Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Predicting Student Performance through Bloom's Taxonomy using Data Mining Framework for Educational Decision Making

Dinesh Singh[1], Amit Kumar[1*], Mahabir Singh[2]
[1]*Department of Computer Science and Engineering, DCRUST, Murthal, India-131001*
[2]*Ex-COE, DCRUST, Murthal, India-131001*

ARTICLE INFO

Received: 24 Sept 2024

Revised: 16 Oct 2024

Accepted: 28 Nov 2024

ABSTRACT

Predicting student performance is an invaluable undertaking within the domain of educational decision-making. A novel methodology is presented in this study, which combines sophisticated data mining techniques with Bloom's Taxonomy, a widely recognised educational framework. The study seeks to improve the precision and comprehensiveness of predictive models pertaining to student performance through the utilisation of this fusion.By providing a structured hierarchy of cognitive processes, Bloom's Taxonomy enables educators to gain a more comprehensive comprehension of the learning outcomes of their students. By implementing data mining techniques, including classification and clustering algorithms, this framework empowers instructors to derive significant insights from a wide range of student datasets. These observations not only facilitate the forecasting of individual student achievement but also provide guidance for the design of instructional approaches and curricula. The framework that has been proposed signifies a substantial progression in the realm of educational decision-making. It offers instructors a potent instrument to enhance learning experiences and promote scholastic achievement.

Keywords: Bloom's Taxonomy, Data mining, Educational decision-making, Cognitive processes, Instructional strategies, Curriculum development.

INTRODUCTION

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna

1. INTRODUCTION

Predicting student performance through Bloom's Taxonomy using a data mining framework for educational decision making is a topic that has gained attention in the field of educational data mining [1]. Several papers have explored the use of machine learning techniques to analyze student data and improve decision-making processes in educational institutions [2]. These papers discuss the development of models that can accurately predict student performance and identify factors that influence it [3]. The papers emphasize the potential benefits of such models, including helping low-performing students improve their academic metrics, increasing institutional effectiveness, and improving rankings and reputations [4]. Overall, these papers provide valuable insights into the use of data mining frameworks for educational decision making and highlight the potential impact on student performance and institutional outcomes [5].

1.1 Understanding Bloom's Taxonomy and Assessment Criteria

Bloom's Taxonomy is an educational framework employed to classify distinct tiers of cognitive abilities that may be acquired by an individual [6]. Each of the six categories in the taxonomy corresponds to a distinct cognitive ability, ranging in complexity from most basic to most advance. Educators may employ the assessment criteria that correspond to each level to gauge students' comprehension and mastery of the subject matter.

1.1.1 Knowledge

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

A foundational element of Bloom's Taxonomy is knowledge. This tier encompasses the retrieval of data, facts, or information. By recollecting previously learned material, pupils exhibit their comprehension. Criteria for evaluation at this level may consist of:

- Recall of specific facts or information
- Recognition of terms, concepts, or principles
- Ability to identify key elements or details
- Memorization of dates, events, or formulas

1.1.2 Comprehension

Comprehension is a cognitive process that augments prior knowledge by elucidating the significance of the acquired information. Students are permitted to define concepts or notions in their own words. Criteria for evaluating comprehension may comprise:

- Interpretation of information
- Summarization of main ideas
- Paraphrasing or restating concepts
- Clarification of relationships between concepts
- Demonstration of understanding through examples or illustrations

1.1.3 Application

Students must apply their comprehension and knowledge to novel situations or contexts in order to obtain application. They are capable of applying conceptual understanding to practical situations and resolving issues. The application evaluation criteria may comprise:

- Application of theories or principles to real-world situations
- Problem-solving using learned concepts
- Demonstration of skills in practical tasks
- Utilization of knowledge in different contexts
- Transfer of learning to novel situations

1.1.4 Analysis

Analysis entails deconstructing data into its constituent elements and comprehending their interrelationships. Students are capable of discerning patterns, trends, and causes and effects in the subject matter. Criteria for analysis evaluation may consist of:

- Identification of patterns or trends
- Differentiation between parts or elements
- Examination of relationships or connections
- Recognition of underlying themes or concepts
- Evaluation of evidence or support for arguments

