MJIT 2017

Malaysian Journal of Industrial Technology

KNOWLEDGE REPRESENTATION ON STUDENT'S ACADEMIC PERFORMANCE USING RULES BASED METHOD

Ismail. Sarerusaenye¹, Ismail.Shahrinaz²

¹Asia Pacific University of Technology & Innovation, Kuala Lumpur, Malaysia

²University Malaysia Kuala Lumpur, Wilayah Persekutuan, Malaysia

asare@apu.edu.my, shahrinaz@unikl.edu.my

Abstract

There are various models introduced in recent years that investigated students' academic performance based on their results, using artificial neural network, rule-based, machine learning, among others. In common scenario at tertiary level education, at least four courses are compulsory for completion by students in the first two semesters, because it would be harder for them to get distinction in academic results when they embark second year onwards. In addition, they need to fulfil course requirements individually or in group, such as tests, quizzes and assignments, ending with exams. It is important to ensure that students are consistently performing throughout their study in ensuring that they graduate on time. With this reason, we propose rule-based method to predict students' potential success in academic, based on their current progress in coursework. A case study on prediction of students' performance in assignments is presented by using GUSC factors adapted from personal knowledge management concept, and techniques of decision tree classifier. Analysis is done on a dataset of group coursework results, categorized into Get, Understand, Share and Connect components (i.e. GUSC), which are considered as attributes for classification in performance prediction. The result includes the analysis process of producing frequency patterns by decision tree with information gain classifier.

Keywords: GUSC; Predictor; Binary classification; Decision tree; Rules Based

1.0 INTRODUCTION

Numerous research studies have attempted to determine the factors that influence program success and to develop suitable prediction models [1]. Example is WISRAS that simplifies the registration process of the graduate students by providing an electronic and interactive registration process. In the WISRAS process, the student fills the course plan page on the intranet site of the department and submits it online, which generates an email confirmation of the course plan submission to both the student and the advisor [2].

*Corresponding author : sare@apu.edu.my

Course non-completion and dropout rates are more common as less able or academically less prepared students are admitted to university. Public funding authorities are also increasingly concerned by the potential waste of public expenditures on students who subsequently fail at university [3]. Every semester, students face difficulties to score each of the registered modules or courses. At least four courses are compulsory for completion and it becomes harder to get distinction result when it turns to second year onwards. In addition, they need to fulfill many requirements of the courses, for example test, quiz, assignment and exam, either as individual or in group.

This paper uses decision tree as a predictor of students' performance in a semester. A part of this study is related to knowledge management, which is useful to understand how to measure students' way of managing knowledge throughout a coursework preparation as an indicator to their potential grades, especially for group assignment. According to [4], agent-mediated personal knowledge management (PKM) processes of individuals could contribute to the emergence of personal intelligence in achieving the collective organizational (i.e. team) goals, demonstrating the bottom up approach from PKM to organizational knowledge management (OKM). For this case, it is an individual student's PKM that would contribute to team's or group's knowledge management, resulting a successful grade in their assignment.

The objectives of this research are to analyze the relationship of GUSC participatory on students' academic performance and to propose a rulebased classification for predicting the value of student's grade. This paper explores the use of decision tree to make predictive analysis for current findings.

2.0 LITERATURE REVIEW

This section is divided into three parts, in order to encapsulate the theories bounded in this research, as per adopted from the previous literature. The three sections are academic performance prediction, rule-based method and decision tree classifier, and GUSC Model for knowledge representation.

Academic Performance Prediction

Assessment is the process of identifying, gathering and interpreting information about students learning. The main purpose of assessment is to provide information on students' achievement and progress and set the direction for ongoing teaching and learning [5]. Based on David Kolb's Model [6], the process of learning involved four stages: concrete experience that diveraes human's feelina; continuous observation on reflection, which assimilates 'think and watch': human perceived abstract and conceptual environment that converge thinking; and active experimentation. For example, a study examined 25 factors that could influence introductory programming performance in students' performance [1]. Each of these factors can be identified at the start of a module when students have had minimal exposure to programming concepts.

A similar study was conducted to determine the relationship between students' demographic attributes, qualification on entry, aptitude test scores, performance in first year courses and their overall performance in the program, to provide the finest quality of education to students [7]. Based on a case study by Koutina and Kermanidis [8], the diagnosis process of students' performance improves as new data becomes available during the academic year, such as students' achievement in written assignments and their in-class presence and participation.

