

Employing artificial intelligence techniques for student performance evaluation and teaching strategy enrichment: An innovative approach



Lalbihari Barik*, Omar Barukab, Adel Ali Ahmed

Faculty of Computing and Information Technology in Rabigh, King Abdulaziz University, Rabigh 21911, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:

Received 9 February 2020

Received in revised form

12 May 2020

Accepted 27 June 2020

Keywords:

Intelligent tutoring systems

Student performance

Teaching strategies

AI techniques

ABSTRACT

An intelligent tutoring system is an excellent Artificial Intelligence (AI) alternative for the haunting problems of the teaching and evaluation system in university education. It evinces a paradigm shift in the current system by employing AI techniques to evaluate students' performance and enrich the myriad teaching strategies. Unlike in regular classes where a teacher has to control 30 to 50 students, a teacher has to monitor hundreds of students, which is quite difficult and mentally exhausting. In such circumstances, mentors or teachers alone are not enough for monitoring the students and offering each student's optimum attention and care. A new and original approach is needed to facilitate reliable and flexible methods of university student monitoring systems. The system should be able to evaluate the performance of many students, predict the final grade, and formulate intelligent decisions in real-time. Several computer-based models of AI are progressively performing an important role in teaching and performance evaluation of students. This paper proposes a new strategy to illustrate the advantages of applying AI techniques to predict the final grade of students. The validation process was carried out with the real-time 1000 students' dataset of 12 core and 18 elective courses in Bachelor of Computer Science during the academic year 2018-2019. In this paper, hybrid SVM with a Fuzzy Expert System is proposed to show the techniques proficiency for teaching and students' final grade prediction and the possibility of future work.

© 2020 The Authors. Published by IASE. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The recent decades have witnessed the development of various specialized tutoring systems within a particular knowledge domain. A major drawback of any knowledge-based Intelligent Tutoring Systems (ITS) is writing rules for generating teaching strategies that involve the experts to write different rules for the same problem given (Ata and Koçyigit, 2010). Although the adaptation technique changes dynamically for different kinds of students depend on the student model, a single teaching method would not work in a multiple domain teaching environments (Bae et al., 2010).

At times, making a proper evaluation of students' learning, primarily online, can be quite formidable. Recent times have witnessed the burgeoning of

online examinations, assisted by the distribution of soft copies of question banks. However, this system is fraught with certain obvious drawbacks—there is neither flexibility nor profusion of questions to ask. Further, this method suffers from the absence of a well laid out system to assign levels of difficulty for the questions (Banakar and Azeem, 2008).

The authors proposed an open architecture ITS for the university that makes efficient use of available resources. ITS helps to develop various intelligent agents that monitor and diagnose based on the user's responses and help to teach procedural knowledge as well as to facilitate the acquisition of conceptual knowledge in multiple subjects and provide a fair evaluation (Cheng et al., 2010). Here, the hybrid SVM with a Fuzzy Expert System is proposed to show the efficiency of the technique for teaching and students' final grade prediction and the possibility for future work.

In this work, two methods are being employed. In the first method, classification and regression algorithms are used to find the patterns for study-oriented data and the data of students' public activities (Dimitriou et al., 2008). The latter comes into the picture because a clear understanding of the

* Corresponding Author.

Email Address: lalbihari@gmail.com (L. Barik)

<https://doi.org/10.21833/ijaas.2020.11.002>

Corresponding author's ORCID profile:

<https://orcid.org/0000-0002-5977-6319>

2313-626X/© 2020 The Authors. Published by IASE.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

students' social characteristics is very much essential in framing their career and hence plays a significant role in forecasting the future course of action (Dong et al., 2010). The collaborative filtering technique is the second approach which is proposed in this work.

The final grades are decided based on the marks obtained and the results achieved by the same candidates. Then the outcomes of both the approaches get comparable average results are constructively utilized for grade prediction (Esfahanipour and Aghamiri, 2010). A scrutiny of the differences between both methods is very enlightening. The classification and regression algorithms approach provides significant results only when the number of students is minimal. However, the second method reaches significant results for the mathematical type courses with many numbers of students (Fan et al., 2009).

Besides, the recognized course group's grade dependability is identified. Finally, it enables us to correctly identify half of all failures and predict the teacher's skills only with the error of one degree using the grade scale. In this work, the validation process was completed with the real-time 1000 student's dataset of 12 core courses and 18 elective course subjects in Bachelor of Computer Science (B.Sc.) course during the academic year 2018-2019.

2. Literature review

The educational ITS main difficulty is to design the student model and predict the performance of the students. So, based on the design of the student model, we can design a reliable performance prediction. However, the design of a student model is beneficial in many criteria like, identifying weak students, recognizing the areas that need improvement (Romero et al., 2013). The design of a student model is essential in formulating the adaptive performance of IT Systems (Nižnan et al., 2015), which is also an essential component in students' feedback generator system. An ITS system is created and collect and organize the students' data, study achievements, and their performance. Then all these data are saved in the university database system. Finally, the students' final grades are predicted based on the saved database.

The students' final grades performance predictions are essential in the commencement of each semester. This helpful for the students; they can plan their workload for the next semester accordingly. Also, this information is beneficial for planning a course conscription system. However, this system is complicated; at the initial stage, the grade prediction is more difficult because there is no necessary information regarding the students or their interest in the course (Harackiewicz et al., 2002). From the above said, students' activity report, the undertakings of students throughout the semester could be populated based on their prediction report. The difficulties in student's ranking prediction within a course are newly

processed using data mining approaches (Bydžovská, 2016).

The authors frequently observe the interrelated student reports like age, sex, and their department of study (Koprinska et al., 2015). Since the interrelated student reports are available in the university information systems, it leverages remarkable ease of access. Likewise, categorizing supplementary characteristics such as their behavior, their parents' qualifications (Nghe et al., 2007) are great indicators for understanding the propensities of an individual student.

This is the majority distinctive manner of acquiring the data to accomplish the questionnaire. In this work, the authors have used the King Abdulaziz University (KAU) student's data; it is the 54th University Impact Ranking 2019, and 201-250th World University Rankings 2020. KAU was instituted in 1967 as a private university to spread higher education in the western area of Saudi Arabia. In 1973, KAU joined the Saudi public universities' system. Now the university consists of nine faculties and 82,152 students, and authors aim to calculate the grades as precisely as possible. This paper cannot suppose the knowledge obtained by questionnaires since they incline to possess an inferior response rate. Thus, the requisites intended for this experiment solitarily based on the databases of the Information System of KAU (KAU, 2018).

The course supervisor is a supporter of the course. The resemblance is assigned with 1 when the comparison is made between the same supervisor and 0 elsewhere. The feature mass is assigned to 0.4 units. While calculating the resemblance between courses with the process as mentioned above, an average link clustering model is utilized (Strecht et al., 2015). The objective of this study is, towards predicting the students' grades among their essential prominence recognition of students who cannot meet their course necessities. Consequently, course necessities dealing with the most important tasks are the calculation of students' achievement or failure and calculation of the students' passing grades (Murtagh and Contreras, 2012).

In this work, two approaches are proposed. The first is based on classification and regression algorithms, which are used to find the patterns for study-oriented data and the data of students' public accomplishments. Students' social characteristics are very much essential in extrapolating on their future needs as well as forecasting their career. The collaborative filtering technique is the second approach which is proposed in this work. The final grades are based on previous achievements of the same students (Dilek et al., 2015). As discussed earlier, the outcomes of both the approaches are favorably comparable, which can later be profitably exploited for making grade prediction. Classification and regression algorithms approaches are limited because it offers significant results when the number of students is minimal. However, a collaborative filtering technique offers significant results for

mathematical courses with many numbers of students (Manouselis et al., 2011).