1.1.5 Synthesis

To produce something novel or original, synthesis involves the combination of diverse components. On the basis of their comprehension of the subject matter, pupils are capable of generating concepts, proposing solutions, and forming unusual viewpoints. Consider the following as evaluation criteria for synthesis:

- Creation of new ideas or solutions
- Integration of diverse perspectives or sources
- Development of original concepts or theories
- Design of projects or products based on learned principles
- Synthesizing information to form coherent arguments or explanations

1.1.6 Evaluation

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Assigning judgements or assessments grounded in criteria and standards constitutes evaluation, the pinnacle of Bloom's Taxonomy. Students are capable of carrying out informed judgements, defending their positions, and analysing information critically. The evaluation process may incorporate the following criteria:

- Critical analysis of arguments or viewpoints
- Assessment of the validity or reliability of information
- Justification of opinions or decisions
- Application of criteria to evaluate outcomes or solutions
- Synthesis of information to form reasoned judgments or recommendations

Bloom's Taxonomy [7] offers an organised framework for instructors to evaluate and cultivate the cognitive abilities of pupils spanning various degrees of intricacy, commencing with rudimentary knowledge acquisition and progressing to critical analysis and synthesis. Through the process of aligning assessment criteria with each level of the taxonomy, educators are able to efficiently gauge student learning and foster a more profound comprehension and mastery of the subject matter.

1.2 Significance of data mining in educational decision making

Through the discovery of hidden patterns in educational data and the prediction of student performance, data mining plays a crucial role in educational decision making [8]. It helps in improving student academic achievement, enhancing the teaching/learning environment, and supporting decision-making in educational systems [9]. Educational data mining tools and techniques provide valuable information about students, such as enrollment, weak students, and factors that contribute to dropout. It enables institutes to allocate resources effectively, apply corrective strategies, and enhance student success in courses [10]. Data mining also aids in curriculum planning, student profiling, and predicting student performance [11]. Large educational databases can be analysed using data mining, which yields useful information for decision-making and a deeper comprehension of students and their learning environments [12].

2. LITERATURE REVIEW

A Dutta et.al (2023) [13] proposed the educational data mining (EDM) has surfaced as a method to improve the academic progress and learning experiences of students through the application of machine learning and data mining techniques. EDM techniques were commonly utilised to improve assessment systems, with an emphasis on evaluations that were driven by marks. The endeavour of creating an assessment framework to determine the cognitive ability of candidates in response to difficult questions proved to be a formidable one. This investigation introduced an innovative approach to dynamically rank candidates and create an online assessment system that is outcome-driven to effectively assess cognitive abilities. Using Bloom's taxonomy of education, the questions were divided into discrete cognitive areas. The questions were segmented utilising the Jenks Natural Breaks Optimisation technique, which produced discrete question clusters that corresponded to different cognitive levels. The candidates for each cognitive group were evaluated using questions from these groups, and their performance was assessed through a combination of question mark totals and their aptitude for answering questions of different levels of difficulty.

Y Cheng et.al (2021) [14] discussed the learning process evaluation was beneficial for personalised online learning. Real-time cognitive assessment during online sessions helped monitor participants' cognitive states and improve learning tactics. Current cognitive evaluation methods included manual coding or traditional machine learning, which were time-consuming and laborious. These methods failed to extract implicit cognitive semantic information from unstructured text input, making cognitive evaluation ineffective. The study used unstructured interactive text data from 9167 MOOC forum students to empirically analyse cognitive objectives using Bloom's taxonomy. The AM-BiGRU-CNN method was the most accurate, measuring six cognitive levels with 84.21% accuracy and a 91.77% F1-Score at the creating level. Deep neural network approaches can uncover latent cognitive traits in text data, enabling automated cognitive assessment of learners. This study provides a technical framework for assessing students' cognitive levels in online learning settings, affecting personalised learning, teacher interventions, and resource suggestions.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

U Omer et.al (2020) [15]In initial level programming courses, students often encountered difficulties in grasping programming concepts. Learning analytics studies primarily relied on anecdotal evidence for assessment, neglecting to evaluate learning across different levels of cognition for particular concepts. Moreover, previous research failed to adequately investigate how learners' cognitive performance influenced later stages of the course. Addressing this gap required implementing more detailed and systematic cognitive analysis approaches to effectively sustain courses on programming in computer science and associated disciplines. This framework provided structure to concept data through concept mapping and examined learners' cognitive progression on related concepts using assessment data. The evaluation of performance prediction on more complex programming concepts was done utilising metrics derived from learners' cognitive maps, generated through the framework's layers. The maximum prediction accuracy ranged from 64.81% to 90.86%, surpassing accuracies reported in most related studies.