Educational data come in different and very complex formats [9]. Based on a survey by Pena-Ayala [10], classification can be established from student performance modeling, student behavior modeling, assessment, curriculum, domain knowledge, sequencing teachers support and student support and feedback, to produce a mechanism for predicting students' success in academic performance.

Rule-based Method and Decision Tree Classifier

Various models have been introduced to investigate student performance based on students' results, for example using a decision tree, artificial neural network, training optimization, data mining, machine learning and more. A successful learner should be able to progress from individual examples to broader generalization. This is also referred to as inductive reasoning or inductive inference [11]. A decision tree is a kind of non-cyclic flowchart, as shown in Figure 1 [12].

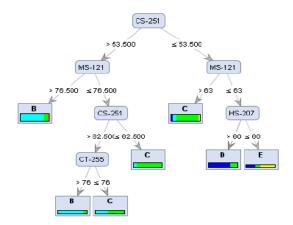


Fig. 1. Example a Decision Tree with Information Gain (Asif et al., 2015)

In rule-based method, classification is the first step to do reasoning, and it is something normal where the numbers of training examples from different classes are very imbalanced [13]. Previous studies have been reviewed in similar topic, for example a predictor that applied learning to match method and formalize as the learning problem [14]. Here, a fundamental solution to the problem performs semantic matching and uses regularized mapping method. A significant study is found in analyzing student academic performance by Asif et al. [12], in which the result showed that it is possible to predict the graduation performance in the students' fourth year of study using only pre-university marks and marks of their first and second year courses.

In addition, Jia and Maloney [3] also used predictive modeling to identify students at risk of poor performance in their university. They developed simple rule-based tools that allowed university to identify and intervene on vulnerable students when they first arrived on campus. Cognitive predictor can also be used to make advance decision making. Recent research has found the measures of inductive reasoning, such as Raven's Advanced Progressive Matrices, which is applied across the setting [15]. On the other hand, an investigation has found that the mean GCSE was not the best predictor of success for individual A-level subjects using aggregating assessment results method, when they predict future examination in University of Cambridge [16]. For this case, the best predictor proved to be the sum of the square roots of the best five GCSE grades. This measure rewards a steady performance on a limited range of courses.

In relation to this paper, Minaei-Bidgoli et al. [17] conducted a similar study, in which they tried to predict the final test grades for students enrolled in a web-based course. In this study, three different

classifications for the students' results were used: dividing results into two classes (pass and fail), three classes (high, middle and low), or into 9 classes, according to their grade. Several learning algorithms were compared: decision trees, neural networks, naïve Bayes, logistic regression, support vector machines, and KNN with feature weights adjusted by a genetic algorithm. The study concluded that each of the machine learning algorithms performed satisfactorily with all performing roughly equally.

GUSC Model for Knowledge Representation

Knowledge representation is the way how to make the system intelligent by four major studies: logic, rules, semantic net and frame. It used inference mechanism to make it intelligence based on expert knowledge inside [18]. From the understanding of knowledge as explained by Owen and Horvath [19], the challenge lies in the interpretation of data that leads to information and finally knowledge, in which an intelligent system is expected to be able to handle and manage. Nowadays, this machine language called as knowledge based system. Knowledge based is techniques to support human decision making which covers the implementation of decision support knowledge acquisition, rules based, meth representation and system architecture.[20]. Looking at the needs to make sense of data in order to get to good interpretation of knowledge, researchers have dived into the personal level of human's knowledge management. This is called personal knowledge management (PKM), in which recent research has postulated that effective PKM processes could be observed at individual level as well as at software agent level. Since software agent is part of artificial intelligence domain, it is deemed fit to apply the PKM concept in this research too.

From PKM domain, the GUSC Model is introduced to define the processes at agent level, namely get knowledge (G), understand knowledge (U), share knowledge (S), and connect to knowledge sources (C) [21]. These four processes are simplified and derived from a range of previous literature on PKM, both from the social science and technical computer science. The validation of this model is presented in various forms, including assigning software agents with GUSC roles [21], and analyzing human's PKM processes through online medium [22]. With these contributions, it is proven that reasoning can be done at software agent level using GUSC concept, in which rule-based method is found to be suitable. In other words, PKM processes can be used for knowledge representation, in which the GUSC is used to represent knowledge in prediction analysis.

