International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 11, No.1 (Mar-2022)

https://dx.doi.org/10.12785/ijcds/1101184

Identifying "At-Risk" Students: An AI-based Prediction Approach

Ghazanfar Latif [1, 2*], Runna Alghazo [3], Maura A. E. Pilotti [3], Ghassen Ben Brahim [4]

[1]Department of Computer Science, Prince Mohammad bin Fahd University, Saudi Arabia.
[2]Université du Québec à Chicoutimi, 555 boulevard de l'Université, , Quebec, Canada.
[3]College of Sciences and Human Studies, Prince Mohammad bin Fahd University, Saudi Arabia.
[4]Department of Information Technology, Prince Mohammad bin Fahd University, Saudi Arabia.

** Corresponding Email address: glatif@pmu.edu.sa*

Received 26 Jan. 2021, Revised 31 Jan. 2022, Accepted 2 Mar. 2022, Published 31 Mar. 2022

Abstract: Student retention is of the utmost importance to higher education institutions. It is a metric used by legislators, accreditation agencies, and governing bodies. Providing students with remedial assistance at the right time has often proven an effective method for student retention. Identifying students that require this type of support is usually cumbersome though. A variety of stakeholders, such as educators, counselors, advisors, and other staff members, may have to be involved in identifying students who are "at-risk". Following recent developments in machine learning algorithms, automated systems may be developed to predict students' performance and refer students to remedial instruction. This paper proposes the utilization of Artificial Intelligence (AI) algorithms to predict a student's grades in a university course at any given semester based on the initial performance of the student in a combination of course assessment tools, such as quizzes, assignments, and tests. The prediction model is based on a dataset of real cases compiled from courses at a private university in Saudi Arabia. The model, however, is general enough to be applied to any course at universities around the world. The prediction classifiers used in this study are Random Forest (RF), Sequential Minimal Optimization (SMO), Linear Regression (LR), Additive Regression (AR), and Multilayer Perceptron (MLP). Various metrics are employed to measure the prediction models' performance and assess the accuracy and validity of the proposed AI-based algorithms. Results indicate that the best classifier for predicting the final exam grade is the SMO, with a minimum mean absolute error of 2.350. The best prediction classifier for the midterm exam is LR with a minimum mean absolute error of 1.978. As tools for the early identification of students' difficulties in the particular courses in which they are enrolled, the effectiveness of the proposed models is discussed.

Keywords: Grade Prediction; Artificial Neural Networks; Artificial Intelligence; Students at-Risk; Machine Learning

1. Introduction

Learning is generally defined as the process through which individuals acquire and retain information as well as skills [1]. The need to understand the key properties of this process, including the antecedents that can promote or hinder it, has led to a treasure chest of empirical studies that illustrate how the human mind may approach the task of learning, as well as to theories of the different ways such a task can be successfully executed [2]. Understanding learning is undoubtedly a complex enterprise because of its many facets. It is closely intertwined with an array of human abilities (e.g., perception, attention, motivation, memory, decision-making, and language), individual difference variables, and assessment modes. Because learning is essentially latent until it emerges in task performance, the sensitivity of assessment tools to the acquired information or skills is critical. If ignored, it is a potential source of measurement bias.

Yet, grasping the properties of this multifaceted and opaque phenomenon and the conditions under which it produces academic success is a key enterprise for a multitude of individuals, including learners, educators, administrators, and policymakers, as successful learning in any given domain of knowledge and practice is necessary for optimal performance in that domain. Thus, it is not surprising that scientists from different academic disciplines have focused their efforts on algorithms that can predict learners' performance. In the unending loop of

E-mail: glatif@pmu.edu.sa

information sharing that defines scientific exchanges, algorithms are built from knowledge acquired from theoretical speculations and empirical investigations. They are then tested through computer simulations whose outcomes are used to refine such knowledge. In essence, algorithms offer valuable insights into the process of learning and its observable outcomes, which can then be applied to the identification of at-risk students and to remedial actions intended to minimize the likelihood of withdrawal.

The search for the ideal algorithm is a prolific area of scholarly research whose main challenge is, first and foremost, the selection of valid and reliable predictors (i.e., variables that substantially contribute to individual differences in performance) [3]. Another challenge is the optimal generality of the algorithm in relation to the cognitive demands that diverse tasks place on equally diverse learners. The latter is often addressed by relying on large data sets of online learners [4]. The hope is of developing a versatile tool that can overcome the idiosyncrasies of smaller sets which may bias the relevance of different predictors and thus offer a distorted picture of their impact on measurable performance.

In the quest for a versatile algorithm with high predictive validity, some issues have remained in the background. To this end, it is noteworthy to point out that the ways with which human beings acquire information and skills vary from those requiring little or no effort at all, such as observational learning and simple associative learning modes (e.g., classical and operant conditioning), to those that are engaging and demanding of attentional resources, such as active learning. The medium may include traditional face-to-face, online, or hybrid forms [5]. Because learning refers to a mental state whose outcome needs to be measured through sensitive performance assessment tools, modes of assessment also vary, from those that require reiteration to those that demand application, analysis, and elaboration, as illustrated by Bloom's taxonomy [6, 7]. Similarly, individual differences in acquisition and retention of information are not merely eccentricities that distinguish one learner from another, but they are likely to reflect uniform patterns of preferences often shaped by cultural practices.