The presented two different approaches are the soul of our work. The first approach is the ITS based professional data mining classification with regression scrutiny (Bydžovská and Popelínský, 2014). Newly we formed a group of learners to exploit the force of the current technique. Also, an innovative kind of dataset concerning students' societal performance originates as of university dataset to develop the prediction. The next method is based on the collaborative filtering techniques (Bydžovská, 2015) functional to the educational environment. Then the results are plotted between users rating difficulty to course grade difficulty. In addition to that, the predicted final grades are mapped or arranged purely base on the previous achievement of comparable student's results.

From the above-said works of literature, it is evident that the classification and regression are the primary techniques which are used for student performance prediction (De Nooy et al., 2018). In the proposed work, the improvement has been achieved during the accumulation of the Social Behavior (SB) dataset to the unique dataset. With the Pajek method, they added the computed extra standard graph features like friends, their strengths and weakness, their network (Manouselis et al., 2011). In this work, persistent methods utilize in the recommender system (Matuszyk and Spiliopoulou, 2014). Here, the baseline user-based method is used for prediction.

Consequently, the proposed method is designed accordingly with the average link clustering to collect the investigate course with their comparison determination. This research describes in cooperation of the two approaches in detail. Here, the combined learning algorithms of SVM with Fuzzy Expert System is proposed to improve the efficiency of the technique for teaching and students' final grade prediction and the possibility for future work. In this work, the validation process was done with the real-time 1000 student's dataset of 12 core courses and 18 elective courses in the B.Sc. course during the academic year 2018-2019. Then the comparison between the two approaches and finally the reports of their advantages and disadvantages are made in the experiment setup.

3. Methods and materials

In this work, the hybrid SVM with a Fuzzy Expert System is proposed to be used to showcase the efficiency of the technique for teaching skills and students' improvement and the possibility for future work. In this work, the validation process was done with the real-time 1000 student's university dataset of 12 core courses and 18 elective courses in the B.Sc. course during the academic year 2018-2019. The results are tested with the forty-five teaching teacher's subjects results of midterm exams 1, 2, and final exams.

3.1. Dataset

Practically chronological datasets are designed for experiments to allow the authors to assess for the proposed method. In this paper, the authors process the dataset of 38 core subjects and 18 elective subjects in the B.Sc. course offered to students. We access the dataset, which has saved in the university's memory units at the moment of students' enrollments. Fresher students' details are not taken in this account for the reason that there is no data about them in that dataset. Here, the hybrid SVM with a Fuzzy Expert System is proposed to be used to showcase the efficiency of the technique for teaching skills and the possibility for future work. In this work, the validation process was done with the real-time 1000 student's dataset of 12 core courses and 18 elective courses in the B.Sc. course during the academic year 2018-2019. The dataset consists of nearly 1000 student details. The training and testing datasets are validated. The training set is validated with the dataset of 300 students from the university's memory units.

Moreover, these are used for the classification of more appropriate methods through their set. Then the testing set consists of a dataset of 450 students from the university's memory unit's student details. Also, which were used for the evaluation of both methods with the different dataset (Murtagh and Contreras, 2012). Here, the fivefold cross-validation scheme was evaluated.

From the students' dataset, we developed intelligent agents, where some set of questions asked to the students and monitor based on their responses. Also, develop a secure ITS to teach procedural knowledge as well as to facilitate the acquisition of conceptual knowledge in multiple subjects. Then provide the best possible quality services to the student as well as to teacher for learning and improvement in various expertise areas. Finally, the system shows enough intelligence to change the way a student asks questions based on the results. This helps the teacher to draw attention in student learning capabilities, their past, and present performances in particular courses, subjects, or topics.

We validated the performance of the approaches through mean absolute error, i.e., T-Statistics. These method dealings with the prediction of nearby approximation predictions are getting and compare it with the existent outcome. Lesser values are a symbol of superior outcomes. These determinations are frequently used for ranking calculation assessment. Within the educational environment, the main issues are weak student's identification. Hence, we also compute accuracy, or it is also called a recall. Student's categorization merely based on only as best or worst. Here, the accuracy procedures are deal with the percentage of worst students who are correctly secret as unsuccessful. Finally, the students' achievements are based on the teacher's teaching and coaching skills, so they calculated

through the utilization of the F1 score to facilitate the stability among precision and recall.

3.2. Study outcomes

This proposal is beneficial for the University in a heterogeneous environment to develop a new teaching model where the student is at the center of the learning process to knowledge acquisition and competences. Students can plan and schedule for their examination based on their knowledge, preparation, and appear in the regular exams set by the teacher members of the University.

3.3. Study-related data

From the above-said works of literature, it is evident that the classification and regression are the primary techniques which are used for student performance prediction. In the student performance prediction, the authors frequently examine the Study Related (SR) data first. In this paper, the study-related information enclosed the general attribute like gender, the date of birth, admission year, passed course credit scores, and the average grades. A classifier is introduced to investigating the course base training set, and then the performances of training sets are evaluated with a two-fold cross-validation method. This technique achieves the most excellent outcome be consequently validating by the test set.

3.4. Existing grade prediction

In the student grade prediction model, the regression is a frequently used method. Commonly used regression algorithms are SVM with the decision tree classifier regression algorithms, random forest algorithms, IBK algorithms, Rep Tree algorithms, and linear/additive regression algorithms. The baseline representations predict the standard grade of the training set of the particular selected course. From the results, the best result was produced by the SVM regression algorithm.

A decision tree classifier method is the best one for the grade with the case-based prediction. Tree classifier techniques are well trained one through without assumption resting on the distribution of information and the structure of the true model. A decision tree is a simple method which uses for regression and classification problem as well as supply a useful and simple tool for data interpretation. Fuzzy logic takes account of the true and false statements, all the algorithms can compare these true and false cases, but it works like the human brain. Fuzzy logic gives the impression more rapidly to the human brain works.

4. Results and discussion

The most important aim of this work is to predict the student's final grades at the beginning of each

semester. It is essential for every semester through the eminence scheduled to identify ineffective students. In this work, there are two unique approaches presented for finding the ITS. Initially, the classifications among the student's grade prediction values are through by the classification algorithm. And regression algorithms are used for finding the students' grade results in case-based classification. Finally, the SVM with the decision tree classifier attains the most excellent outcome of the students' ITS skill calculation results. This method is usefully utilized in grade prediction of small student dataset. The next novel approach utilized is the collaborative filtering technique and the predict grades base on the resemblance of students' achievement. The benefit of these approaches can save every university data of students' grades, not like the students' social performance. Similarly, these approaches succeed in course dependency identification.

4.1. Existing student pass or failure prediction algorithms

In this work, the primary assignment was to expose failure students. Here, there are two prediction classes were considered. They are class 3 as student's success defined grades and the failure defined in class 4. In this method, the class-based classification technique is performed with four grades. The grade 1 denotes as 'excellent'; grade 1.5 denotes as 'very good'; grade 2 denotes as 'good.' Then the grade 2.5 denotes as 'satisfactory'; grade 3 denotes as 'sufficient'; finally, grade 4 denotes as 'failed.' Commonly used algorithms are used for finding the comparative study. Naive Bayes, J48, Random Forests, Part, IB1, OneR, Support Vector Machines (SVM) algorithms are used, and the determined values are shown in Fig. 1. Here, there is a baseline developed for each model, which constantly predicts the failure details. Fig. 1 shows the intention of our work, i.e., SVM with the decision tree classifier model achieve the greatest performance.