3.0 METHODOLOGY

This research takes a sample of group assignments, submitted by 64 students in a semester of 2015 for one course. The students are from two programs offered at a Malaysian private university: software engineering and intelligence system. These 64 students made up 23 groups, meaning that the total group assignments submitted was 23. The data is coded by group, with identification codes generated from G001, G002, G003, and so on until G023. On the other hand, the data for individuals in the group is coded as \$1, \$2 and \$3 for each group, hence the repetition in the codes for each group. These codes will be presented in the next section on data analysis.

The overall research process performed in this study is as follows:

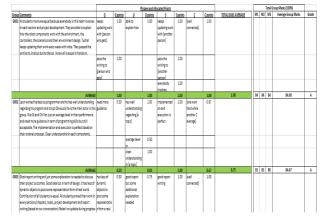
- 1. Gather the marks of each student; Group the cohort based on marking scheme
- 2. Categorize and split the marks into individual and group portions; Identify the components of GUSC from the marks and comments
- Generate a decision tree with information gained and perform analysis; Perform rulebased classification by individual portion (45%) and group portion (55%)
- 4. Update the GUSC variables with the fitness and generate the pattern

The process started with gathering the data from the selected semester, based on criteria that seemed fit for the purpose of this research. The assignment results of this class were categorized according to the original portions: individual and group portions.

The group assignment for this particular course inclusive of four portions: individual portion (45%), comprising question and answer portion performed during presentation (15%) and individual workload (30%); and group portion (55%), comprising proposal (5%), design (10%), project (10%) research and module integration (20%) and documentation (10%). The GUSC components were identified through content analysis on comments or remarks written for each assignment, and quantitative measurement was assigned to each components found based on previous work. This study carries out a process of categorizing the dataset into GUSC using the basic scale of 5-Likert scale with 0.00 for 'not exist', 0.25 for 'somewhat exist', 0.50 for 'half exist', 0.75 for 'almost completely exist', and 1.00 for 'obviously exist'. The content of the comments written for each group assignment was analyzed to produce the GUSC

average. Average of total GUSC is calculated for both, group portion and individual portion (as shown in Table 2 on group differential, and Table 3 on individual differential), based on the range of GUSC average versus potential grade shown in Table 1, presented in the next section. The purpose of this calculation is to measure and decide on the weightage in each component to be used in classification of grading. In this analysis, assume the ranges of potential grade are distinction (A), credit (B), pass (C) and fail (D).

Fig. 2. Step one of GUSC Analysis from Sample	Э
Comments on Group Assignment	



The decision tree was then generated for further analysis, with rule-based classification performed on the individual and group portions. The GUSC variables were then updated with the fitness to generate the pattern for prediction. We tried to extract all possible interesting frequent patterns based on the individual variables. Then, we applied classification method in order to predict potential factor types based on GUSC. Figure 3 shows the second step in this research process, in which the GUSC average is calculated for individual level, to produce a personal GUSC values and total of group GUSC values.

	Qualitative Analysis based on Comments						
Student name	G	U	S	C	GUSC Total	GUSC Average	
ID	1.000	1.000	1.000	1.000	4.000	1.000	
TJ	0.500	1.000	1.000	1.000	3.500	0.875	
KJW	1.000	1.000	0.500	1.000	3.500	0.875	
GUSC Total	2.500	3.000	2.500	3.000	11.000	2.750	
GUSC Average	0.833	1.000	0.833	1.000	3.667	0.917	
LWQ	0.500	0.500	0.500	1.000	2.500	0.625	
MCY	0.500	0.500	0.500	1.000	2.500	0.625	
NL	1.000	1.000	1.000	1.000	4.000	1.000	
GUSC Total	2.000	2.000	2.000	3.000	9.000	2.250	
GUSC Average	0.667	0.667	0.667	1.000	3.000	0.750	
CPE	0.000	0.750	0.750	0.000	1.500	0.375	
D	0.000	0.750	0.750	0.000	1.500	0.375	
AD	0.000	0.750	0.750	0.000	1.500	0.375	
GUSC Total	0.000	2.250	2.250	0.000	4.500	1.125	
GUSC Average	0.000	0.750	0.750	0.000	1.500	0.375	
KJT	0.250	0.750	0.500	0.250	1.750	0.438	
CHC	0.250	0.750	0.500	0.250	3.500	0.438	
S	0.250	0.750	0.500	0.250	1.750	0.438	
GUSC Total	0.750	2.250	1.500	0.750	5.250	1.313	
GUSC Average	0.250	0.750	0.500	0.250	2.333	0.438	
TW	0.000	0.250	0.250	0.250	0.750	0.188	
G	0.000	0.750	0.500	0.250	1.500	0.375	
GUSC Total	0.000	1.000	0.750	0.500	2.250	0.563	
GUSC Average	0.000	0.500	0.375	0.250	1.125	0.281	