Bloom's Taxonomy (BT) is still one of the most popular frameworks for creating and navigating educational exams, according to Makhlouf et al. (2020) [16]. A considerable quantity of time and effort was recently devoted by a number of academics to automating the process of categorising test queries into groups based on BT. In the paper, a comprehensive overview of researchers was provided. Further investigation was intended to broaden the scope of the instruction set through the testing of different categories of queries. It was suggested to integrate several feature extraction and selection methods.

Z Ullah et.al (2019) [17] determined the evaluation of computer programming students posed a difficulty for educators, especially those teaching at the introductory level due to the substantial number of enrolled students. An innovative methodology for assessing the programming proficiency of students was implemented, which made use of Bloom's taxonomy. In contrast to prior investigations, this methodology employed a direct extraction of competencies from written code according to Bloom's cognitive levels, as opposed to manually mapping queries to these levels. An introductory course in Java is implemented the rule-based assessment method on student code, with an emphasis on fundamental concepts of structured Java programming, including modularity, recurrently, and selection. The rule-based approach was utilised to acquire and process data from 213 students, which was subsequently validated against manual assessment. A comparative statistical analysis validated the proposed rule-based assessment method's dependability.

Uma et al. (2017) [18] stated that a significant factor element influencing success of a school and its students is the calibre of its teaching and learning. To meet the rigorous standards expected of educators and to mitigate the discrepancy between instructional approaches and learning outcomes, a substantial restructuring of the assessment system was necessary. As a significant metric of performance evaluation, assessment was intended to afford students the chance to apply their aptitude for critical thinking. A revised iteration of Bloom's Taxonomy was utilised to assess critical thinking abilities, which were examined in greater detail in this study. Moreover, the overall quality of the system was ascertained through an analysis of the evaluation's intricacy. A weighted data mining approach was employed by the researchers in this study to methodically categorise Bloom's taxonomies and the logical apex states that are associated with them.

A Parkavi et al. (2017) [19] suggested the employing data mining techniques, the patterns of data in various disciplines were analysed. Dialogue with decision-making authorities was conducted in accordance with the findings of the analyses. Data mining techniques were implemented within the educational sector with the aim of enhancing outcomes. The authors of this paper conducted research by developing a tool and algorithm for determining the level of course knowledge among students through the application of data mining techniques. This facilitated the faculty and students in implementing essential corrective measures to enhance their academic performance.

3. METHODOLOGY & EXPERIMENTATION

3.1 Data collection

To measure students' skill in hands-on programming, this study used practical assessments from an introductory course at a prestigious university. To protect participant privacy, the university name is kept secret. A prior study was supported by practical test data. Only 261 of 298 registered students took the test. Students were given easy, medium, and hard questions. Furthermore, the development of the assessment queries was guided by the programming instructors' perspectives. In addition to the allotted time of two and a half hours, which was the intended duration,

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

fifteen minutes were granted in response to student requests. The additional time allotment provides support. 213 pupils submitted their as programming code files.

The files underwent evaluation in accordance with the standards outlined in Section 2. The competencies that were attained were then associated with the appropriate cognitive levels of Bloom's taxonomy. The process of mapping was executed utilising a rule-based approach that was informed by benchmark criteria derived from the literature on theory. A rule-based approach consists primarily of if–then statements, with the "if" component denoting the criteria and the "then" component specifying the achieved cognitive level. Given the established fact that every subject possesses unique frameworks and that a student may employ various frameworks within a single programme or the same framework multiple times, the percentage of competencies attained for each subject was computed utilising the subsequent equation:

$$\text{Struct} = \frac{\text{TStructs} - \text{EStructs}}{\text{TStructs}} \times 100$$

In which Struct calculates a specific percentage for a given structure; EStructs represent the defects, while TStructs denotes the total number of structures utilised in a programme.