The present research, as others before it, aims to test the efficiency of different algorithms in predicting students' test performance measured as a continuous variable. It is grounded on the recognition that researchers' emphasis on overall test performance in formal educational settings, and their appreciation of large data sets, have often led variations in learning and assessment modes to being placed on the backburner. Ours is an attempt to contribute to the extant literature by taking a slightly different approach. Specifically, instead of the customary predictors, such as grade point average (GPA), engagement, and demographic measures, the main predictor for testing is students' initial performance as measured by particular types of course assessment tools, such as quizzes and assignments [8]. Furthermore, our research relies on a preliminary data set of grades produced by undergraduate students from the Middle East, thereby focusing on a population largely neglected in the literature on learning. The educational background of these students can be described as instructor-centered, a pedagogical approach that reinforces from an early age the exact reproduction of information rather than its elaboration [9]. Students are enrolled in engineering courses in which assessment entails cognitive operations that are above the reiteration of concepts and include application, analysis, and elaboration. Thus, these students' academic success relies on their ability to alter well-established learning habits quickly to adjust to the demands of higher education courses.

The specific aim of the present pilot study is to design a grade prediction system that is capable of forecasting students' midterm exam grades and final exam grades in a particular course through using Artificial Neural Networks (ANNs). In the following sections, we describe how artificial intelligence- (AI-) based algorithms can be applied to a dataset including different assessment tools. The ultimate goal of our approach is to aid the identification of students at risk of failure so that prompt and effective remedial interventions can be made available. The inclusion of other variables, such as a student's academic history, covering past grades and missed assessment activities in prior courses, is not considered in the current work. The reason is that we advocate an approach that is both minimalist and practical, which is based on the recognition that selected initial performance records are readily available to educators and that educators can identify the precise difficulties experienced by a student before they become more substantial and unmanageable to make remediation and individualized instruction ineffective. However, educators often teach large numbers of students and are expected to satisfy multiple professional commitments, which make the identification of at-risk learners in their classes, informed guidance, and targeted assistance challenging. Thus, we examine whether grades in course activities that precede the midterm or final exam can be useful predictors of either exam performance, independently of other more complex and less accessible contributions (e.g., a student's performance history). Obviously, we recognize that such contributions may add to the predictability of our minimalist approach. Yet, we aim to assess whether readily accessible information in a course is sufficient for an accurate prediction of students' performance on key assessment activities in that course. We believe that a tool that can draw educators' attention to learners with difficulties and that relies on readily accessible information is of no negligible utility. Midterm and final exams are selected as key assessment activities for two reasons: (a) They are common types of assessment in university-level courses around the world, thereby increasing the generalizability of our proposed approach. (b) Each marks a specific timeframe of learning activities in such courses, and thus the amount of assessment data available for prediction. The relative sensitivity of different algorithms to the amount of available data allows us to test whether it can determine their relative effectiveness as predictors.

The novelty of our approach is that it proposes the testing of various ANN classifiers for the prediction of midterm exam grades and final exam grades hypothetically in any of the courses in which a student enrolls. As a result of testing, the selection of ideal classifiers may then translate into a software package given to educators for the accurate and effortless identification of at-risk students in their courses. Albeit our research concerns grades of courses at a private university in Saudi Arabia, we believe its findings can generalize to other courses at the host institution and at other institutions of higher education as well if a sufficient number of cases is added to the dataset to ensure representation of such courses and institutions. In the dataset, though, each student's performance must be indexed by a collection of assessment outcomes throughout the course in which he/she is enrolled (i.e., quizzes, assignments, projects, and midterm and final exams).

In sum, the task of identifying students at-risk or even those who require temporary guidance to succeed in a course or degree plan is a cumbersome one. Educators are usually teaching multiple sections. Thus, even irrespective of the size of individual sections, each semester educators are likely to deal with a large number of students. Thus, it becomes imperative to find an alternative approach for identifying students at-risk of failure or even students who are experiencing challenges in selected courses. Unfortunately, students who experience difficulties tend not to report them, thereby gradually accumulating withdrawals and more substantial failures. The advent of AI offers the possibility of a seamless and timely identification of students either at-risk or merely requiring support in a certain course without much effort exerted by the educator. Although the current work involves a newly developed dataset of modest size, it is expected that its size will be increased to include additional courses in a variety of disciplines, as well as that it will be made available to other researchers for further testing.

2. Literature Review

Our research stems from the field of learning analytics whose goal is the development of algorithms that successfully predict students' test performance [10]. Due to its useful applications, the field of learning analytics is often cited as one of the key emerging trends in higher education [11, 12].

By and large, learning analytics focuses on predicting students' test performance from a defined set of individual difference factors, whose selection is guided by theoretical considerations, empirical evidence, as well as test simulations. The evidence produced by the testing of different algorithms often supports established empirical knowledge in the fields of neuroscience and education. For instance, the belief that past behavior is the best predictor of future behavior is a well-known truism in the fields of neuroscience and education. Not surprisingly, one of the most impactful factors in predicting test outcomes in research involving learning analytics has been reported to be past performance [3, 8].