Fig. 1, blue color indicates the existing methods F1 value; Orange color indicates the T-Statistics, and the ash color indicates the Accuracy values. From the existing models, the graph shows the SVM is the best one among the teaching-learning AI techniques. So, in this work, a novel SVM with a Fuzzy Expert System is proposed to improve the efficiency of the technique for teaching and students' final grade prediction.

4.2. Support vector machine decision tree regression algorithm

For each assignment, the most elegant technique was preferred, and an ensemble learning model was built. First, the SVM or SVM with the decision tree regression classifiers were predicted the failure cases, i.e., grade 4. Subsequently, the ensemble learners as well predict the failure cases.

Alternatively, else, the result values of the grade have been obtained by the SVM with the decision

tree regression classifier.

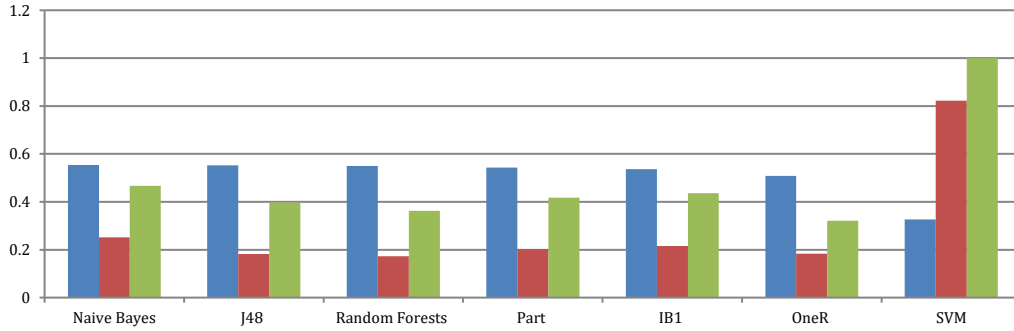


Fig. 1: Comparative existing classification algorithms results

At last, the final results of the SVM with the decision tree classifier advance shown in Fig. 2. Then the test set values are evaluated by the use of classifiers. The outcome indicates that approximately half of the students from the sample

are failed students. Still, the job was tricky because the fact is all failed students amount is less than a part of whole students. The calculation mistake was regarding 0.75 quantities on common, which was nearly 1.5 quantities in this grading scale.

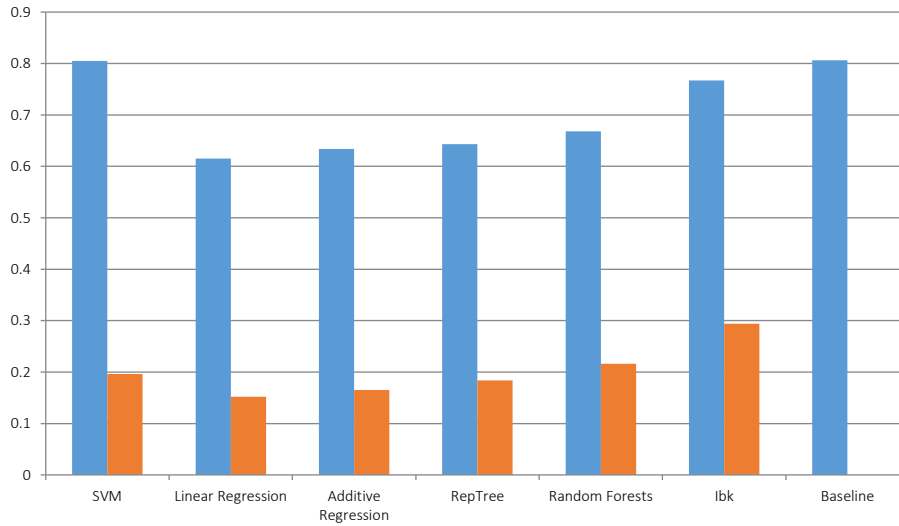


Fig. 2: Existing regression algorithms results

In Fig. 2, the blue color indicates the existing methods F1 value; orange color indicates the T-Statistics, and the ash color indicates the accuracy

values. Fig. 3 shows the SVM decision tree regression algorithm testing and training models.

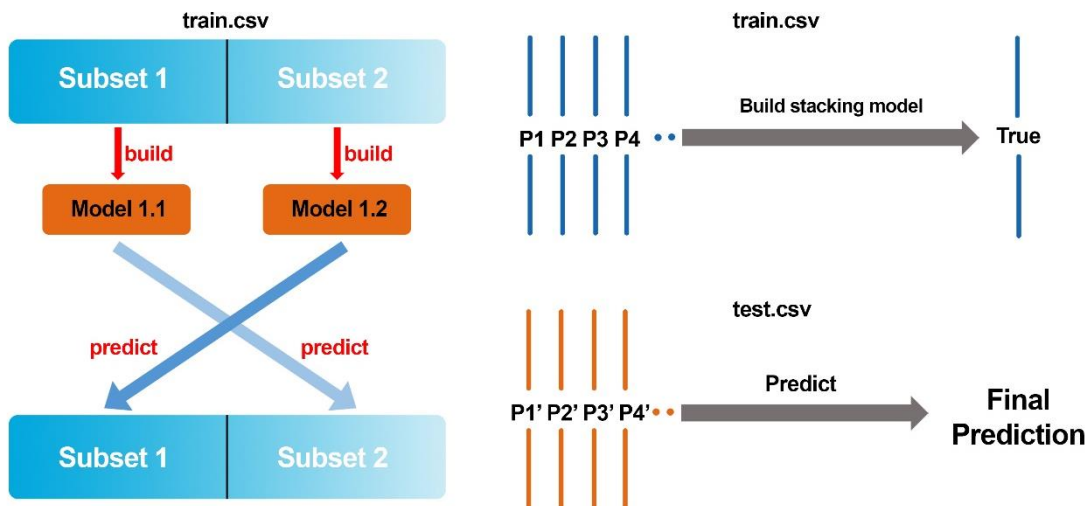


Fig. 3: Support vector machine decision tree regression algorithm testing and training

4.3. Social behavior data

Current research works are regularly based on finding supplementary data to improve prediction accuracy. In the proposed work, the improvement has been achieved during the accumulation of the SB dataset to the unique dataset. The exact kind of data originates from the university dataset describes the students' performance and their common support. We mainly pay attention to a statistical dataset that represents and dealings with students. The statistical dataset consists of posts and comments in the conversation forum, e-mails statistics, publications coauthoring, and file sharing. This information is served on the basis of the social relations between students and develops a sociogram.

Upon starting this sociogram, the latest features of weighted average grades of friends are obtained. This method is then calculated from additional standard graph characteristics such as the number of friends, the degree weighted by the strong relationships, and the measurements for each student in the network. Furthermore, the collected data of the students discover from various classification sections. Unfortunately, the university dataset never provides whole records of students

because of students' seclusion. Students are enthusiastically revealing them to develop in front of their classmates.

Also, using this method can determine the individual student's class attendance for a particular teacher's class. Amongst the other students, they can mention their desired courses from the offered courses. 'H' stands for Hypothesis is supposes with the purpose of students' communal interrelated into the student's performance. Further, the ensemble learner's train the dataset which contain social attribute was build. The evaluation results between other methods are shown in Table 1. From Table 1, we can evidence that the T-Statistics scores are somewhat lesser than average value. Though for the thirty-two different courses from the test set, the dissimilarity in T-Statistics was considerably enhanced through social action data.

The minimum value is 0.1, the average value is 0.178, and the maximum obtained value is 0.734. Merely five courses were achieved more unsatisfactory results like the minimum are 0.1, the average is 0.12, and the maximum is 0.21. Then for the remaining courses were get only very low differences, which are slight negligible value.