It is noteworthy to mention that in contrast to the previous research, which employed Bloom's taxonomy to generate evaluation questions and pre-map them to cognitive levels, the current study evaluates student competency through written programmes without pre-establishing questions in accordance with cognitive levels.

3.2 Data Analysis

A statistical method that combines several approaches to construct a series of interactions between one or more dependent and independent variables, structural equation modelling (SEM), was used to evaluate the assessment data. Experimental data are utilised to model, estimate, and test hypotheses using SEM. This research employed variance-based partial least square SEM (PLS-SEM), a technique that has been recognised by numerous scholars as an advantageous approach for estimating SEM. In addition, PLS-SEM offers numerous advantages to researchers utilising SEM, including the ability to handle non normal data, small sample sizes, and formative models.

The structural model is an inner model, while the measurement model is an exterior model. The conceptual path model of competency assessment, illustrated in Figure 2, is founded upon the PLS-SEM theory. This model effectively delineates the measurement and structural components.

3.2.1 Model specification

The path model depicted in Figure 2 was constructed in accordance with the elucidated logic and theory. The ovoid shapes represent the constructs, the arrows symbolise the relationships between them, and the rectangle shapes denote the variables or indicators. The six variables represent the cognitive levels mentioned earlier, while the indicators represent fundamental programming concepts—specifically, modularity, repetition, and selection.

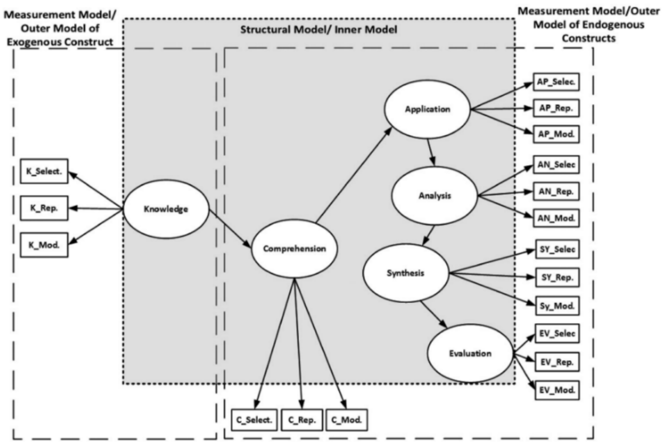

FIGURE 2: Conceptual path model for competency assessment

Independent variables are those that exert a causal influence without being influenced by other variables. Independent variables do not have arrows pointing at them. Thus, "knowledge" is the independent variable in Figure 2. It has the capacity to influence others but is not susceptible to being influenced by them; thus, it is not denoted by an arrow.

The variables with an arrow pointing to them are known as dependent variables since they are impacted by the actions of other people. Application, analysis, synthesis, appraisal, and comprehension are the dependent variables.

3.2.2 Measurement model

The correlation between construct and indicators is manifested in measurement model. It is additionally denoted as a peripheral model in PLS-SEM. The purpose of this model evaluation is to ascertain the validity and reliability of the construct measures, which are fundamental requirements for the structural model evaluation. The evaluation of the measurement model in this research was conducted using the procedures outlined: assessing indicator loadings, ensuring consistency and reliability, and examining composite and convergent validity.

3.2.3 Structural model

An inner model, the structural model illustrates the relationship between the measures of the constructs. The assessment of the structural model's quality is predicated on its predictive ability with respect to the dependent variables. For this specific model type, it is imperative that the measuring model's valuation be properly finished.

The structural model was assessed in the current study using the following assessment criteria: collinearity, path coefficient, coefficient of determination (R2), and predictive relevance (Q2).

4. RESULTS ANALYSIS

The Smart PLS software was employed to analyse the PLS-SEM outcomes. Furthermore, the assessment and interpretation of PLS-SEM analysis were conducted in a dual-stage manner, consisting of evaluations of the measurement and structural models. The evaluation of the structural model is predicated on the assessment of the measurement model.