The field of learning analytics relies heavily on identifying students' features that are not only predictive of performance, as either a global indicator (e.g., GPA as a single marker of a student's academic history) or as a contextualized indicator (e.g., a course grade), but also convenient to use by the purported beneficiaries (e.g., educators). Yet, not all students' features may receive the same attention as some may be more readily available and easier to measure than others. For instance, behavioral evidence exists that missed tests, often conceptualized as exam avoidance, are overall indicators that a given student is facing academic challenges [13]. Evidence also exists that students' missed test questions can be used to improve future performance through remedial actions [14]. Notwithstanding their implied utility, research devoted to learning analytics has yet to examine in depth the relative contribution of these factors to test performance as a function of time enrolled in a course. The reason may be the difficulty of measurement, as the real sources of a targeted behavior are not always obtainable, the unavailability of adequate datasets, or other practical considerations related to the development of models that educators can readily use [15].

The quest for the ideal algorithm to predict test performance is a problem that may involve the treatment of performance as either a binary variable (success versus failure), thereby requiring a coarse classification of students, or a continuous variable (actual marks within an established range), thereby capturing subtle differences among students. In addition to the choice of the format of measures of students' performance, researchers may select one of three categories of extant approaches: (a) Similarity-based approaches, which attempt to identify patterns of features that are close matches, and thus focus on finding students that are similar to each other. Notable instances are K-Nearest Neighbours (k-NN) and distance-weighted k-NN [16, 17]. (b) Model-based approaches, which are driven to uncover and exploit correlations between the properties of the sample and the selected performance outcome variable. Instances are Support Vector Machines (SVM) [18], Artificial Neural Networks (ANN) [19], linear or logistic regression [20] and Decision tree [21]. (c) Probabilistic approaches, which exploit the probability distribution characteristics in a dataset to determine either the extent to which they can predict whether a student belongs to a certain class (success versus failure) or his/her academic outcomes defined on a probabilistic continuum [22]. Exemplars of this approach are Naïve Bayes (also called Maximum A Posteriori, MAP) and Bayesian linear regression.

The relative effectiveness of each approach in predicting students' performance is entangled with variations in the identification of the actual students' features that are used for simulation and testing. The approach proposed below addresses the need of educators for measures of students' features that are readily available and sufficiently predictive of students' performance.

3. Proposed System

The model of the proposed grade-prediction system is shown in Figure 1. Educators (e.g., professors, instructors, and teaching assistants) are expected to enter students' grades in the various assessment activities of the courses in which they are enrolled. Assessment activities may include assignments, quizzes, exams, and projects. Data entry, which is done for the various courses taken by each student, may start at the freshman year and continue over the various years of his/her educational journey. The combined grades entered by educators produce a sizable dataset that can be used by machine learning algorithms to predict a student's midterm and final exam grades in a variety of courses based on the information available from other students with a similar profile in the dataset. The benefit of the proposed system becomes intuitively apparent, as predicted grades immediately flag at-risk students so that they can be given the appropriate support to succeed in the courses in which they are enrolled and ultimately complete their degrees. In the proposed system, the dataset's attributes are fed to well-known machine learning algorithms with data divided into 90% training and 10% testing. The selected algorithms are Sequential Minimal Optimization (SMO), Linear Regression (LR), Random Forest (RF), Multilayer Perceptron (MLP), and Additive Regression (AR). Thus, the key research question is whether an algorithm exists that is a more accurate predictor than the others based on the performance information available during the first half of the semester or up to the final exam.

A. Work Accomplished: Students' Grades Dataset

A dataset was developed including students' grades in various courses from a private university in Saudi Arabia. The dataset comprised 250 students. Courses with the same assessment tools and grading schemes were chosen to standardize the grading scheme for each assessment item. To this end, undergraduate courses from the College of Computer Engineering and Science were included. The dataset covered a period of three semesters with different students each semester. The dataset described the max grade for each assessment tool (Assignment 1, Assignment 2, Quiz 1, etc.). Furthermore, it provided the average, median, minimum value, and maximum values for all students (see Table 1).

Figure 2 shows a measure of the symmetry of the grade distribution (Skewness). Figure 3 shows a measure of the peakedness of the grade distribution (Kurtosis).