Table 1: Behavior attributes dataset

Dataset	Attributes	T-Statistics	Accuracy
Training Dataset	Study Related (SR)	0.734	0.596
	Study Related (SR) + Social Behavior (SB)	0.698	0.598
Test Dataset	Study Related (SR)	0.767	0.456
	Study Related (SR) + Social Behavior (SB)	0.709	0.478

Then the list of particular common attributes was developed in Table 1. We portray the pinnacle five communal performance attribute, which drastically affects the result. Fig. 4 shows the Students Performance Attributes with their course basis.

Hypothesis 1 data are confirmed. Dataset concerning students' performance enhanced the prediction. Base on the mainly important attribute, we believed that the support of students' contacts had enlarged the possibility of getting pass the courses.

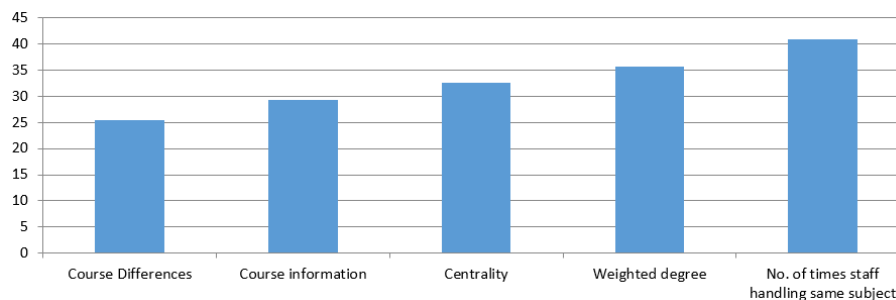


Fig. 4: Students performance attribute

4.4. Student grades

In this work, persistent a method utilized in the recommender system. Our focus is to find the similarities among students' grades from the dataset. In hypothesis 2 of students' acquaintance can be categorized by the results of course that students enroll through their studies. Based on this information, we might choose students with comparable interests and knowledge and consequently forecast whether an exacting student

has satisfactory skills necessary for an exacting course.

In this model, the preliminary work is classification regression, but it has some disadvantages it can be conquered in this work. Initially, a resemblance matrix G has been build whereas the matrix rows represent students' names or their details, and the columns represent the equivalent courses. The predicted grades of 64 courses resemblance matrix G have 500 columns in mat lab while analyzing all the students' grades. The Grades which have been obtained with all students

are made in a matrix in the training set. Suppose a student who is absent in an exacting course; the equivalent cell remains unfilled. The plan is to fill cells with the significant students' grades of the investigate courses enroll with students in the 2018 batch. In this matrix, authors have been mentioned by a representation null. From the grade, a vector of resemblance matrix, the comparison among all students enrolls from 2018 to 2019 are made. Fig. 5 shows the example of Matrix G.

		Courses				
		1	2	3	m
Students	1	4		3.67	4	
	2	3	3.33			
	3			2		
	⋮	3	2.67		3.67	
	n		3.33			4

Fig. 5: Example of matrix G

Here, extensively utilize resemblance metrics for the computation of students compared correctly for every single metric between the grades of students with the shared courses. Here, the common amount of courses collectively joined in students is 25 numbers. Consequently, the suitable region of general related to students to the examined student could be selected to influence the predicted final grade. Here, the baseline user-based method is used for prediction. The particular students are selected for the nearby student details that were extra comparable to the investigated student than the baseline candidate. In this work, the calculated types of baseline students are the average grade for each course and an average grade through all the courses of 3.5.

The nearby student of the major 35 students shows a realistic result. Although in the context of the same courses, there are 27 average students are enrolled in a single year. So, there are three-course categories occupancies used; 25 as small students, medium 50 students, and large >50 students. Consequently, the appropriate sizes of the nearby student for courses through the variable possessions are made. The correlation among the T-Statistics and the cardinality of N is shown in Fig. 6. The following are the selected nearby student of 15 for small courses, 25 for medium courses, and 50 for large courses. From the graph, the prediction for smaller course value is challenging assignments.

The concluding grades are estimated from the grades of comparable students belong to the compute neighbor one. For finding the optimum values, the major criterions such as mean, max, median as well as superior models like implication weighting are utilized. In Table 2, propagate the top five combinations of the comparison models are shown. The models for the neighbor value collection and the estimated grade functions are determined and plotted.

A baseline based model is utilized here for the consumer required huge neighbor values for each student on an average value. In the production system, reduce the tie among students are essential. Because at the time of recompilation in similarity at the course conscription process to be currently updated for students. Consequently, dissimilar neighbor values are chosen when the T-Statistics attains a high value. In the efficiency cases, only particularly the third row only for the performance of the system.

4.5. T-statistics for independent means and hypothesis testing with regressions

The first Hypothesis, 'H0', is a Null Hypothesis. It is initiated with the aim of class size 'c' as well as the class average (β) of the student, which are pass in similar courses based on the self-determining factor. The strength of students within a particular class does not affect every personality in that classroom. In this dataset, the student strength in each class of having either 50 or more than 50 students, i.e., 'H0': $\beta_1 = \beta_2$. Unconventional Hypothesis 'H1' proposes with a negative link exists among the student's strength and their performance. Exclusively the high strength class must result low than the low strength class result. These average results are calculated between the students' results as well as the small strength class leads to high average results, i.e., 'H1': $\beta_1 > \beta_2$.

In the second Hypothesis, the 'H2' hypothesis states that the teaching efficiency of teachers does not influence the class average. If the number of students in opposition to the teachers' class must be no evident teacher-student relationship, so the slope would be 0, i.e., 'H2': $\beta_1 = 0$.

Optional Hypothesis 'H3', is reflecting the relationship between the teaching efficiency of teachers' average value. There would likely exit a negative relationship, as bigger class sizes lead to lower lecturer concentrations. Hence, the graph lines of most excellent have a positive slope, i.e., 'H3': $\beta_1 = 1$.

Table 2: Various hypothesis results

Hypothesis	High student strength	Low Student strength	Relationship between teacher and student	Teachers teaching efficiency	Number of classes taken
H0: $\beta_1 = \beta_2$	73%	81%	92%	70%	67%
H1: $\beta_1 > \beta_2$	63%	68%	88%	66%	61%
H2: $\beta_1 = 0$	Not true 55%	True 40%	True 30%	Not true 60%	Not true 50%
H3: $\beta_1 = 1$	77%	80%	75%	70%	60%
Data 1000 students	500 students	500	500 and 45 Teachers	45 Teachers	45 Teachers

From Table 2 hypothesis results, the student's performance is based on the teacher's class handling basis and their subject knowledge-based efficiency. So, to improve the teaching and learning efficiency, there are four matters must be taken into account. The first one is student-teacher relationships based on their academic basis. Moreover, the second one is the teachers' particular subjects skill efficiency, and the third one is the number of classes taken by the same subject. Finally, the teacher handling the same subject repeated semester means they have comparatively much knowledge than others.

4.6. Student learning outcome

In this work, the performances of the students are calculated based on three attributes. They are problem analyzing, the local and global impact of computing, and process understanding. Based on these strategies, only the student learning outcome improves. In our proposed work, around 750 male and female student's dataset of every individual semester are taken. It is beneficial for analyzing their previous semesters for comparisons. Also, it is more helpful for analyzing the teacher's performance and the criteria to improve their performance in the teaching method. In Table 3, 'B' implies "an ability to analyze a problem, and identify and define the computing requirements appropriate to its solution." The 'G' implies "an ability to analyze the local and global impact of computing on individuals, organizations, and society." The 'J' implies "an understanding of processes that support the delivery and management of information systems within a specific application environment."