4.1 Assessment of the measurement model

To begin with, the reliability of the indicators is deemed satisfactory as each indicator loading surpasses the predetermined threshold of 0.70. Furthermore, the model demonstrates internal consistency and reliability by virtue of the composite reliabilities and Cronbach's α indices surpassing the preferred threshold of 0.70. Furthermore, the average variance extracted (AVE) is utilised to assess convergent validity. The obtained AVE values exceed 0.50, thereby providing further evidence in favour of the construct measures' convergent validity. In a similar vein, three widely employed techniques were utilised to estimate the discriminant validity: (1) Cross loadings; (2) Fornell and Larcker criteria; and (3) Hetrotrait–Monotrait ratio (HTMT). Discriminant validity is established when cross-loadings indicate that the indicator loadings of one construct are greater than the loadings of other constructs in the same model. The Fornell–Larcker criterion estimates discriminant validity by comparing the square root of values of the average variance extracted (AVE) with the construct's correlation. Every construct's square root of the AVE in this procedure should be higher than its highest correlation with other constructs. Each construct's square root of AVE values is higher than its correlations with other constructs, according to the analysis shown in Table 3. In conclusion, Table 2 displays the HTMT values, which fall below the conservative rule of thumb of 0.85, thereby confirming the model's discriminant validity.

The results presented in Tables 2 and 3 demonstrate that the measurement model was effectively evaluated, as the data exhibited internal consistency and reliability. The structural model was assessed subsequent to the measurement model's effective evaluation.

4.2 Assessment of structural model

The outcomes of the structural model were evaluated based on previous PLS-SEM research. Among the requirements are:

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

- Collinearity
- Path coefficient
- Coefficient of determination (R2)
- Blindfolding and predictive relevance (Q2)
- Model fit

Table 1: Assessment of the measurement model

Construct	**Indicators reliability**	**Internal consistency and reliability**		**Convergent validity**
		α	ρ	AVE
Test criterion	>0.70	>0.70	>0.70	>0.50
Knowledge	0.90	0.79	0.88	0.71
	0.88			
	0.74			
Comprehension	0.89	0.86	0.92	0.79
	0.90			
	0.86			
Application	0.86	0.80	0.88	0.71
	0.89			
	0.78			
Analysis	0.91	0.89	0.93	0.82
	0.91			
	0.89			
Synthesis	0.91	0.90	0.94	0.83
	0.91			
	0.91			
Evaluation	0.92	0.80	0.88	0.72
	0.71			
	0.91			

Table 2: Mean, standard deviations, correlations, and discriminant validity results

Construct	**Mean**	**Standard deviation**	**1**	**2**	**3**	**4**	**5**	**6**
Knowledge	92.96	20.54	0.84	0.75	0.3	0.20	0.15	0.12
Comprehension	86.78	27.54	0.78	0.89	0.61	0.29	0.22	0.17
Application	51.07	40.08	0.35	0.51	0.85	0.74	0.60	0.44
Analysis	19.26	34.09	0.18	0.26	0.62	0.90	0.79	0.73
Synthesis	12.63	29.20	0.14	0.2	0.51	0.89	0.91	0.83
Evaluation	5.29	18.46	0.10	0.14	0.35	0.61	0.71	0.85

To calculate collinearity, the tolerance value and variance inflation factor (VIF), two independent variables, must be estimated. As a general rule, tolerance should be greater than 0.2 and VIF should be less than 5. The absence of collinearity in this investigation is demonstrated by the tolerance values exceeding 0.2 and the VIF values falling

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

below 5. In addition, the coefficient of determination (R2) is a commonly employed metric for assessing the predictive capability of the structural model. The R2 values fall within the interval of 0 to 1, with 1 denoting absolute predictive accuracy and 0.75, 0.5, and 0.25 representing substantial, moderate, and feeble predictive accuracy, respectively, according to a rule of thumb. The R2 values for Comprehension (0.62), Application (0.26), Analysis (0.39), Synthesis (0.79), and Evaluation (0.50) are all within the acceptable range, according to this study.