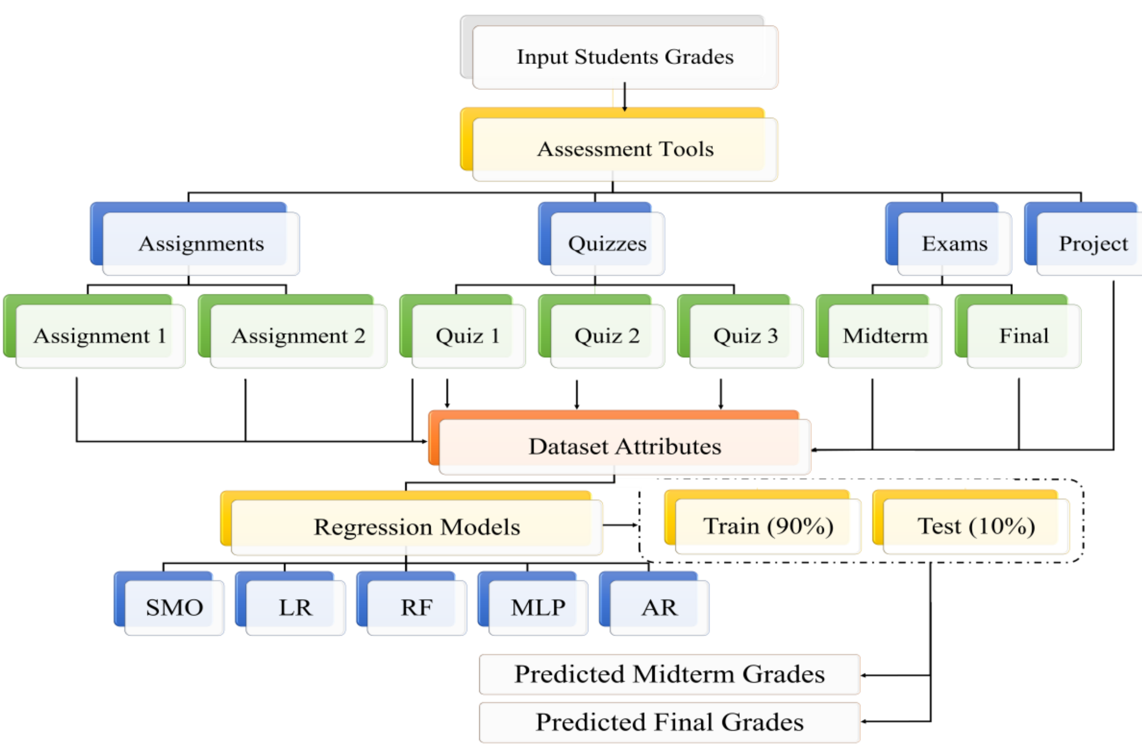


Figure 1. Proposed system model for the prediction of students' grades

Table 1. Descriptive statistics of the dataset of 250 students' grades

	A1	A2	Q1	Q2	Q3	Mid	Project	Final
Total	7.5	7.5	5	5	5	20	20	30
Avg.	6.7	6.7	4.0	4.1	4.1	14.5	17.6	19.7
Median	7.1	7.5	4.2	4.5	5	14.7	18	19.5
Min	0.5	3.3	0.7	0.5	0.2	6.5	6.8	8.5
Max	7.5	7.5	5	5	5	20	23.5	30

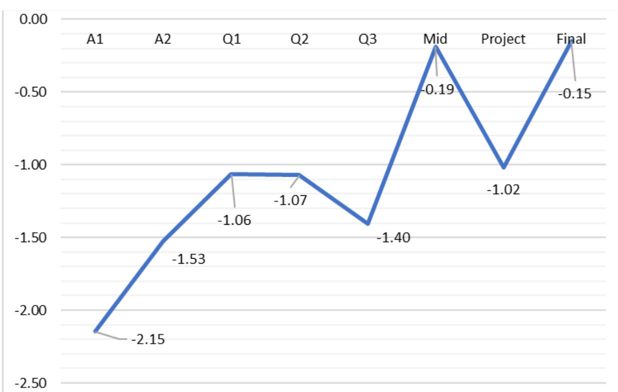


Figure 2. The skewness of the students' grade distribution

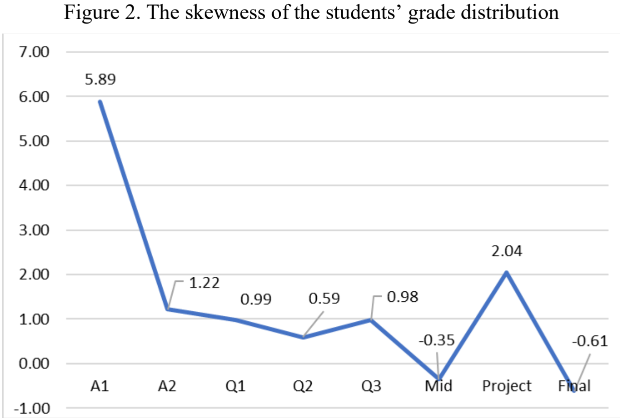


Figure 3. The Kurtosis of the students' grade distribution

B. Work Accomplished: The Selection of Machine Learning Algorithms

Machine Learning (ML) models along with statistical methods have often been used for performance predictions in various application domains. Statistical models mainly focus on studying and describing the quantitative relationship between problem input variables and hence model dependency using mathematical equations to make predictions. Instead, ML-based models attempt to explore various algorithms that reason from external datasets describing problem instances (i.e., training data) to produce general hypotheses, which will eventually serve to make predictions about future events [23]. For various problems, ML-based models have shown their effectiveness in handling predictions in case of problems with many domain variables and large databases.

In the present research, we examined the accuracy of students' exam predictions using ML-based models. Five models were considered, including (1) Random Forest (RF), (2) Sequential Minimal Optimization (SMO), (3) Linear Regression (LR), (4) Additive Regression (AR), (5) Multilayer Perceptron (MLP). These models are briefly described next.

The *Random Forest* (RF) technique can handle large datasets with a considerable number of attributes, while it weighs the importance of each of the problem features. It is prone to noise, outliers, and overfitting [24]. Contrary to other techniques, the RF relies on a combination of classification techniques contributing to a single vote during the classification process.

The *Sequential Minimal Optimization* (SMO) is a variation of the support vector machine (SVM) classification algorithm. It is designed to handle applications of a large-scale nature and develop efficient solutions to complex Quadratic Program (QP) optimization problems [25]. SMO attempts to divide the initial large-scale optimization problem into a set of sub-QP, which are, in turn, solved analytically. SMO avoids draining the memory space because no extra matrix storage is needed.

