International Journal on Advanced Science Engineering Information Technology
Vol.9 (2019) No. 6
ISSN: 2088-5334

Predicting Students' Course Performance Based on Learners' Characteristics via Fuzzy Modelling Approach

Chee Siong The[#], Shu Hsien Lee[#], Mohamad Hardyman Barawi[#]

[#]*Faculty of Cognitive Sciences and Human Development, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia*
E-mail: csteh@unimas.my; shlee_pg@hotmail.com; bmhardyman@unimas.my

***Abstract*— Frequent assessment allows instructors to ensure students have met the course learning objectives. Due to lack of instructor-student interaction, most of the assessment feedbacks and early interventions are not carried out in the large class size. This study is to proposes a new way of assessing student course performance using a fuzzy modeling approach. The typical steps in designing a fuzzy expert system include specifying the problem, determining linguistic variables, defining fuzzy sets as well as obtaining and constructing fuzzy rules is deployed. An educational expert is interviewed to define the relationship between the factors and student course performance. These steps help to determine the range of fuzzy sets and fuzzy rules in fuzzy reasoning. After the fuzzy assessing system has been built, it is used to compute the course performances of the students. The subject expert is asked to validate and verify system performance. Findings show that the developed system provides a faster and more effective way for instructors to assess the course performances of students in large class sizes. However, in this study, the system is developed based on 150 historical student data and only a total of six factors related to course performance are considered. It is expected that considering more historical student data and adding more factors as the variables help to increase the accuracy of the system.**

***Keywords*— assessment feedback; fuzzy logic; fuzzy inference system; student modeling; intelligent tutoring system.**

I. INTRODUCTION

Examining learner characteristics is one of the crucial principles in instructional design [1]. Insights on learner characteristics when designing instruction and learning help to assure that most of the target audience is receptive to the instruction. It also often leads to effective, efficient, and appealing teaching and learning.

An instructional designer should determine the learner's skills, knowledge, and attitude toward learning before designing the instruction. Among the characteristics that should be taken into consideration include specific prior knowledge, prerequisite knowledge, academic motivation, attitudes and levels of education and/or ability [1]. Learner analyses are often done via self-reported instruments [2], [3]; surveys [4], [5]; as well as interview [6].

Some studies use statistical methods to predict student performance and to examine factors that are associated with student performance [7], [8]. Other methodologies are later employed to perform such prediction. Recent studies such as [9] use eye-gaze activities to examine how individuals perform different gazing activities when holding different team roles in pair programming and derive predictive insight into their post-test performance. Data mining techniques that employ classification algorithms like Decision Tree, Naïve Bayes and Support Vector Machine were used to discover the significant predictors of student's performance [10]. Such prediction enables teachers and parents to keep track of student performance and provide appropriate intervention. Some studies apply genetic programming to predict online student performance and investigates the efficacy of Singular Value Decomposition (SVD), a matrix factorization technique, for predicting student performance [11], [12].

To provide a better analysis of students' learning characteristics, student modelling is to be built. Student modelling is defined as the process of creating representation of a student's characteristics [13]. As the characteristics involve many factors and facts that are vague and unmeasurable, student modelling should be able to deal with such uncertainty [14]. One possible method to handle such uncertainty is by using fuzzy logic. Fuzzy logic was introduced by Lofti Zadeh in 1965 as an approach for computing with words in which words are used in place of numbers for computing and reasoning. This approach tackles ambiguous problems caused by human subjectivity. By using fuzzy modelling approach, an expert's knowledge in verbal descriptions can be transformed into mathematical models through computer algorithms [15].

Studies also reveal the various factors that influence a student's course performance, such as socio-economic, psychological, environmental and individual factors [16].