Fig. 3. Step two on sample GUSC Analysis from Comments on Group Assignment

4.0 RESULTS AND FINDINGS

In a rule induction algorithm, IF-THEN rules are extracted sequentially, i.e. one after the other, from the training dataset based on weight of GUSC, as opposed to a decision tree that generate IF-THEN rules in parallel. Each rule for a given class should have a high coverage and a high accuracy, where coverage is measured by the proportion of the data to which the rule applies. Once a rule is learned, the corresponding subset is excluded from the data and a new rule is learned on the remaining dataset. In this study, we used Rule Induction with information gain as a criterion to learn rules. As for decision trees, the results of a rule induction algorithm are easily interpretable for humans.

As mentioned in the previous section, the range of GUSC average is pre-determined to facilitate further analysis in producing academic grading prediction.

Table 1. The Range of GUSC vs Potential Grade

Range of GUSC Average	Potential Grade		
0.80 - 1.00	A including A- and A+		
0.65 - 0.79	B including B- and B+		
0.50 - 0.64	C including C- and C+		
Below 0.50	D		

Table 1 shows the range of GUSC average that was decided to be used in this research during data analysis. This becomes the benchmark for group and individual differentials shown in Table 2 and 3.

Based on the range of GUSC average in Table 1, Table 2 is produced to show the comparison between the calculated GUSC average and actual grade attained by the students, in groups. If the range in Table 1 is perceived true, then the fourth column in Table 2 shows the accuracy of the total GUSC average in predicting the actual grade. The difference between the expected grade based on total GUSC average and the actual grade is calculated and presented in the last column.

Table 2. Sample of Group Differential

Group	Total GUSC Average	Actual Grade Attained	IF Table 3 is TRUE, THEN	Difference from potential Grade	
G001 1.00		Α	TRUE	0	
G002	0.75	А	FALSE	0.05	
G003	0.74	A-	FALSE	0.06	
G004	0.77	В	TRUE	0	
G005	0.56	В	FALSE	0.09	
G006	1.00	A-	TRUE	0	
G007	0.47	A+	FALSE	0.33	
G008	0.93	А	TRUE	0	
G009	0.65	А	FALSE	0.15	
G010	0.22	В	FALSE	0.43	

Based on the range of GUSC average in Table 1, Table 3 is produced to show the comparison between the calculated GUSC average and actual grade attained by the students, as individuals. If the range in Table 1 is perceived true, then the sixth column in Table 3 shows the accuracy of the total GUSC average in predicting the actual grade. The difference between the expected grade based on total GUSC average and the actual grade is calculated and presented in the last column. The total GUSC average values shown in Table 2 and 3 are different. This is due to the content analysis done at different level of understanding. In Table 3, the total GUSC average values takes into account the total of individual GUSC average, but Table 2 measure the GUSC average as collective at group level.

Then the values from these GUSC components are analyzed to find differences among group and individual portions that would reach each grade. Study on both, group and individual differentials, are taken as a result that used in classification of rules based. Induction rule is powerful to state all facts and relationships about the problem. For example, from the result of distinction groups (A+, A and A-), it is clearly shown that A+ group has strong connectivity among the members, hence being given different grade than others.

