **2016**, *64*, 939–955. [[CrossRef](#)]
32. Sathya, R.; Abraham, A. Comparison of supervised and unsupervised learning algorithms for pattern classification. *Int. J. Adv. Res. Artif. Intell.* **2013**, *2*, 34–38. [[CrossRef](#)]
33. Deng, Z.; Zhu, X.; Cheng, D.; Zong, M.; Zhang, S. Efficient kNN classification algorithm for big data. *Neurocomputing* **2016**, *195*, 143–148. [[CrossRef](#)]

34. Perez, A.; Grandon, E.E.; Caniupan, M.; Vargas, G. Comparative Analysis of Prediction Techniques to Determine Student Dropout: Logistic Regression vs Decision Trees. In Proceedings of the 37th International Conference of the Chilean Computer Science Society (SCCC), Santiago, Chile, 5–9 November 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–8. [[CrossRef](#)]
35. Herlambang, A.D.; Wijoyo, S.H.; Rachmadi, A. Intelligent Computing System to Predict Vocational High School Student Learning Achievement Using Naïve Bayes Algorithm. *J. Inf. Technol. Comput. Sci.* **2019**, *4*, 15–25. [[CrossRef](#)]
36. Beaulac, C.; Rosenthal, J.S. Predicting University Students' Academic Success and Major Using Random Forests. *Res. High. Educ.* **2019**, *60*, 1048–1064. [[CrossRef](#)]
37. Burman, I.; Som, S. Predicting Students Academic Performance Using Support Vector Machine. In Proceedings of the 2019 Amity International Conference on Artificial Intelligence (AICAI), Dubai, United Arab Emirates, 4–6 February 2019; IEEE: Piscataway, NJ, USA. [[CrossRef](#)]
38. Hung, J.-L.; Shelton, B.E.; Yang, J.; Du, X. Improving Predictive Modeling for At-Risk Student Identification: A Multistage Approach. *IEEE Trans. Learn. Technol.* **2019**, *12*, 148–157. [[CrossRef](#)]
39. Hung, J.-L.; Wang, M.C.; Wang, S.; Abdelrasoul, M.; Li, Y.; He, W. Identifying At-Risk Students for Early Interventions—A Time-Series Clustering Approach. *IEEE Trans. Emerg. Top. Comput.* **2017**, *5*, 45–55. [[CrossRef](#)]
40. Arnold, K.E.; Pistilli, M.D. Course signals at Purdue: Using learning analytics to increase student success. Proceedings Of the 2nd International Conference on Learning Analytics and Knowledge (LAK '12), Vancouver, BC, Canada, 29 April–2 May 2012; Association for Computing Machinery: New York, NY, USA, 2012; pp. 267–270. [[CrossRef](#)]
41. Yu, T.; Jo, I.-H. Educational technology approach toward learning analytics: Relationship between student online behavior and learning performance in higher education. In Proceedings of the Fourth International Conference on Learning Analytics and Knowledge (LAK '14), Indianapolis, IN, USA, 24–28 March 2014; Association for Computing Machinery: New York, NY, USA; pp. 269–270. [[CrossRef](#)]
42. Kruse, A.; Pongsajapan, R. Student-centered learning analytics. *CNDLS Thought Pap.* **2012**, *1*, 98–112.
43. Chutinan, S.; Riedy, C.A.; Park, S.E. Student performance in a flipped classroom dental anatomy course. *Eur. J. Dent. Educ.* **2018**, *22*, e343–e349. [[CrossRef](#)]
44. El Sadik, A.; Al Abdulmonem, W. Improvement in student performance and perceptions through a flipped anatomy classroom: Shifting from passive traditional to active blended learning. *Anat. Sci. Educ.* **2021**, *14*, 482–490. [[CrossRef](#)]
45. Ferrer-Torregrosa, J.; Jiménez-Rodríguez, M.; Torralba-Estelles, J.; Garzón-Farinós, F.; Pérez-Bermejo, M.; Ehrling, N.F. Distance learning ects and flipped classroom in the anatomy learning: Comparative study of the use of augmented reality, video and notes. *BMC Med. Educ.* **2016**, *16*, 1–9. [[CrossRef](#)]
46. Yang, C.; Yang, X.; Yang, H.; Fan, Y. Flipped classroom combined with human anatomy web-based learning system shows promising effects in anatomy education. *Medicine* **2020**, *99*, e23096. [[CrossRef](#)]
47. Akçapınar, G.; Bayazit, A. MoodleMiner: Data Mining Analysis Tool for Moodle Learning Management System. *Elem. Educ. Online* **2019**, *18*, 406–415. [[CrossRef](#)]
48. Silberschatz, A.; Tuzhilin, A. What makes patterns interesting in knowledge discovery systems. *IEEE Trans. Knowl. Data Eng.* **1996**, *8*, 970–974. [[CrossRef](#)]
49. Han, J.; Kamber, M. *Data Mining: Concepts and Techniques*, 2nd ed.; University of Illinois at Urbana Champaign: Champaign, IL, USA; Morgan Kaufmann Publishers: Berlington, MA, USA, 2006; pp. 1–14.