This study proposes the use of fuzzy modelling approach that considers the various student performance predictors to provide prediction of student course performance at the beginning of a semester. Prediction of performance at the beginning of a semester enables appropriate adjustment and/or intervention to be incorporated into the course design to produce the desired learning outcomes. A fuzzy-based predicting system was developed and a simulation study was conducted to test the performance of the developed system. The applicability of the proposed system was then demonstrated in an undergraduate course. The related works are as follows:

A. Intelligent tutoring system

World Wide Web (WWW) opens new ways of learning. However, "one-size-fits-all" approach used in online tutorials is inappropriate for many students as those tutorials are either too difficult for beginner or too simple and insufficient for more advanced students [13]. This problem could be solved by more intelligent solution that provides adaptive learning material. Intelligent tutoring systems (ITSs) refer to advanced computer-based instructional systems which could adapt the content according to individual learner's needs [13]. The representation of an ITS to an individual student is based on the student's needs and preferences during the teaching process, providing the student a highly personalized learning experience [17]. It serves as an alternative to the traditional "just-put-it-on-the-Web" approach employed in web-based educational courseware [18]. This adaptive feature heavily relies on a student model which keeps the relevant data about a student [13]. In general, the adaptation process will start by getting information about a student, then the system processes this information in order to initialize and update the student model, hence affording adaptation via the student model [17]. Design Patterns Teaching Help System (DEPTHS) is one of the examples of ITS [13].

B. Student Modelling

A student model is a representation of a student's characteristics, containing all the information about the student such as his or her learning style, background knowledge, preference, knowledge level, ability, need and so forth [18]. The process of creating such representation is called student modelling [13]. Reliable student modelling allows experts to diagnose the learner's mental state and knowledge status in order to check the efficiency of teaching as well as to detect possible learning deficiencies. A student model enables an intelligent tutoring system to provide individualized course content in order to support and help students with different backgrounds and knowledge status for achieving their learning objectives. After all, high adaptability in teaching has been proven as a fruitful way to maximize learning results.

C. Fuzzy student modelling

Students' characteristics are often not able to be measured precisely. In addition, the instructional process and teaching strategies provided by an ITS are based on human subjectivity and conceptualizations [14]. A way to deal with such subjectivity is to employ fuzzy logic which was introduced by Lofti A. Zadeh. The integration of fuzzy logic into the student model of an ITS can increase a learner's satisfaction and improve its adaptivity [14]. This is because fuzzy logic can analyze a student's knowledge level, needs and behaviors more precisely and thus, enables the model to make more valid and reliable decisions.

Educational researchers often utilize fuzzy logic techniques in student modelling to deal with uncertainty occurred in diagnosing and assessing students' performance. For example, fuzzy models are used to represent students' profiles as to provide personalized learning materials, quizzes and advices to each student. Moreover, a fuzzy student model to facilitate every student's reasoning process based on the imprecise information acquired through the interaction between the student and his or her computer, predicting the errors possibly to be made in the student's subsequent attempt [19]. The application of fuzzy logic principles in student modelling is attractive since it overcomes computational complexity issues.

II. MATERIALS AND METHOD

A. Designing a fuzzy expert system

The typical steps in designing a fuzzy expert system include specifying the problem, determining linguistic variables, defining fuzzy sets as well as obtaining and constructing fuzzy rules [20]. This study focuses on predicting the course performance of an undergraduate course on Computer Vision.

1) Specify the problem and determine its linguistic variables: Before building an expert system, it is important to first specify the problem to be solved. The problem should be described in terms of knowledge engineering, defining the problem input and output variables as well as their linguistics values and ranges [20]. In this study, the proposed fuzzy-based system is to predict students' course performance by considering several factors that contribute to it. These factors are the input variables of the system while student course performance is the output variable.

The factors that would influence a student's course performance were identified via literature review. A student's effort is a factor that influences his or her course performance [21]. A student's effort considers his or her participation in the learning process as well as the student's daily engagement in studies [16]. Previous educational level will also affect a student's course performance [21]. The student's English literacy and his or her previous educational achievement were identified to define previous educational level. Academic achievement (or test results) is also the major issue an instructor would consider while assessing a student's course performance. Thus, a student's prior knowledge assessment score and the score of a closely related course are used to define his or her academic achievement. Figure 1 depicts the hierarchical structure of all input variables for the proposed predicting system.