Table 3. Sample of Individual Differential

Group	Individual	GUSC Average	Total GUSC Average	Actual Grade	IF Table 3 is TRUE, THEN	Difference from potential Grade
G001	S1	1	0.917	Λ	TRUE	0
	S2	0.875		А	TRUE	0
	S3	0.875		Λ	TRUE	0
G002	\$1	0.625	0.75	В	FALSE	0.025
	S2	0.625		В	FALSE	0.025
	S3	1		Α	TRUE	0
G003	S1	0.375	0.375	D	FALSE	0.125
	S2	0.375		D	FALSE	0.125
	S3	0.375		D	FALSE	0.125
G004	S1	0.438	0.438	D	FALSE	0.062
	S2	0.438		D	FALSE	0.062
	S3	0.438		D	FALSE	0.062
G005	S1	0.188	0.281	D	FALSE	0.312
	S2	0.375		В	FALSE	0.125
G006	S1	1	0.938	А	TRUE	0
	S2	0.875		А	TRUE	0
	S3	0.938		А	TRUE	0

Another example of this classification is the comparison between the cases of distinction: group A with average total GUSC is 0.65 and group A- with average total GUSC is 0.66. This study finds that strong connectivity has relationship to their result. Besides, for analysis of groups that attained credit (i.e. grade B), this study finds near to no reflection on all GUSC in group activities, but the lower value = 0.25 in Get (G) component still exists and it could be possible as one rule.

During the analysis process, this study generated 22 probabilities based on grades in group portions, in which the number of grades are A = 9, A = 4, A = 3, B = 3, B = 1 and B = 1, C + = 1 and none for grade C and C-. On another hand, 17 probabilities based on individual portions, in which the number of grades are A = 5, A = none, A = 3, B = 4, B = none, B = none,

C=none, C-=none, C+=none and grade D=5. Based on these possibility results, we created the rules using GUSC components for both group and individual.

5.0 DISCUSSION

After experimenting 39 probabilities of analyses and rules generation, the knowledge could be presented using a decision tree. It is found that the result of group portions clearly show the overall component of Connect (C) that reach the highest score. The students are strongly recommended to achieve minimum 50% score, in which they need to be well-connected. Based on the case study, many students that have highest score on connection are active, always keep updating their assignment progress, they are the one who works hard in the group and have good teamwork skill.

This study also found that highest score in component of Get (G) and Share (S) clearly gives a big impact to student's result. The sample from the groups that attained grades A and B proves this. Students are strongly recommended to collect all possible information for their assignment and from among team members even though they could not create great project, because teamwork could help them to score distinction, by providing good justification during presentation, balanced or equal in workload portion, and having everybody aware on his/her own work.

It is found that the result of individual portions clearly shows that overall component of Get (G) and Understand (U) reached the highest score. Students are strongly recommended to achieve minimum 50% in both components in order to get distinction. This means that the students are encouraged to be more interactive with the lecturer during classes, progress report and presentation, and become main actor in the group (i.e. shows leadership), have a good understanding of knowledge and are able to handle all portions as required in the assignment.

This study has found that minimum 75% of Get (G), 50% to 75% of Understand (U) and 50% of Connect (C) will be awarded B grade. This means that the students need to keep in touch with their group members, and keep updating their parts with team members.

6.0 CONCLUSION

From the analysis performed on both group and individual portions of a group assignment, it is found that is it important for these components to support one another, especially during progress report and presentation. For example, there are certain cases in this sampling dataset that other group members helped to prepare proposal, and answered on behalf of other students during progress report and presentation. These activities are supposed to be under group portion but indirectly it gives a big significance and contributes to individual marks.

Strong connectivity in teamwork is obviously a key to great grade achievement as an individual. As proven in this case, when the students helped their group members, the lecturer would be able to give some marks for the studnets' individual portion. In addition, it is advised that the lecturers teaching the course should be the one administering the GUSC analysis on assignment comments, because they understand how the knowledge is managed within the groups. Lecturers are also advised to provide comprehensive and descriptive comments based on his/her observation because these comments are important in generating well-defined prediction. The lecturers are the ones who have better insight on the quality of the students' assignment, which is the tacit knowledge.

As recommendation for future work, optimization algorithms are the best solution for dealing with problems in which a best solution can be represented as a point or surface in an n-dimensional space. Hypotheses are plotted in this space and seeded with an initial velocity, as well as a communication channel between the components. Components then move through the solution space, and are evaluated according to some fitness criterion after each time step.