From the achievement arrangement Table 3, it is proved that the total 100 marks are distributed by 'B,' 'G,' and 'J' accordingly. The university assigns the maximum marks in each component for male and female students by their norms. The students' annual results are predicted based on the two assignments, quizzes, and exams. Here, the exams are conducted by three stages, i.e., before the final exams, there are two early-stage exams conducted; after that, the final exams are conducted. The allotted mark of 'B' value is around 37, 'G' is 32, and 'J' is around 31 marks. Table 3 shows the students gathered marks for a semester.

Table 3: Actual achievement in each component

Outcome	B			G		J		
	B1	B2	B3	G1	G2	J1	J2	J3
Assignmen1	2			2		1		
Assignmen2		1.7		1.9		1		
Quiz 1		2		1.6				0.9
Quiz 2		2		1.5		0.9		
Exam 1			6.1	3.2	1.2	2.1	0.9	
Exam 2	1.2		4.7	2.6	1.1	2.6	1.5	
Final Exam	4.4	3.7	4.9	3.1	2.5	3.7	2.1	3
Sub Total	7.6	9.4	15.7	15.9	4.8	11.3	4.5	3.9
TOTAL		32.7			20.7			19.7

The achievement arrangement Table 3 shows the actual value of the results out of 100 marks; the

resultant outcomes of the student's achievements are 73.1%. It is proved that the total 100 marks are distributed by 'B,' 'G,' and 'J' accordingly. The university assigns the maximum marks in each component for male and female students by their norms. The students' annual results are predicted based on the two assignments, quizzes, and exams. Here, the exams are conducted by three stages, i.e., before the final exams, there are two early-stage exams conducted; after that, the final exams are conducted. The obtained mark of 'B' value is around 32.7, 'G' is 20.7, and 'J' is around 19.7 marks. Table 4 shows the achievement percentage for each outcome.

Table 4: Achievement percentage in each outcome

Outcome	B	G	J
Achievement Percentage (Male Section)	88.4%	64.7 %	63.5%
Achievement Percentage (Female Section)	91.2%	71.4%	68.6%
Average	89.8%	68.05%	66.05%

The resultant outcome percentages of the overall achievement of the students are plotted in Table 4. The university assigns the maximum marks in each component for male and female students by their norms. The students' annual results are predicted based on the two assignments, quizzes, and exams. Here, the exams are conducted by three stages, i.e., before the final exams, there are two early-stage exams conducted; after that, the final exams are conducted. The obtained mark of B value is 89.8%, G is 68.05%, and J is around 66.05% marks. Table 4 shows the male and female students gathered percentage marks for a semester.

4.7. Student success and failure prediction

Commonly most of the students pass in their examination courses in our study, so that next is to finds a small neighbor value to facilitate the additional ineffective students. Fig. 6 shows the highest T-Statistics value reached. It was achieved when the added number of nearby comparable students. Though, this method is affected by a criterion as is very lower precision. Thus, we can expect a letdown in this method is that the prediction grade must be lower than the grade average. Then the precisions were enhanced and also can find an adequate quantity of failed students. Finally, the results of the proposed methods are T-Statistics is 0.754, and the obtained Accuracy is 0.378, which is shown in Fig. 6.

Any transform in the resemblance matrix 'G' might guide to the recompilation in the comparison of students was calculated from all students' grades. H3 is the third hypothesis value. The third hypothesis value has been calculated with the related courses necessitate the same skill having students to pass. From this hypothesis, the computational cost is reduced and does not extensively reduce the prediction accuracy. Instead

of predicting all attended course results, we are using only grades of similar courses.

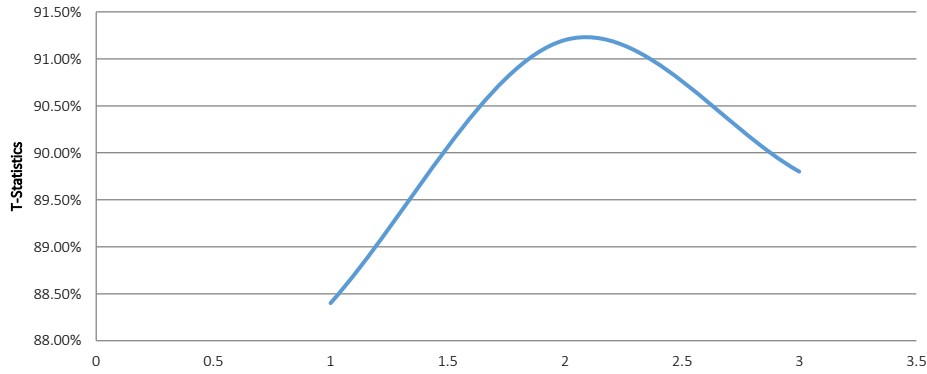


Fig. 6: Relationship between T-Statistics and the size of the neighborhood

4.8. Course characteristics

Students search for valuable data regarding the courses in the course list, facilitate them to choose whether they should register their courses or not. Also, find some certain dissimilar course descriptions and attempt to recognize dependency between courses. The relationship between the courses was defined with the weighted sum of the similarity of the certain course descriptions.

$$\text{Weighted Mean} = \frac{\sum_{i=1}^n xi * wi}{\sum_{i=1}^n wi} \dots \quad (1)$$

In the above equation, 'w' is defined as the weighting factor, and 'x' is the feature weight. The accumulations of the features are located respectively to exploit the grade calculation precision. The comparisons between each adjacent course are calculated. The preferred features and expanse metrics named dist are explained below.

Requisite is defined as a group, of course, which would be passing before students might register in a particular course. The resemblance value is '1' when the compared courses belong to requisite or else '0'. The mass of this feature is '1' since the requisite denotes a considerable dependency value.

$$\text{Jaccard Coefficient} = \frac{\sum_{k=1}^n (a[k] * b[k])}{\sum_{k=1}^n (a[k] + \sum_{k=1}^n b[k] - \sum_{k=1}^n (a[k] * b[k]))} \quad (2)$$

Works of literature contain the suggested teaching designed for a particular course, which is featured with a group of selected teachers. The resemblance between the group of teachers 'a' and the group of teachers 'b' is specified through the Jaccard coefficient. The featured mass of a group of value is 0.9, which was acquired because of the hypothesis, and the teachers never regularly publish in various fields. As a result, the works of literature might comprise a strong tie amongst courses. Then course content is represented by the textbook on the subject of study and summary about what be supposed to learn for that particular course. In the tutoring system, the checking slice STOP terms in the textbook furthermore utilize branch to obtain the extraction of words. It is beneficial when checking

the book content while tutoring. Consequently, the cosine similarity determination is used for dealing with ending vector symbols of the words' significance. Here, feature value is indicated by 0.7.

$$\text{Biased Jaccard Coefficient } J(A, B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} = \frac{|A \cap B|}{|A \cup B|} \quad (3)$$

The Biased Jaccard coefficient is utilized in finding the difference between the teachers of 2 courses. The mass value of the literature is made by 1, and 0.5 is assigned for the seminar tutors. The course supervisor is a supporter of the course. The resemblance is assigned to 1 when the comparison is made between the same supervisor and 0 elsewhere. The feature mass is assigned to 0.4 units. While calculating the resemblance between courses with the process as mentioned above, an average link clustering model is utilized. From the total 64 courses, 41 only selected for course selection, and there are 30 cluster values are formed. From the cluster formation, there are 18 investigate courses were presented. The present count of courses in every cluster ranges from 2 to 14 with common courses 12, and average shared courses are taken 3.

While calculating the resemblance between courses with the process as mentioned above, an average link clustering model is utilized. From the total 500 courses, 64 only selected for course selection, and there are 41 cluster values are formed. From the cluster formation, there are 30 investigate courses were presented. The present count of courses in every cluster ranges from 2 to 18 with common courses 14, and average shared courses are taken 3.