2) Determine the universe of discourse: A questionnaire consists of two sections was employed to identify a student's participation in the learning process and his or her daily engagement in studies. The total scores for these sections are 100% and 70% respectively. They correspond to the ranges (or universe of discourse) of both components mentioned while the range of the "individual effort" output variable is given a scale of 100%.

As for the "previous educational level" factor, the input variable "previous educational achievement" is defined by the student's CGPA (Cumulative Grade Point Average) whereas the variable "English literacy" is defined by his or her MUET (Malaysian University English Test) grade. The universe of discourse for these variables are 0.00 to 4.00 and Band 1 to Band 4 respectively while the range of the output variable "previous educational level" is set to the scale from 0 – 10.

As for the "academic achievement" factor, the ranges for input variables "prior knowledge assessment score" is set from 0 to 20% and "related course score" is set from 0 to 100% whereas the universe of discourse of the output variable "academic achievement" is fixed from 0% to 100%.

3) Define the fuzzy sets of variables: One of the most useful approaches for forming variable fuzzy sets is by acquiring knowledge from a single expert or multiple experts [19]. In this study, the number of fuzzy sets, the shapes of the membership functions and the linguistic values of the variables for both "individual effort" and "previous educational level" factors were determined based on an expert's opinion.

On the other hand, a different method was used to define the fuzzy sets and the shapes of membership function of "academic achievement" factor. Historical data was analyzed to gain knowledge instead of eliciting information from the expert. First, the historical data of students from the previous batch of Computer Vision course was collected from the instructor. In order to use the data for analysis, the outliers in the data were removed as they have a significant effect on the mean and the standard deviation. After finding the mean and the standard deviation of the total scores (or grades) obtained by previous batch of students, the parameters were used for plotting a normal distribution. By using the density curve, the probabilities of students gaining each grade were identified. The probabilities were then used to find the minimum marks a student should obtain for prior knowledge assessment score and related course score in order to achieve the targeted grade, via the percent point function (inverse of normal distribution). The minimum scores are useful in this study for determining the range of fuzzy sets for the features "prior knowledge assessment score" and "related course score" in the "academic achievement" category. At the same time, the shapes of the membership functions for each fuzzy set of the variables in the category were defined.

Obtain and construct fuzzy rules. Fuzzy IF-THEN rules play the most important part of a fuzzy-based system in decision making. As mentioned by [19], field experts can be asked to describe the way to solve a problem using its fuzzy linguistic variables to elicit fuzzy rules. In this study, the IF-THEN fuzzy rules were obtained through interviewing the course instructor. His expertise is the source of knowledge for the fuzzy prediction system to do reasoning when predicting student course performance of that course.

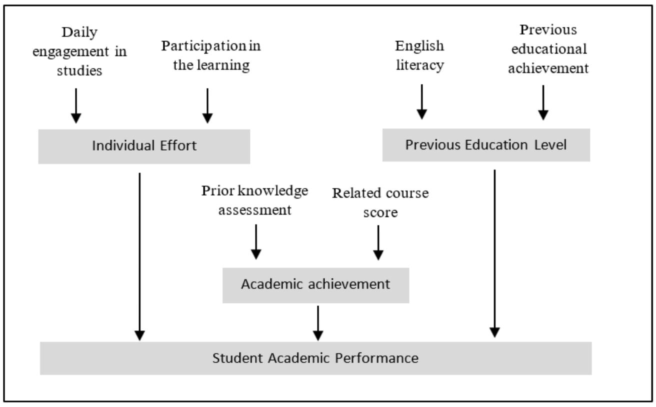

Fig. 1 The hierarchical structure of the student course performance prediction system

B. Developing a fuzzy expert system

It is necessary to encode the fuzzy sets and fuzzy rules for a fuzzy expert system to perform reasoning or inference. The MATLAB Fuzzy Logic Toolbox with graphical editors for designing, analyzing and simulating a fuzzy logic system, is highly recommended as it is intuitive and convenient to use. However, the Fuzzy Logic Toolbox does not support building a Fuzzy Inference System (FIS) using data. In this study, the MATLAB command-line (.m files) was developed. The Mamdani's fuzzy inference method was employed to compute the output (student course performance) of the system.