Linear Regression (LR) is a supervised ML technique used to forecast data of a continuous format [26]. In the case of a multi-variable problem (similar to what we have in hand), the multi-variable linear equation model is defined as follows:

$$f(x_1, x_2, \dots, x_n) = w_1.x_1 + w_2.x_2 + \dots + w_n.x_n$$

Where x_i and w_i represent attributes and weights, respectively. The LR-based model attempts to minimize a cost function capturing the difference between observed values and predicted values until the weight constants are optimized. The cost function is as follows:

$$\frac{1}{2N}\sum_{i=0}^{n}(y_i - (w_1.x_1 + w_2.x_2 + \dots + w_n.x_n))^2 \qquad (1)$$

Where y_i and N represent the predicted value and data size, respectively.

Additive Regression (AR) provides more flexibility in handling the dependency between variables. Contrary to the traditional linear regression model, where variables are assumed to have a linear relationship, the AR model uses an arbitrary function to capture the marginal and additive effect of a variable. Such a model with additive features is expected to decrease the estimated variance between actual and predicted values [27]. The predicted and the expected values are related through the following function:

$$g\big(f(Y)\big) = \beta_0 + \sum_{i}^{m} f_i(x_i) \qquad (2)$$

Where, β_0 is a constant and f_i is a smoothing function. Such flexibility in terms of the actual relationship between prediction and response is assumed to lead to an improved fit to input data compared to that exhibited by traditional models.

Multilayer Perceptron (MLP) is a supervised learning-based approach [28]. It is based on the concept of perceptron in Neural Networks, which is capable of generating a single output based on multidimensional data inputs through the exploitation of their linear (non-linear in some instances) relationships along with their corresponding weight as follows:

$$y = \alpha \sum_{i=1}^{n} w_i x_i + \beta \qquad (3)$$

Where w_i, x_i, β, and α are the weight, input variable, bias, and non-linear activation function, respectively. The MLP is composed of three or more layers of nodes. These include the input/output layer and one or many hidden layers. The training phase in the case of MLP consists of adjusting the model parameters (biases and weights) through a back and forth mechanism (feed-forward pass followed by back-forward pass) concerning the prediction error.

4. Experimental Results

This section describes the experimental results of the selected student performance prediction models. The models' performance was assessed by considering quantitative metrics, including (1) Mean Absolute Error (MAE), (2) Root Mean Squared Error (RMSE), (3) Relative Absolute Error (RAE), (4) Root Relative Squared Error (RRSE), and (5) Correlation coefficient (CC) [29].

The five prediction classifiers were evaluated using performance metrics to validate the accuracy of the proposed classifiers. The attributes used to predict midterm grades were the grades of Assignment 1, Assignment 2, Quiz 1, Quiz 2, and Quiz 3. The following attributes were used to predict final grades: Assignment 1, Assignment 2, Quiz 1, Quiz 2, Quiz 3, Project, and Midterm grades. The classifiers were evaluated to identify the one that achieved the best prediction accuracy. Table 2 shows the average performance metrics calculated for all classifiers using 90% data for training and 10% for testing. In the table, the prediction accuracy for the final exam grade of all classifiers is displayed in descending order.

The best prediction of the final exam grade was achieved using SMO with a correlation coefficient of .608 and a minimum mean absolute error of 2.350. The coefficient of determination (i.e., the percentage of variance accounted for by the model) was 36.966%. AR was the worst prediction classifier for the final exam grade with a correlation coefficient of .312 (coefficient of determination: 9.734%) and a mean absolute error of 3.187.

Table 2. Predictions of final exam grades using 90% data for training and 10% for testing with different algorithms

Algorithm	MAE	RMSE	RAE	RRSE	CC
SMO	2.350	2.887	75.291	73.630	0.608
LR	2.404	2.934	77.045	74.835	0.611
RF	2.842	3.535	91.070	90.149	0.458
MLP	2.936	3.680	94.081	93.845	0.315
AR	3.187	4.052	102.107	103.335	0.312

In the course of the validation process, a 10-fold cross-validation was performed. As shown in Table 3, it confirmed that the best prediction classifier for the final exam grade was SMO and the worst prediction classifier was AR.

Table 3. Predicted final exam grades using a 10-fold cross-validation

Algorithm	MAE	RMSE	RAE	RRSE	CC
SMO	2.599	3.167	80.428	79.952	0.635
RF	2.630	3.438	81.389	86.788	0.561
LR	2.674	3.373	82.745	85.143	0.570
AR	2.853	3.865	88.304	97.554	0.460
MLP	2.914	3.720	74.818	75.354	0.664

Table 4 shows the performance of prediction classifiers for the midterm exam in descending order, which was obtained using 90% of the data for training and 10% for testing. The table shows that the best prediction classifier for the midterm exam grade was LR with a correlation coefficient of .563 (coefficient of determination: 31.697%) and a minimum mean absolute error of 1.978, while the worst prediction classifier for the midterm exam was AR with a correlation coefficient of .286 (coefficient of determination: 8.180%) and mean absolute error of 3.013.

A 10-fold cross-validation was performed (see Table 5). Results of the cross-validation confirmed that the best classifier for predicting the midterm grades was LR and the worst is AR.

In Figure 4, the waveform of the actual grades is shown as well as the waveforms of all the five prediction classifiers. The figure allows one to visually compare the predicted final exam grades and the actual final exam grades. The waveform of the classifier that closely resembles the waveform of the actual grades is to be considered the best predictor while the one that is further away from the waveform of the actual grades is to be considered the worst predictor. The graph confirms that the best prediction classifier with the minimum error was the SMO, while the maximum error was exhibited by AR.