4.9. Relationship with other methods

Evaluation amid the various models with a group of grades, the proposed model attains optimistic things has happened for the iterations. Out of the 64 selected courses, there are 41 courses belong to the predicted cluster. The final grade might predict base the grades of just 3 further courses standard. Through cluster formation methods, possibly attain 18 of the investigate course belongs to various

clusters. From Table 5, a little improved T-Statistics is attained through the technique which utilizes the course features for the particular courses. Consequently, while in grade prediction, the equivalent courses are initially searching within SC2

after that in SC1. From the selected courses, the T-statistics, mean, and standard deviations are calculated. Table 5 shows a comparison between the hypotheses of the proposed model.

Table 5: Comparison between hypotheses of the proposed model

Method	T-Statistics	Accuracy	Mean	Standard Deviation (SD)
All grade values	0.740	0.778	0.630	0.598
For equivalent course grade prediction	0.721	0.762	0.641	0.582
Teachers Performance	0.690	0.710	0.650	0.530
Grade values from same classes	0.721	0.732	0.711	0.582

From Hypothesis 2 and Hypothesis 3, we can say that we achieved a prescribed model. Because of the use of students' grades and course characteristics, our proposed models predict students' performance. The predicted outputs are shown in Table 6. Also, the novel methods performances are tabulated in Table 6. The primary significance of the proposed model: Every data's of the university students' details are stored. Also, the data regarding student grades are required for prediction except for the students' social behavior calculation data. Here, we have recognized the course dependency to reduce the computation charge. Besides, the system is capable of predicting the ending grade considering the grades as of just five extra courses from the major investigate courses.

Table 6: The proposed models' Results

Dataset	T-Statistics	Accuracy
Training Dataset (SVM)	0.590	78%
Training Dataset (FUZZY)	0.610	80%
Training Dataset (SVM and FUZZY)	0.680	89%
Test Dataset (SVM)	0.625	76%
Test Dataset (FUZZY)	0.595	81%
Test Dataset (SVM and FUZZY)	0.698	88%

4.10. Teachers performance evaluation

Artificial intelligence is the most excellent technique to calculate the concert of teachers based on student's feedback and calculate grades. It can attain the teacher's presentation evaluation step by step base on the student's performance. AI allocates the performance assessment to get a position in immediate on a reliable source. It is individually supportive for them to get better improvement in

their carrier. Otherwise, if the teacher is distracted in their work immediately, bring up to them and help them to obtain a favored point. As a result, a variety of intelligent agents are residential to work together with the system by making several questions, and start supervises and identify based on the user's responses. The following hypotheses were checked with the teacher for their learning and improvement to provide the best possible teaching skills.

- Hypothesis A:

Ha: There is a considerable connection among teacher and student in academics

Hb: There is no considerable connection among teacher and student in academics

- Hypothesis B:

Ha: There is a considerable connection among teachers self-confidence

Hb: There is no considerable connection among teachers self-confidence

- Hypothesis C:

Ha: There is a considerable connection among teachers' past academic performance.

Hb: There is no considerable connection among teachers' past academic performance.

Hypothesis D:

Ha: There is a considerable connection among teachers' past similar subject performance.

Hb: There is no considerable connection among teachers past similar subjects.

The following Table 7 shows the different hypothesis values of our proposed method.

Table 7: Proposed hypothesis values

Hypothesis	Students Result	Study Eagerness	Academic Performance	Increased Results	Regression Analysis	SVM and FUZZY
Hypothesis A	63%	71%	82%	73%	0.691	0.682
Hypothesis B	74%	59%	76%	47%	0.702	0.632
Hypothesis C	True 45%	Not True 60%	Not True 80%	True 45%	0.714	0.672
Hypothesis D	67%	77%	64%	68%	0.743	0.682

4.11. Artificial intelligence learning and development

Performance reviews also provide a chance for students to establish the needed skills to be learning also new skills to obtain improved at their job. The collaborative filtering technique is moved towards on comparison of two grades or any other value

comparison utilization. Moreover, the adjusted cosine similarity value is calculated by the use of the already distinct resemblance matrix G for every couple of courses. Consequently, the proposed method is accordingly designed with the average link clustering to collect the investigate course with their comparison determination. Then the resulting clusters define the groups of same course clusters. At

last, the predicted students' grades of some courses the computation between grades obtain commencing courses belong to a similar cluster of investigating the course. From 41 of every investigate courses belong to single 25 clusters. The amount of course in a single cluster varies from 2 to 12. The average amount of certain courses in a single cluster is 3. The nominal amount of students' mutually sharing courses is to three.

Table 8: Comparison of both approaches

Dataset	Models	T-Statistics	Accuracy	Regression Analysis
Training Set	Students' Attributes	0.642	0.643	0.631
	Students' Attributes and Teachers Attributes	0.680	0.739	0.712
Test Set	Students' Attributes	0.628	0.769	0.704
	Students' Attributes and Teachers Attributes	0.698	0.735	0.643

In Hypothesis D, each row is more equivalent in various course groups. Here, selects the category based on basic course characters are described below.

The first one is difficult. Here, the average grade is considered as 2.4 from the student's grade. So the courses are divided into two subsystems. The grade values at 2.4 are considered as easy cases, and greater than 2.4 grades are in difficult cases. Finally, the occupancy rate values are defined by small, medium, and large. The course interests are separated into four parts as mathematics and basic sciences (M), fundamental computing (F), advanced computing (A), and others (O). The student's data of the following is plotted in a tree structure. The university requirements of credit are 26 points, and the applied science credit requirements are 15 points. There are 48 points for faculty requirements, 33 points for core courses, and 18 points for electives courses. There are 140 total credit hours in the university requirements for every subject. From the dataset, core course 12 and elective course 18 with student strength can take a maximum of 1000 students.

Every investigate course belong to any set of the group for every distinct category. Then the abovementioned categories related by the following three layers of tree structures. It is different in the category order distributing. In this experiment, the valuations are done between every transformation of category. The full tree is made up of the courses in the training set, which are distributed afterward with every category. Every node of the tree saved the details regarding the course from splitting. In every node, Harmonic Mean (HM) is evaluated then in comparing the two methods are generated an appropriate connection among Accuracy and T-Statistics. Afterward, the trees and combined branches cannot produce more attraction in a considerable phenomenon. Attractive branches contain any of the subsequent situations.

Student's academic performance reviews are essential for the individual student based determination about their knowledge and skills. Also, it needs to be refined their learning ability in new skills to get a better job. Advanced analytics researches are very helpful through personalized

The predicted students' attributes, teacher attributes, and maximum received similar average results are portrayed in Table 8. However, the values are being different in some situations. This proposal aims to categorize the course groups, which are providing constant predictions and identifying the comparatively worst model. The proposed novel SVM decision tree Regression classifier with Fuzzy Expert System comparison results is shown below.

and employee-based learning experience. Nowadays, skills are shorter projection in the student's life than ever before. Artificial Intelligence is the best and sophisticated technique to identify the employees or students who need to reinvent or renovate. It helps them to skills them as a great deal earlier than become superseded or get replaced with enhanced technologies.

The difference value is 0.3 in the HM node. The grade values of difficult and easy also present in the same node. Afterward, is there any equal dissimilarity rate within HM the nodes belong to sibling nodes. Finally, for the same difference in HM, the nodes are belonging to the parent and child nodes. The resultant tree structure is represented in Fig. 7.

From Fig. 7, our proposed methodology has the following benefits. The course groups that are predicting extensively superior compared with the average value are identified. The results contain only all the mathematical as well as English courses. The course groups do not predict considerably, not as good as regular values are recognized. The prediction consists nearly all courses belong in other categories as well as medium or large simple theory courses.