1) Conduct Simulation Studies: Simulation studies are conducted to check whether the developed fuzzy system meets the requirements specified at the beginning [19]. This process also helps to test the system performance and accuracy as well as to increase its reliability. To evaluate the developed fuzzy-based assessing system, 30 cases of student course performance predicted were computed.

2) Participants: As the historical data collected refers to the scores obtained by previous batch of Computer Vision course students offered by an undergraduate program, the targeted population of this study comprised students who enrolled to the program and took the course at the time of the study. Thirty students were involved as participants of this study.

3) Student Engagement Questionnaire: This study used an simplify version of a well-known instrument to survey student engagement. The instrument collects information on a student's participation in educationally purposeful activities and on how his or her spending of the time weekly. The areas focused are similar with the input variables of the fuzzy predicting system - a student's "participation in a learning procedure" and "daily engagement in studies".

4) Data collection: The collected student data were recorded in Microsoft Excel so that it could be read by the MATLAB software. The fuzzy predicting system developed using MATLAB will then be used to compute course performance of each student in the list. This process is to test the performance and accuracy of the system as well as increases its reliability.

5) Validation and verification: Validation process towards an expert system is to check whether the system does what it was intended to and at an adequate level of accuracy. On the other hand, verification of an expert system is to determine whether the system's knowledge satisfactorily represents the domain knowledge. Results produced via the fuzzy predicting system and human expert were compared to check the similarity. The similarity indicates the accuracy and performance of the fuzzy system; the higher the similarity, the better the system performance.

C. A case Study

The system was applied to an undergraduate Computer Vision course as a case study. Students' performance of the course was computed by the fuzzy system. The predicted performance provides insights to the instructor on how to redesign the learning of the subject matter as well as plan for appropriate interventions for students of different performance level. Appropriate personalized assistance and guidance during the course may prevent them from failing the course.

III. RESULTS AND DISCUSSION

A. Probability distribution of academic achievement

Unlike "individual effort" and "previous educational level" factors, in which the number of fuzzy sets, shapes of the membership functions and linguistic values of the variables were determined by an expert, a historical data was analyzed to build the "academic achievement" probability distribution. The normal distribution of the total marks by the senior students (historical dataset) was generated. In this study, the probabilities of students getting Fail(F), Grade C-, C or C+, B-, B or B+, A- and A were derived from the data set.

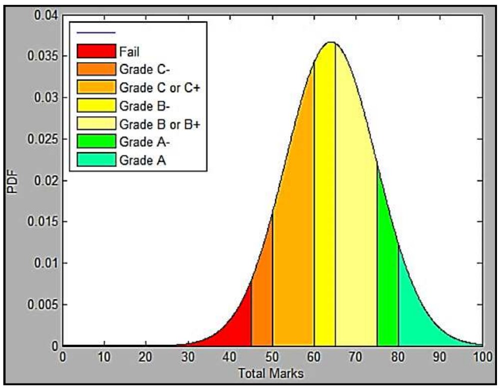

Fig. 2 The normal distribution of the grade

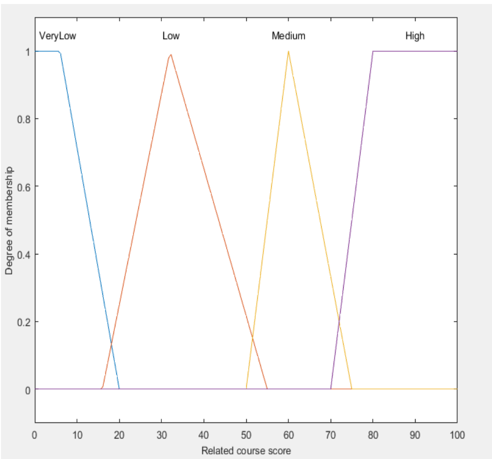

Fig. 3a the fuzzy sets of the variable "Related course score"

Figure 2 depicted the probability (or shaded area under the curve) of the historical dataset. The result reveals that the probability of students failing the Computer Vision course is 0.0399, whereas the probabilities of students gaining Grade C-, C or C+, B-, B or B+, A- and A are 0.0583, 0.2570, 0.1802, 0.3082, 0.0856 and 0.0704 respectively.