Table 4. Predicted midterm exam grades using 90% data for training and 10% for testing

Algorithm	MAE	RMSE	RAE	RRSE	CC
LR	1.978	2.327	80.052	78.769	0.563
SMO	2.285	2.840	92.486	96.143	0.416
RF	2.520	3.043	101.979	103.029	0.341
MLP	2.663	3.330	107.795	112.742	0.379
AR	3.013	3.507	121.949	118.734	0.286

Figure 5 shows a comparison between the predicted midterm exam grades and students' actual midterm exam grades. As explained above, the waveform of the classifier that closely resembles that of the actual values through visual inspection is deemed the best predictor. The graph clearly shows that the best prediction classifier with minimum error is LR, while the maximum error is displayed by AR.

Table 5. Predicted midterm exam grades using a 10-fold cross-validation

Algorithm	MAE	RMSE	RAE	RRSE	CC
LR	1.748	2.234	70.908	75.222	0.647
RF	1.798	2.276	70.424	75.724	0.656
SMO	1.819	2.307	71.277	76.763	0.647
MLP	1.824	2.447	71.439	81.435	0.631
AR	2.212	3.441	85.318	85.736	0.520

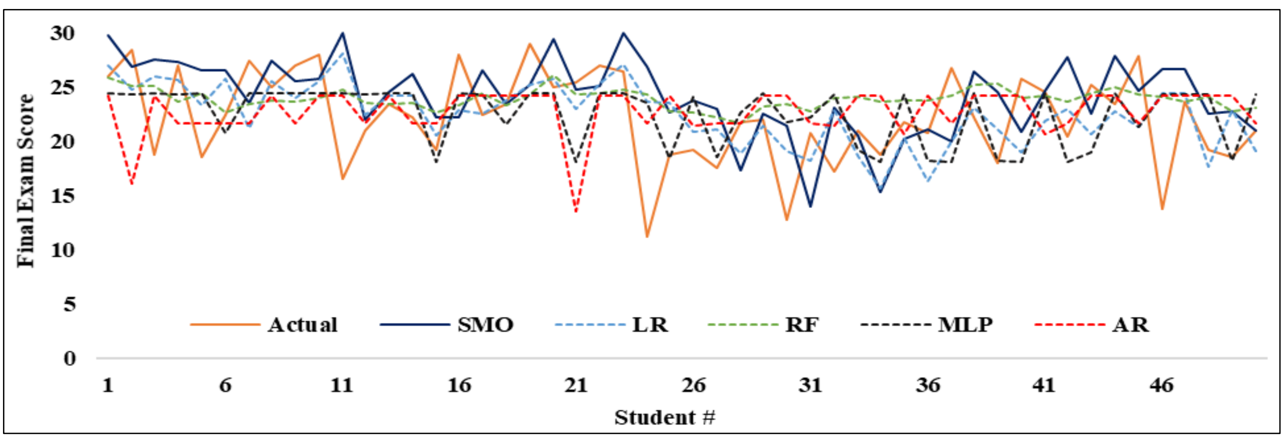


Figure 4. Comparison of final exam grade predictions using different machine learning algorithms

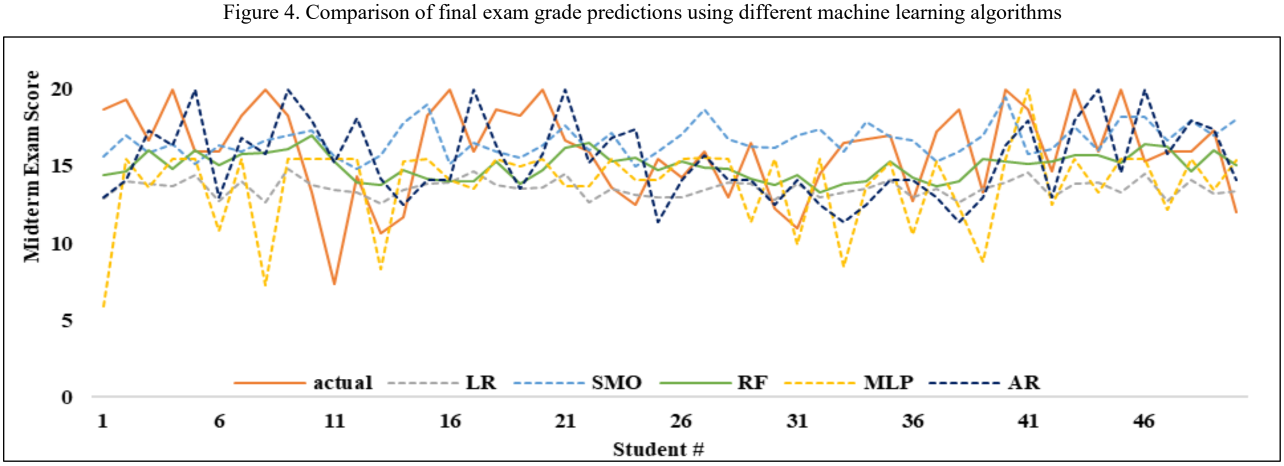


Figure 5. Comparison of midterm exam grade predictions using different machine learning algorithms