Hypothesis four is also established in this work. Here, the course group's values are predicted considerably enhanced through the easy model and which were represented in blue color. The results enclosed nearly every mathematics course. Alternatively, else the red node represents the enhanced outcome obtains via the difficult model. Also, it comprises nearly all the small courses too. In the yellow nodes, dissimilarity in prediction accurateness is a very insignificant value.

When all the outliers as well recognized the single courses of the sets how dissimilar performance with others like the English course is predict without any difficulty in associate through all courses belong to other categories. The small mathematical course only differs in this model, which achieves a superior outcome compared to other mathematical courses.

A new approach intended for prediction in student presentation while the test set performance in this work. Also, without any trouble in establishing some exacting course on a tree basis

and the proposed method attain a proper result. In addition to that, there is no prediction designed for course in reliable. Table 9 shows the final validated

results on the Test set. Here, the results are considerably enhanced with other models because of the utilization of SVM with decision tree classifiers.

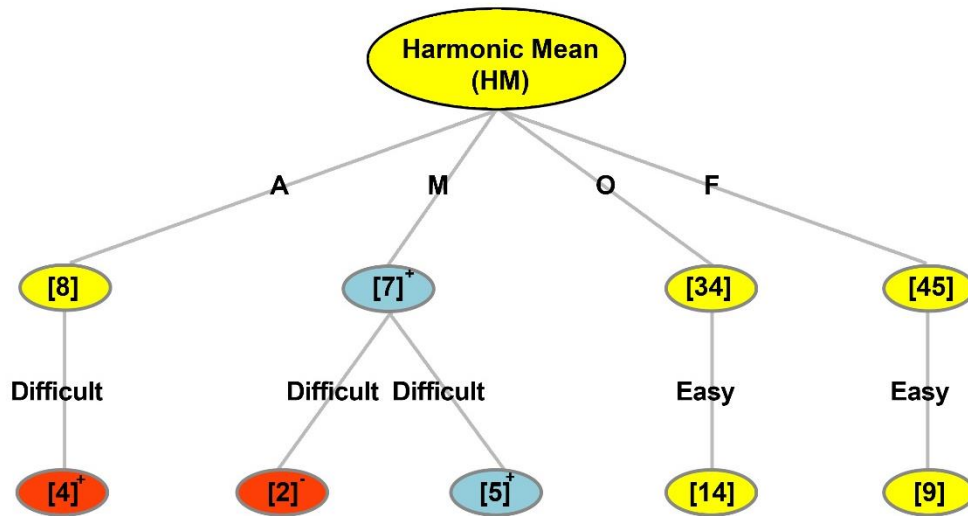


Fig. 7: Regression tree structure

Table 9: Predicted results of validated on the Test set

Methods	T-Statistics	Accuracy	Accepted Courses	Rejected Courses
SVM Regression model	0.639	0.686	22	12
SVM decision tree Regression classifier with Fuzzy expert System	0.758	0.789	25	0

In this model, a new approach is used for prediction in student presentation while the training set performance. This is used to find the students' subject difficulties and how the teacher can improve the same. Also, without any trouble in establishing some exacting course on a tree basis and the proposed method attain a proper result. In addition to that, there is no prediction designed for course in reliable. Table 10 shows the final validated results on the Training set. Here, the results are significantly improved with other models because of the utilization of SVM with decision tree classifiers.

Artificial intelligence is the best method for calculating the performance of teachers based on their student's feedback and their predicted grades. The data collection, make decisions based on the collected data, eliminate biasing are the basic

benefits of the Artificial intelligence method for calculating the performance of the teacher. These three steps are done based on the simple and clear cut process, but while processing, some difficult challenges are there. It is on behalf of the student performance and their final grade values. It is not only for a single teacher; it is team up with several teams or departments from moment to moment. Hence, it gets complicated to accumulate data from all the input data holders. Most of the data collection times, the accumulated input data holders miss certain sources of information. They fail to notice some precious aspect of the teacher's assistance; it leads some good teachers to get bad addresses. So the presentation review might be exactly most important to the teacher demotivated.

Table 10: Predicted results of validated on the training set

Approaches	T-Statistics	Accuracy	Accepted courses	Rejected courses
SVM Regression model	0.594	0.736	30	10
SVM decision tree Regression classifier with Fuzzy expert System	0.698	0.749	32	0

Artificial Intelligence is capable of obtaining the teacher's performance review step by step based on the student's performance and gets rid of the required a permanent interval completely. AI allows the performance assessment to obtain a position in instantaneous on a dependable source. For the time of the ambition location procedure, persons have permanent target or quota they necessitate to accomplish in a precise time limit. This personally helps them to improve the teachers' improvement, and they can improve them based on the students' performance analysis. AI is proficient help to administer the whole improvement concurrently

and provide immediate feedback to improve them. If the student's performance is in the right direction, the teacher is doing their works correctly, so they are eligible for appreciation in practical.

Table 11 shows the students' assessment marks; the students must score through the teacher's teaching method. So, to improve the teacher's performance is the reciprocal of the student's achievements.

Course Values are calculated with the sum of assessment value 1, assessment value 2, assessment value 3, and other values. The teachers are only

responsible for the student's final results and their achievements.

Table 11: Predicted results of validated on the training set

Students assessment vs. teachers perceptions	T-Statistics	Accuracy	Regression analysis	Perception
Language and Basic Sciences	0.594	0.626	0.643	0.631
Mathematics	0.698	0.649	0.539	0.612
Computer Science	0.624	0.716	0.569	0.604
Others	0.641	0.789	0.535	0.543

Alternatively, the teachers are unfocused in their work instantaneously inform them and help them to obtain a preferred stage. Various intelligent agents are developed to inquire about a number of questions and start monitoring and diagnosing based on the student's responses. A secure ITS developed to teach procedural knowledge as well as to facilitate the acquisition of conceptual knowledge with four parts, such as basic sciences (M), fundamental

computing (F), advanced computing (A), and other (O) subjects. The results are tested with the forty-five teaching teacher's subjects results of midterm exams 1, 2, and final exams. ITS helps to provide the best possible teaching skills to the student as well as to the teacher for learning and improvement in various expertise areas. Table 12 shows descriptive information.

Table 12: Descriptive information

Statistics	Numbers	Min/Max (%)	Mean	Standard Deviation (SD)
Students Performance	500	19/64	5.88	.689
Course Outcomes	400	34/81	6.23	.534
Assignment Outcomes	350	37/88	5.81	.721
Exam Outcomes	500	54/89	6.13	.632
Course Interest	300	53/76	8.01	.564
Teachers Performance	45	80/92	9.02	.842

The resultant outcome percentages of the overall achievement of the students are plotted in Table 12. Also, the indirect method of grade predictions is done, and the average values are taken for the teacher's performance evaluation. The university assigns the maximum marks in each component for male and female students by their norms. The students' annual results are predicted based on the

two assignments, quizzes, and exams. Here, the exams are conducted by three stages, i.e., before the final exams, there are two early-stage exams conducted; after that, the final exams are performed. The obtained mark of 'B' value is 89.8%, 'G' is 68.05%, and 'J' is around 66.05% marks. Table 13 shows the male and female students gathered percentage marks for a semester.

Table 13: Predicted results of teachers performance on the training set

Method	T-Statistics	Accuracy	Mean	Standard Deviation (SD)
Regression model	0.770	0.539	0.661	0.676
SVM Regression model	0.731	0.516	0.649	0.772
SVM decision tree Regression classifier with Fuzzy expert System	0.790	0.628	0.691	0.781

From Table 13, results outcome of the regression analysis, the teacher's care for students is very much necessary. Also, the assignments are submitted based on the teachers' request or punishment only. So the individual care for each student is a must. The teachers who are handling maths and physics must take extra care for every student because the students are feeling high difficult compared to the other subjects. In this proposed research, authors design and construct a heterogeneous environment, develop in an open architecture, identify several intelligent agents, and assign to their predefined role to track different aspects of students learning and performance patterns, behavior, and personal traits. All this information is stored in the knowledge base system. Our focus is on more precious types of question bank along with annotating the pedagogical material using metadata for facilitating its reusability to the teacher members and students.