References

[1] Bergin, S., Mooney, A. Ghent, J., Quille, K. Using Machine Learning Techniques to Predict Introductory Programming Performance, International Journal of Computer Science and Software Engineering (IJCSSE), 4(12), (2015).

[2] R. Naini, R. S. Sadasivam, M. M. Tanik, A Web-based Interactive Student Advising System using Java Frameworks (2008), Southeastcon.

[3] Jia, P., Maloney, T. Using Predictive Modeling to Identify Students at Risk of Poor University Outcomes, (2014).

[4] Ismail, S., Ahmad, M.S., Hassan, Z. Emerging personal intelligence in collective goals: data analysis on the bottom-up approach from PKM to OKM, 17(6), Emerald Group Publishing Limited, (2013), pp.973-99.

[5] Opara, I.M., Onyekuru, B.U., Njoku, J.U. Predictive Power of School Based Assessment Scores On Students'Achievement in Junior Secondary Certificate Examination (JSCE) in English And Mathematics, Journal of Education and Practice, 6(9), (2015).

[6] Kolb, D. Experiental Learning Experience as the Source of Learning and Development (2nd ed.), Pearson Education, (2014).

[7] Golding, P., McNamarah, S. Predicting Academic Performance in the School of Computing & Information Technology (SCIT), in Proceedings of 35th ASEE /IEEE Frontiers in Education Conference, (2005).

[8] Koutina, M., Kermanidis, K.L. Predicting Postgraduate Students' Performance Using Machine Learning Techniques, IFIP AICT, 364, (2011), pp.159-168.

[9] Saneifar, R., Abadeh, M.S. Association Rule Discovery for Student Performance Prediction Using Metaheuristic Algorithms, CSIT, (2015).

[10] Peña-Ayala, A. Educational data mining: A survey and a data mining-based analysis of recent works, Expert systems with applications, 41(4),(2014), pp.1432-1462.

[11] Shwartz, S.S., David, S.B. Understanding Machine Learning: From Theory to Algorithm, Cambridge University Press, (2014).

[12] Asif, R., Merceron, A., Pathan, M.K. Predicting Student Academic Performance at Degree Level: A Case Study, International Journal of Intelligent Systems and Applications, (2015), pp.49-61.

[13] Jordan, M.I., Mitchell, T.M. Machine Learning : Trends, Prespectives and Prospects, 349(6245), (2015).

[14] Xu, J., Lu, Z, Chen, T., Li, H. Learning to Match, (2014).

[15] Klein, R.M. et al. Cognitive Predictors and Age-Based Adverse Impact Among Business Executives, Journal of Applied Psychology, 100(5), American Psychological Association, (2015), pp.1497–1510.

[16] Bell, J.F. Methods of aggregating assessment results to predict future examination performance, Research and Evaluation Division University of Cambridge Local Examinations Syndicate, (2014).

[17] Minaei-Bidgoli, B., Kashy, D.A., Kortmeyer, G. and Punch, W.F. Predicting student performance: An application of data mining methods with an educational Web-based system, in 33rd Annual Frontiers in Education, 1, (2003), pp.T2A–18.

[18] Poonar Tanwar, T.v Prasad, Kamlesh Datta, Hybrid Techniques for Knowledge Representation and Comparative Study, Springer (2011).

[19] Owen, R., Horváth, I. Towards product-related knowledge asset warehousing in enterprises, in Proceedings of the 4th International Symposium on Tools and Methods of Competitive Engineering TMCE (2002), p. 155-70.

[20] H.Fujita, J.Lu, Knowledge Based Systems, Elsevier, (2015).

[21] Ismail, S., Ahmad, M.S. Emergence of Personal Knowledge Management Processes within Multiagent Roles, in D. Richards and B.H. Kang (Eds.), Knowledge Management and Acquisition for Intelligent Systems, Lecture Notes in Artificial Intelligence, LNAI 7457, Springer: Heidelberg, (2012), pp. 221–228.

[22] Ismail, S., Abdul Latif, R., Ahmad, M.S. Learning Environment over the Social Network in Personal Knowledge Management, in Proceedings of International Workshop on Collaboration and Intelligence in Blended Learning, (2012), pp.2-12.