5. Analysis of Results

The current study aimed to test the effectiveness of AI-based prediction models using various ANN classifiers for students' midterm exams and final exams. The practical utility of an effective model is clear. It can provide educators and counselors a tool to identify students at-risk early in the semester. Early identification of difficulties experienced by a student, which are likely to increase in severity as the semester progresses, offers the opportunity for individualized assistance and guidance. It can also justify the referral of the student to support services for specialized tutoring. In the manuscript, we demonstrated that AI-based algorithms can be applied to datasets including different assessment tools (e.g., assignments, quizzes, midterm exams, and projects) to predict final exam grades. Furthermore, with a smaller input set, we demonstrated that AI-based algorithms can be used to predict midterm grades. Demonstrations relied on a dataset including grades of real courses from a private university in Saudi Arabia. The dataset consisted of the detailed grades of 250 students obtained in courses offered by the College of Computer Engineering and Science of the selected university. Five different prediction classifiers were used in the study: RF, SMO, LR, AR, and MLP. To measure the performance of the prediction models and thus the accuracy of the proposed AI-based algorithms, several measurement metrics were used: MAE, RMSE, RAE, RRSE, and CC.

Students' final exam grades were best predicted by the SMO algorithm, while midterm exam grades were best forecasted by the LR algorithm. The amount of available data might explain why the best prediction for the midterm exam and the best prediction for the final exam were achieved by different classifiers. Namely, the efficiency of classifiers changed as a function of the magnitude of the dataset. Thus, it is reasonable to conclude that as a dataset increases in size, the SMO may be the best choice. Alternatively, as the dataset shrinks, LR may become the most suitable choice. It is also reasonable to assume that the less-than-ideal accuracy of each model is due to the size of the current dataset, as it includes only the grades of 250 students.

6. CONCLUSION

Predictive models in learning analytics are intended to infer a single aspect of the data (i.e., the outcome variable, such as performance in a course or GPA) from some combination of other aspects of the data (predictor variables) [30]. Albeit the ability of a variable to predict performance is key, its accessibility to educators is no less important. Learning analytics can offer tools that can be predictive of students' performance as well as of practical use to assist educators in identifying at-risk students. Usually, class sizes are large, teaching loads are substantial, face-time to address academic difficulties is often scarce, and the varied professional demands of an academic job are heavy. In this context, it is hard for an educator to identify in a timely manner students who are facing difficulties in a given course, let alone give individualized attention and practical assistance to each student. The use of AI-based tools can lift the burden of early identification of at-risk students from educators by offering reliable pointers for informed interventions. In our research, we have demonstrated that such tools can be informed by students' initial performance in course activities. Of course, the development of a more comprehensive dataset may improve the prediction rate of the proposed models, and their generalizability to a variety of disciplines and localities.

A key objective of both educators and scientists in higher education is to improve student persistence [31]. Yet, in many countries, drop-out rates are still unbearably high [32, 33]. Research in AI, including ours, offers hope for the amelioration of this state of affairs. Namely, when used effectively, AI-based tools can help higher education institutions monitor learning and improve students' performance, thereby reducing dropout rates and increasing graduation numbers.

REFERENCES

[1] De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin & Review, 20*(4), 631-642.

[2] Anderson, T. (2016). Theories for learning with emerging technologies. In G. Veletsianos (Ed.), *Emergence and innovation in digital learning: Foundations and applications* (pp. 35-50). Athabasca University.

[3] Tomasevic, N., Gvozdenovic, N., & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers & Education, 143*, 103676.

[4] Kuzilek, J., Hlosta, M., & Zdrahal, Z. (2016). Open university learning analytics dataset. Scientific Data, 4, 170171.

[5] Tastle, W. J., White, B. A., & Shackleton, P. (2005). E-learning in higher education: The challenge, effort, and return on investment. *International Journal on E-learning, 4*(2), 241-251.

[6] Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives*. Cognitive domain. McKay.

[7] Bloom, B. S. (1976). *Human characteristics and school learning*. McGraw Hill.

[8] Minaei-Bidgoli, B., Kashy, D. A., Kortmeyer, G., & Punch, W. F. (2003). Predicting student performance: An application of data mining methods with an educational web-based system. *33rd Annual Frontiers in Education*, *1*, T2A–T18.

[9] Weimer, M. (2002). Learner-centered teaching: Five key changes to practice. John Wiley & Sons.

[10] Corrigan, O., Smeaton, A. F., Glynn, M., & Smyth, S. (2015). Using educational analytics to improve test performance. In *Design for Teaching and Learning in a Networked World* (pp. 42-55). Springer.

[11] Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality indicators for learning analytics. *Journal of Educational Technology & Society*, *17*(4), 117-132.

[12] Sclater, N., Peasgood, A., & Mullan, J. (2016). *Learning analytics in higher education*. London: Jisc.

[13] Jury, M., Smeding, A., Court, M., & Darnon, C. (2015). When first-generation students succeed at university: On the link between social class, academic performance, and performance-avoidance goals. *Contemporary Educational Psychology*, *41*, 25-36.

[14] Favero, T. G., & Hendricks, N. (2016). Student exam analysis (debriefing) promotes positive changes in exam preparation and learning. *Advances in Physiology Education*, *40*(3), 323-328.

[15] Zilvinskis, J., & Willis III, J. E. (2019). Learning Analytics in Higher Education: A Reflection. *InSight: A Journal of Scholarly Teaching*, *14*, 43-54.

[16] Latif, G., Iskandar, D. A., Alghazo, J., & Jaffar, A. (2018). Improving brain MR image classification for tumor segmentation using phase congruency. *Current Medical Imaging*, *14*(6), 914-922.