In a similar fashion, predictive relevance Q2 represents the predictive power beyond the sample. A portion of the data matrix is omitted during the computation of the Q2 value using a blindfolding technique; this is done in order to evaluate the model's parameters and make predictions on the omitted data matrix using measures that have been calculated beforehand. Additionally, a reduced discrepancy between the predicted and original values results in a greater Q2. The criterion for determining the predictive accuracy of a dependent variable is a Q2 value greater than zero. The predictive relevance of the model is supported by the fact that all Q2 values for dependent variables in this study are greater than zero: 0.44 for Comprehension, 0.17 for Application, 0.30 for Analysis, 0.61 for Synthesis, and 0.33 for Evaluation.

In addition, the path coefficients represent the outcomes obtained by applying the PLS-SEM algorithm to assess the hypothesised relationships among the constructs in the structural model. Furthermore, the path coefficient values are standardised values ranging from -1 to +1. Coefficients estimated near +1 indicate a robust positive relationship, while those near -1 indicate robust negative relationships and are deemed statistically significant. On the contrary, coefficients calculated in proximity to zero indicate a tenuous correlation and are, therefore, lacking in statistical significance. Additionally, the bootstrap confidence interval is a practical method that generates p and t values for determining whether a path coefficient deviates significantly from zero. In order to ascertain the significance level of the path coefficients and weights, 213 bootstrap cases and 5,000 subsamples are utilised in the bootstrapping procedure. The path coefficients, t values, significance level, and p values, as well as the 95% bias corrected and accelerated bootstrap confidence intervals, are displayed in Figure 3 and Table 4.

The correlation is deemed significant at the.001 level (two-tailed test), as indicated in Table 4, due to the t values exceeding 3.29 ($t > 3.29$) with a $p <.001$. In addition, the results of the path coefficient analysis and level of significance suggest that comprehension has been positively impacted by knowledge (HI: $\beta =.78$, $t = 13.47$, $p <.001$); therefore, H1 is accepted. In a similar vein, there is evidence that comprehension positively affects application (H2: $\beta =.59$, $t = 9.08$, $p <.001$), providing support for H2. H3 is supported by the finding that application had a positive effect on analysis ($\beta =.66$, $t = 17.92$, $p <.001$). Following this, the results of the evaluation of synthesis (H5: $\beta =.78$, $t = 5.79$, $p <.001$) indicate that synthesis had a positive impact on analysis (H4: $\beta =.93$, $t = 25.66$, $p <.001$), which provides support for H4. Nevertheless, two additional routes demonstrate statistical significance: application to synthesis and comprehension of the analysis ($\beta = -.09$, $t = 3.53$, $p <.001$). It is noteworthy to mention that two nonhypothesized trajectories exhibit significance; however, the significance is compromised in a negative way (sign changed), as illustrated in Table 3.

It is statistically demonstrated through the aforementioned five hypotheses that the only optimal learning path for students is knowledge → comprehension → application → analysis → synthesis → evaluation, as demonstrated and empirically validated in the preceding discussion. The remaining paths illustrated in Figure 3, denoted by dashed lines, do not exhibit statistical significance. Model fit constitutes the ultimate criterion by which the structural model is assessed. It represents the strength of the relationship between the independent and dependent variables in the model. The evaluation of model fit is constructed using the estimated values of the parameters and the entire model. The conservative threshold value of less than 0.08 for the standardised root mean square residual (SRMR), which indicates the quality of model fit, is utilised to assess the overall model fit. This research identified an SRMR value of 0.06, which is below the critical value, thus signifying a satisfactory level of model fit.

As demonstrated by the analysis of the results, Bloom's taxonomy is an effective method for evaluating and learning computer programming. Developing proficiency in computer programming is an area where regrettably the majority of computer science students fall short. As opposed to being a form of magic that can be mastered through the use of techniques and shortcuts, learning computer programming is a competency that is cultivated through consistent practice. Following the cognitive levels delineated in Bloom's hierarchy, Bloom's taxonomy is thus a practical instrument for progressively augmenting programming proficiency. The Bloom's taxonomy framework comprises six

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

cognitive levels that consistently direct students towards skill development, culminating in the expert level. In accordance with the taxonomy concern, it provides a foundation for students to practise basic problems, which will inspire them to advance to more complex programming when they observe the results of their initial programme.