The inverse of the normal distribution (or also known as the percent point function) is a technique of working backwards to find the x values which undergoes the transformation $z = (x\text{-}\mu)/\sigma$. The *norminv* (Inverse of the normal cumulative distribution function) command in MATLAB was used to carry out the inverse process. The computed probabilities were used to calculate the minimum scores that need to be obtained by students for prior knowledge assessment and related course in order to achieve certain grade.

Based on the computed results, if a student scores less than 14.50% for the prior knowledge assessment, he or she has a high possibility to fail the subject. Moreover, students who get marks within the range of 14.50-14.92%, 14.93-15.78%, 15.79-16.21%, 16.22-17.07% and 17.08-17.50% are probably gaining Grade C-, C or C+, B-, B or B+ and A- for the course respectively. Students obtaining more than 17.50% will probably get Grade A. If a student scores less than 6.15% in the related course, he or she has a high possibility to fail the subject. Moreover, students who get the marks within the range of 6.15-16.65%, 16.65-37.80%, 37.80-48.40%, 48.40-69.55% and 69.55-80.20% have possibility to obtain Grade C-, C or C+, B-, B or B+ and A- for the course respectively. Students obtaining more than 80.20% will probably get Grade A.

B. Define Linguistic Variables and Fuzzy Sets

The fuzzy sets for all input and output variables were then created. The fuzzy sets for the input variables of "individual effort" and "previous educational level" factors were determined by an expert while for "academic achievement" factor, as explained in Section III, were based on historical data. Figure 3(a) and 3(b) show the fuzzy sets of prior knowledge assessment and related course score. Fuzzy rules were then constructed to derive the fuzzy sets of output variables

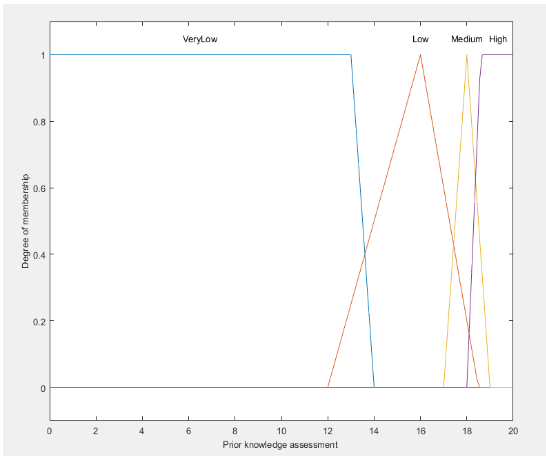

Fig. 3(b) The fuzzy sets of the variable "Prior knowledge assessment

C. Construct Fuzzy Rules

Based on the Matlab software, the fuzzy rule base can be typed in an array form, in which column one represents the index of membership function for first input; column two refers to the index of membership function for second input; column three signifies the index of membership function for output; column four denotes the rule weight; as well as column five stands for the fuzzy operator (where value "1" for fuzzy operator AND; whereas value "2" for fuzzy operator OR). To add a rule base for a FIS, the add rule command can be used. An example of the fuzzy rule set for "individual effort" factor is as follows:

1. If (Participation is Low) and (DailyEngagement is Low) then (IndividualEffort is Low) (1)
2. If (Participation is Low) and (DailyEngagement is Medium) then (IndividualEffort is Low) (1)
3. If (Participation is Low) and (DailyEngagement is High) then (IndividualEffort is Medium) (1)
4. If (Participation is Medium) and (DailyEngagement is Low) then (IndividualEffort is Low) (1)
5. If (Participation is Medium) and (DailyEngagement is Medium) then (IndividualEffort is Medium) (1)
6. If (Participation is Medium) and (DailyEngagement is High) then (IndividualEffort is High) (1)
7. If (Participation is High) and (DailyEngagement is Low) then (IndividualEffort is Medium) (1)
8. If (Participation is High) and (DailyEngagement is Medium) then (IndividualEffort is Medium) (1)
9. If (Participation is High) and (DailyEngagement is High) then (IndividualEffort is High) (1)