[17] Tanner, T., & Toivonen, H. (2010). Predicting and preventing student failure–using the k-nearest neighbour method to predict student performance in an online course environment. *International Journal of Learning Technology Archive, 5*(4), 356–377.

[18] Latif, G., Iskandar, D. A., Jaffar, A., & Butt, M. M. (2017). Multimodal brain tumor segmentation using neighboring image features. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *9*(2-9), 37-42.

[19] Zacharis, N. Z. (2016). Predicting student academic performance in blended learning using Artificial Neural Networks. *International Journal of Artificial Intelligence and Applications*, *7*(5), 17-29.

[20] Gauer, J. L., Wolff, J. M., & Jackson, J. B. (2016). Do MCAT scores predict USMLE scores? An analysis on 5 years of medical student data. *Medical Education Online*, *21*(1), 31795.

[21] Mesarić, J., & Šebalj, D. (2016). Decision trees for predicting the academic success of students. *Croatian Operational Research Review*, *7*(2), 367-388.

[22] Lucas, C. G., Griffiths, T. L., Williams, J. J., & Kalish, M. L. (2015). A rational model of function learning. *Psychonomic Bulletin & Review, 22*(5), 1193-1215.

[23] Kotsiantis, S., Pierrakeas, C., & Pintelas, P. (2004). Predicting Students' Performance In Distance Learning Using Machine Learning Techniques. *Applied Artificial Intelligence*, *18*(5), 411-426.

[24] Alghazo, J. M., Latif, G., Elhassan, A., Alzubaidi, L., Al-Hmouz, A., & Al-Hmouz, R. (2017). An Online Numeral Recognition System Using Improved Structural Features–A Unified Method for Handwritten Arabic Numerals. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *9*(2-10), 33-40.

[25] Zeng, Z. Q., Yu, H. B., Xu, H. R., Xie, Y. Q., & Gao, J. (2008, November). Fast training support vector machines using parallel sequential minimal optimization. In *2008 3rd international conference on intelligent system and knowledge engineering* (Vol. 1, pp. 997-1001). IEEE.

[26] Nguyen, D., Smith, N. A., & Rose, C. (2011, June). Author age prediction from text using linear regression. In *Proceedings of the 5th ACL-HLT workshop on language technology for cultural heritage, social sciences, and humanities* (pp. 115-123).

[27] Fu, J. C., Huang, H. Y., Jang, J. H., & Huang, P. H. (2019). River stage forecasting using multiple additive regression trees. *Water Resources Management*, *33*(13), 4491-4507.

[28] Latif, G., Iskandar, D. A., Alghazo, J. M., & Mohammad, N. (2018). Enhanced MR image classification using hybrid statistical and wavelets features. *IEEE Access*, *7*, 9634-9644.

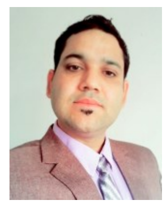
[29] Singh, B., Sihag, P., & Singh, K. (2017). Modelling of impact of water quality on infiltration rate of soil by random forest regression. *Modeling Earth Systems and Environment*, *3*(3), 999-1004

[30] Daniel, B. (2015). Big Data and analytics in higher education: Opportunities and challenges. *British journal of educational technology*, *46*(5), 904-920.

[31] Chen, R. (2012). Institutional characteristics and college student dropout risks: A multilevel event history analysis. *Research in Higher education*, *53*(5), 487-505.

[32] Barefoot, B. O. (2004). Higher education's revolving door: Confronting the problem of student drop-out in US colleges and universities. *Open Learning: The Journal of Open, Distance and e-Learning, 19*(1), 9-18.

[33] Cabus, S. J. (2017). Why do school dropout rates vary (so much) across countries? A survey. In G. Johnes, J. Johnes, T. Agasisti, & and L. López-Torres (Eds.), *Handbook of contemporary education economics* (pp. 43–75). Edward Elgar Publishing.



Ghazanfar Latif received a Ph.D. in computer science from the University of Malaysia Sarawak (MY). He serves as a Research Coordinator for the Deanship of Graduate Studies and Research at Prince Mohammad bin Fahd University, KSA. He is a computer scientist whose research interests include image processing, artificial intelligence, neural networks, and medical image processing.

Runna Alghazo obtained her Ph.D. in Rehabilitation from Southern Illinois University-Carbondale (USA). She is an Assistant Professor in the College of Sciences and Human Studies at Prince Mohammad Bin Fahd University. She previously served as an Assistant Professor in the Special Education Department at King Faisal University (KFU). She is an educational researcher and rehabilitation counselor. Her research interests are special education, accommodations for students with disabilities, assistive technology, inclusion, and cognitive psychology.

Maura A. E. Pilotti is a cognitive psychologist whose research interests include learning and memory processes across the lifespan. Her research focuses on the interrelations of memory, language, and emotion. She received her Ph.D. in cognitive psychology at the City University of New York (USA). She is currently an Assistant Professor in the College of Sciences and Human Studies at Prince Mohammad Bin Fad University.

Ghassen Ben Brahim is an Assistant Professor of computer science at Prince Mohammed University. He received a Ph.D. in computer science from Western Michigan University (USA). He worked as a Systems Analyst Engineer at the Integrated Defense Systems of Boeing and as a research visitor at the US Naval Research Lab. His research interests include machine learning, computer and network security, wireless networks, QoS routing in large-scale MANETs, routing in all-optical networks, and the design and analysis of network protocols.