50. Al Shalabi, L.; Shaaban, Z.; Kasasbeh, B. Data mining: A preprocessing engine. *J. Comput. Sci.* **2006**, *2*, 735–739. [[CrossRef](#)]
51. Demšar, J.; Zupan, B. Orange: Data mining fruitful and fun—a historical perspective. *Informatica* **2013**, *37*, 55–60.
52. Marzuki, Z.; Ahmad, F. Data mining discretization methods and performances. *Lung* **2012**, *3*, 57.
53. Marbouti, F.; Diefes-Dux, H.A.; Madhavan, K. Models for early prediction of at-risk students in a course using standards-based grading. *Comput. Educ.* **2016**, *103*, 1–15. [[CrossRef](#)]
54. Rizvi, S.; Rienties, B.; Khoja, S.A. The role of demographics in online learning; A decision tree based approach. *Comput. Educ.* **2019**, *137*, 32–47. [[CrossRef](#)]
55. Islam, M.J.; Wu, Q.M.J.; Ahmadi, M.; SidAhmed, M.A. Investigating the performance of naive-bayes classifiers and K-nearest neighbor classifiers. In Proceedings of the International Conference on Convergence Information Technology (ICCIT 2007), Gwangju, Korea, 21–23 November 2007; IEEE: Piscataway, NJ, USA, 2007; pp. 1541–1546. [[CrossRef](#)]
56. Khan, R.; Hanbury, A.; Stoettinger, J. Skin detection: A random forest approach. In Proceedings of the IEEE International Conference on Image Processing, Hong Kong, China, 26–29 September 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 4613–4616. [[CrossRef](#)]
57. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
58. He, B.; Shi, Y.; Wan, Q.; Zhao, X. Prediction of customer attrition of commercial banks based on SVM model. *Procedia Comput. Sci.* **2014**, *31*, 423–430. [[CrossRef](#)]
59. Lou, Y.; Liu, Y.; Kaakinen, J.K.; Li, X. Using support vector machines to identify literacy skills: Evidence from eye movements. *Behav. Res. Methods* **2017**, *49*, 887–895. [[CrossRef](#)]
60. Qiao, C.; Hu, X. A Joint Neural Network Model for Combining Heterogeneous User Data Sources: An Example of At-Risk Student Prediction. *J. Assoc. Inf. Sci. Technol.* **2019**, *71*, 1192–1204. [[CrossRef](#)]
61. Doherty, I.; Sharma, N.; Harbutt, D. Contemporary and future eLearning trends in medical education. *Med. Teach.* **2015**, *37*, 1–3. [[CrossRef](#)]

62. Stewart, W.H. A global crash-course in teaching and learning online: A thematic review of empirical Emergency Remote Teaching (ERT) studies in higher education during Year 1 of COVID-19. *Open Prax.* **2021**, *13*, 89–102. [[CrossRef](#)]
63. Bond, M.; Bedenlier, S.; Marín, V.I.; Händel, M. Emergency remote teaching in higher education: Mapping the first global online semester. *Int. J. Educ. Technol. High. Educ.* **2021**, *18*, 1–24. [[CrossRef](#)]
64. Mahmood, S. Instructional strategies for online teaching in COVID-19 pandemic. *Hum. Behav. Emerg. Technol.* **2021**, *3*, 199–203. [[CrossRef](#)]
65. Walker, K.A.; Koralesky, K.E. Student and instructor perceptions of engagement after the rapid online transition of teaching due to COVID-19. *Nat. Sci. Educ.* **2021**, *50*, e20038. [[CrossRef](#)]
66. Wu, F.; Teets, T.S. Effects of the COVID-19 pandemic on student engagement in a general chemistry course. *J. Chem. Educ.* **2021**, *98*, 3633–3642. [[CrossRef](#)]
67. Tatiana, B.; Kobicheva, A.; Tokareva, E.; Mokhorov, D. The relationship between students' psychological security level, academic engagement and performance variables in the digital educational environment. *Educ. Inf. Technol.* **2022**, *1*–15. [[CrossRef](#)]
68. Wilson, A.; Watson, C.; Thompson, T.L.; Drew, V.; Doyle, S. Learning analytics: Challenges and limitations. *Teach. High. Educ.* **2017**, *22*, 991–1007. [[CrossRef](#)]
69. Fritz, J. Using analytics to nudge student responsibility for learning. *New Dir. High. Educ.* **2017**, *2017*, 65–75. [[CrossRef](#)]