D. System Performance

Table 1 shows the student course performance computed by the fuzzy-based predicting system as well as the values of each input variable, namely "participation in learning process", "daily engagement in studies", "CGPA" which represents the previous academic achievement, "MUET grade" that corresponds to English literacy, "prior knowledge assessment score" and "related course score". To carry out the validation process, two subject matter experts were asked to assess the course performance of the 30 student cases in the Table 1.

If the student course performance computed by the fuzzy expert system are as expected by the expert, it shows that the system satisfactorily represents the expert knowledge. To carry out the verification process, student course performance computed through the fuzzy predicting system and evaluated by the human experts were compared to check for similarity (by referring to the remarks column generated by the system in Table 1). Higher similarity indicates better system performance. More than 24 out of 30 (>80%) outcomes computed by the system match the expert's evaluation.

TABLE I
THE FINAL OUTPUT (STUDENT COURSE PERFORMANCE) COMPUTED BY THE FUZZY PREDICTING SYSTEM AND VALUES OF INPUT VARIABLES

No.	Name	Participation (100%)	Engagement (70%)	CGPA	MUET	Prior knowledge (20%)	Related course score (100%)	Performance (Output)	Remarks
1	-	62	60	3.55	3	15	77	74.0	Medium
2	-	68	55	2.87	3	14	67	55.4	Low
3	-	80	67	3.13	3	17	75	87.0	High
4	-	55	43	2.98	4	14	82	72.0	Medium
5	-	76	64	3.56	4	14	78	80.4	High
6	-	76	56	3.34	4	16	67	72.0	Medium
.	-	.	.	.	.	.	.	.	.
.	-	.	.	.	.	.	.	.	.
.	-	.	.	.	.	.	.	.	.
30	-	56	66	3.13	3	15	76	73.3	Medium

IV. CONCLUSION

The developed fuzzy system was verified as the computed outputs are more than 80% similar to what are expected by the expert. It shows that the developed system, to a certain extent, is useful for predicting student course performance before the semester start. Most instructors are the experts in their respective subject matters and able to assess every student's course performance. However, assessment needs a lot of effort and can be time consuming when many students are involved. Prediction becomes more complicated when more factors are taken into consideration. Hence, this proposed fuzzy-based predicting system provides an alternative for instructors evaluate student course performance.

There is only a total of six factors included in the fuzzy assessing system. In fact, there are a lot of factors influencing a student's course performance. Adding more factors as the variables will help to increase the accuracy and performance of the system. Apart from eliciting knowledge from experts, historical data is very useful for gaining information as the data record past real cases. The fuzzy sets of "academic achievement" were determined using the historical data, However, only 150 historical student data were used in this study. Future work may increase the amount of historical data to produce more precise performance of the fuzzy predicting system.

ACKNOWLEDGMENT

This work is partly supported by UNIMAS Scholarship of Teaching and Learning (SoTL) Grant through SoTL/FSKPM/2018(1)/009. The authors would like to thank Universiti Malaysia Sarawak for supporting this work.

REFERENCES

[1] R. M. Gagne, W. W. Wager, K. C. Golas, J. M. Keller, and J. D. Russell, "Principles of instructional design," *Performance improvement*, vol 44(2), pp. 44-46, 2005.

[2] C. M. Cranston, and B. McCort, "A learner analysis experiment: Cognitive style versus learning style in undergraduate nursing education," *Journal of nursing education*, vol. 24(4), pp. 136-138, 1985.

[3] C. Larmuseau, J. Elen, and F. Depaepe, "The influence of students' cognitive and motivational characteristics on students' use of a 4C/ID-based online learning environment and their learning gain," in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 2018, p. 171-180, ACM.