5. Conclusion

In this research, the main difficulty is to predict the student's final grades on the commencement of

the semester with the prominence scheduled to identify ineffective students. Here, there are two unique approaches presented for the ITS. Initially, extensively classification and regression algorithms are used, and the SVM with decision tree classifier reaches the most excellent outcome. This method is usefully utilized in grade prediction of small student dataset. The next novel approach utilized is the collaborative filtering technique and the predict grades base on the resemblance of students' achievement. The benefit of these approaches can save every university data of students' grades, not like the students' social performance. Similarly, these approaches succeed in course dependency identification.

From these, we conclude that we can predict final grades of the investigate course by exploratory grades of merely three other courses. This model might work usefully in grade prediction of mathematical courses. Compared with the predicted courses in the other unusual areas, the present used dataset B.Sc. students are very good in their academics. Moreover, they pass in their large informatics courses. Lastly, the final grade predicted

value is achieved as an error degree in one course only and high-grade scale value for all other courses. The partial number of students' failures is too identifying correctly. Still, the project is tough because the information from failing grades comprises below a quarter of all grades.

Compared with the unusual areas, the B.Sc. students are very good in their academics based on their predicted courses. In addition to they get a pass in their large informatics courses. The rapid growth of ITS had a set of optimistic results and brought many services into our daily lives. This manuscript has accessible advance completed consequently in the field of apply AI techniques for university teacher and student details, and the presentation skills prediction and further improvements also specified the extent for future work.

Acknowledgment

This project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant No. G-446-830-34. The authors, therefore, acknowledge with DSR technical and financial support.

Compliance with ethical standards

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Ata R and Koçyigit Y (2010). An adaptive neuro-fuzzy inference system approach for prediction of tip speed ratio in wind turbines. *Expert Systems with Applications*, 37(7): 5454-5460.
<https://doi.org/10.1016/j.eswa.2010.02.068>
- Bae C, Yeh WC, Chung YY, and Liu SL (2010). Feature selection with intelligent dynamic swarm and rough set. *Expert Systems with Applications*, 37(10): 7026-7032.
<https://doi.org/10.1016/j.eswa.2010.03.016>
- Banakar A and Azeem MF (2008). Artificial wavelet neural network and its application in neuro-fuzzy models. *Applied Soft Computing*, 8(4): 1463-1485.
<https://doi.org/10.1016/j.asoc.2007.10.020>
- Bydžovská H (2015). Are collaborative filtering methods suitable for student performance prediction? In: Pereira F, Machado P, Costa E, and Cardoso A (Eds.), Portuguese conference on artificial intelligence: 425-430. Springer, Cham, Switzerland.
https://doi.org/10.1007/978-3-319-23485-4_42
- Bydžovská H (2016). A comparative analysis of techniques for predicting student performance. In the 9th International Educational Data Mining Society, Raleigh, USA: 306-311.
- Bydžovská H and Popelínský L (2014). The influence of social data on student success prediction. In the 18th International Database Engineering and Applications Symposium, Association for Computing Machinery, Porto, Portugal: 374-375.
<https://doi.org/10.1145/2628194.2628199>
- Cheng MY, Tsai HC, and Sudjono E (2010). Evolutionary fuzzy hybrid neural network for project cash flow control. *Engineering Applications of Artificial Intelligence*, 23(4): 604-

613.
<https://doi.org/10.1016/j.engappai.2009.10.003>
- De Nooy W, Mrvar A, and Batagelj V (2018). *Exploratory social network analysis with Pajek: Revised and expanded edition for updated software*. Vol. 46, Cambridge University Press, Cambridge, UK.
<https://doi.org/10.1017/9781108565691>
- Dilek S, Çakır H, and Aydın M (2015). Applications of artificial intelligence techniques to combating cyber crimes: A review. *International Journal of Artificial Intelligence and Applications*, 6(1): 21-39.
<https://doi.org/10.5121/ijaiia.2015.6102>
- Dimitriou L, Tsekeris T, and Stathopoulos A (2008). Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow. *Transportation Research Part C: Emerging Technologies*, 16(5): 554-573.
<https://doi.org/10.1016/j.trc.2007.11.003>
- Dong Y, Xiang B, Wang T, Liu H, and Qu L (2010). Rough set-based SAR analysis: An inductive method. *Expert Systems with Applications*, 37(7): 5032-5039.
<https://doi.org/10.1016/j.eswa.2009.12.008>
- Esfahanipour A and Aghamiri W (2010). Adapted neuro-fuzzy inference system on indirect approach TSK fuzzy rule base for stock market analysis. *Expert Systems with Applications*, 37(7): 4742-4748.
<https://doi.org/10.1016/j.eswa.2009.11.020>
- Fan YN, Tseng TLB, Chern CC, and Huang CC (2009). Rule induction based on an incremental rough set. *Expert Systems with Applications*, 36(9): 11439-11450.
<https://doi.org/10.1016/j.eswa.2009.03.056>
- Harackiewicz JM, Barron KE, Tauer JM, and Elliot AJ (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology*, 94(3): 562-575.
<https://doi.org/10.1037/0022-0663.94.3.562>
- KAU (2018). Our history. King Abdul-Aziz University. Jeddah, Saudi Arabia. Available online at:
<https://bit.ly/2C9qNXU>
- Koprinska I, Stretton J, and Yacef K (2015). Students at risk: Detection and remediation. In the 8th International Conference on Educational Data Mining, Madrid, Spain: 512-515.
- Manouselis N, Drachsler H, Vuorikari R, Hummel H, and Koper R (2011). Recommender systems in technology enhanced learning. In: Ricci F, Rokach L, Shapira B, and Kantor P (Eds.), *Recommender systems handbook*: 387-415. Springer, Boston, USA.
https://doi.org/10.1007/978-0-387-85820-3_12
- Matuszyk P and Spiliopoulou M (2014). Hoeffding-CF: Neighbourhood-based recommendations on reliably similar users. In the International Conference on User Modeling, Adaptation, and Personalization, Springer, Aalborg, Denmark: 146-157.
https://doi.org/10.1007/978-3-319-08786-3_13
- Murtagh F and Contreras P (2012). Algorithms for hierarchical clustering: An overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(1): 86-97.
<https://doi.org/10.1002/widm.53>
- Nghe NT, Janecek P, and Haddawy P (2007). A comparative analysis of techniques for predicting academic performance. In the 37th Annual Frontiers in Education Conference-Global Engineering: Knowledge without Borders, Opportunities without Passports, IEEE, Milwaukee, USA: T2G-7.
<https://doi.org/10.1109/FIE.2007.4417993>
- Nižnan J, Pelánek R, and Rihák J (2015). Student models for prior knowledge estimation. In the 8th International Educational Data Mining Society, Madrid, Spain: 109-116.

Romero C, López MI, Luna JM, and Ventura S (2013). Predicting students' final performance from participation in on-line discussion forums. *Computers and Education*, 68: 458-472.
<https://doi.org/10.1016/j.compedu.2013.06.009>

Strecht P, Cruz L, Soares C, and Mendes-Moreira J (2015). A comparative study of classification and regression algorithms for modelling students' academic performance. In the 8th International Conference on Educational Data Mining, Madrid, Spain: 392-395.