[4] M. J. Kintu, C. Zhu, and E. Kagambe, "Blended learning effectiveness: the relationship between student characteristics, design features and outcomes," *International journal of educational technology in higher education*, vol. 14(1), pp. 7, 2017.

[5] R. F. Kizilcec, M. Pérez-Sanagustín, and J. J. Maldonado, "Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses," *Computers & education*, vol. 104, pp. 18-33, 2017.

[6] M. L. Tan and J. Harun, "Utilizing Concept Map of the Social Presence Characteristics in Social Collaborative Learning Environment for Nurturing Engineering Students' Knowledge Construction Levels," *Advanced science letters*, vol. 24(6), pp. 4561-4564, 2018.

[7] R. Arredondo, "Multiple Regression Analysis of Student's Learning Style, Student's Characteristics and Learning Environment to Predict Student's Performance in an Introductory Accounting Course," Doctoral dissertation, Northcentral University, 2018.

[8] M. D. Price and S. E. Park, "Can Noncognitive Components of Admissions Data Predict Dental Student Performance and Postdoctoral Program Placement?" *Journal of dental education*, vol. 82(10), pp. 1051-1058, 2018.

[9] M. Celepkolu and K. E. Boyer, "Predicting Student Performance Based on Eye Gaze During Collaborative Problem Solving," *Proceedings of Group Interaction Frontiers in Technology* (GIFT 2018). ACM.

[10] A. Soni, V. Kumar, R. Kaur, and D. Hemavathi, "Predicting student performance using data mining techniques," *International journal of pure and applied mathematics*, vol. 119(12), pp. 221-227, 2018.

[11] R. L. Ulloa-Cazarez, C. López-Martín, A. Abran, and C. Yáñez-Márquez, "Prediction of online students performance by means of genetic programming," *Applied artificial intelligence*, vol. 32(9-10), pp. 858-881, 2018.

[12] E. Jembere, R. Rawatlal, and A. W. Pillay, "Matrix Factorisation for Predicting Student Performance," in *7th World Engineering Education Forum* (WEEF), IEEE, pp. 513-518, (2017, November).

[13] Z. Jeremić, J. Jovanović, and D. Gašević, "Student modeling and assessment in intelligent tutoring of software patterns," *Expert systems with applications*, vol. 39(1), pp. 210-222, 2012.

[14] K. Chrysafiadi and M. Virvou, *Advances in personalized web-based education*, Springer International Publishing. 2015.

[15] I. Rashkovsky and M. Margaliot, "Nicholson's blowflies revisited: A fuzzy modeling approach," *Fuzzy sets and systems*, vol.158(10), pp. 1083-1096, 2007.

[16] S. Hijazi and S. Naqvi, "Factors affecting students performance: A case of Private Colleges in Bangladesh," *Journal of sociology*, vol.3(1), pp.12-17, 2006.

[17] C. González, J. C. Burguillo, and M. Llamas, "A qualitative comparison of techniques for student modeling in intelligent tutoring systems," in *Frontiers in education conference, 36th Annual*, pp.13-18, IEEE, (2006, October).

[18] P. Brusilovsky and C. Peylo, "Adaptive and intelligent web-based educational systems," *International journal of artificial intelligence in education*, vol. 13, pp. 159-172, 2003.

[19] G. Goel, S. Lallé, and V. Luengo, "Fuzzy logic representation for student modelling," in *International Conference on Intelligent Tutoring Systems*, 2012, p. 428-433, Springer, Berlin, Heidelberg.

[20] W. T. Ng and C. S. Teh, "Fuzzy students' knowledge modelling system through revised Bloom's taxonomy," *Journal of telecommunication, electronic and computer engineering*, vol. 9(2-9), pp. 169-174, 2017.

[21] S. Ali, Z. Haider, F. Munir, H. Khan, and A. Awais, "Factors contributing to the students course performance: A case study of Islamia University Sub-Campus," *American journal of educational research*, vol.1(8), pp. 283-289, 2013.