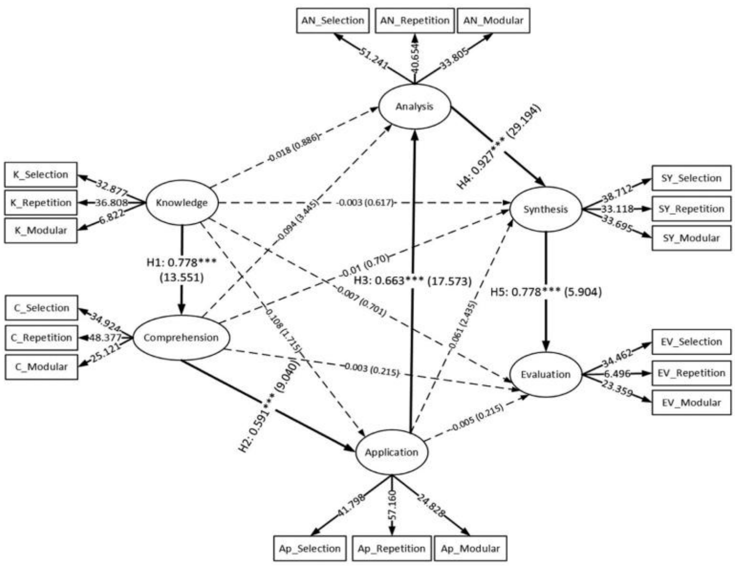

Figure 3: Structural model (path coefficients). Bold lines show the strongly significant values, whereas the dashed lines show nonsignificant values

Table 3: Significant testing results of the structural model path coefficients

Structural path	Path coefficient (β)	t value	Significance level	p value	95% BCa confidence interval	Conclusion
Nonhypothesized relationship						
Knowledge → Application	−.11	1.72		.08	(−0.25, −0.01)	
Knowledge → Analysis	.02	0.92	ns	.36	(−0.02, 0.06)	
Knowledge → Synthesis	.00	0.63	ns	.53	(−0.01, 0.01)	
Knowledge → Evaluation	.00	0.74	ns	.46	(0.00, 0.02)	
Comprehension → Analysis	−.09	3.53		.00	(−0.15, −0.04)	Not supported (−sign)
Comprehension → Synthesis	−.01	0.70	ns	.48	(−0.03, 0.01)	
Comprehension → Evaluation	.00	0.54	ns	.59	(−0.01, 0.01)	
Application → Synthesis	−.06	2.43		.01	(−0.11, −0.01)	Not supported (−sign)

Application → Evaluation	.00	0.22	ns	.82	(−0.05, 0.04)	
Analysis → Evaluation	−.08	0.66	ns	.51	(−0.30, 0.17)	
Hypothesized relationship						
Knowledge → Comprehension	0.78	13.47		.00	(0.65, 0.88)	H1 Supported
Comprehension → Application	0.59	9.08		.00	(0.48, 0.73)	H2 Supported
Application → Analysis	0.66	17.92		.00	(0.59, 0.73)	H3 Supported
Analysis → Synthesis	0.93	25.66		.00	(0.84, 0.99)	H4 Supported
Synthesis → Evaluation	0.78	5.79		.00	(0.48, 1.02)	H5 Supported

Expecting students at the introductory level to generate a comprehensive programme is not feasible. Consequently, assessments should be administered on lower levels, as their progress is insufficient to advance to the third level of Bloom's hierarchy at this stage. Students demonstrate proficiency in the initial two cognitive levels; however, their competence levels diminish as the intricacy of the assessment criteria escalates. Similarly, the high mean values in the first two levels, as shown in Table 3, indicate that the majority of students have successfully completed those levels. Additionally, approximately 50% of the sample did not succeed in attaining the application level. Furthermore, there is a negative correlation between the performance of pupils and the advancement of assessment criteria at the corresponding cognitive level.According to empirical evidence, Bloom's taxonomy is a valuable instrument for evaluating and instructing students in programming. PLS–SEM was used to analyse data collected from 213 students enrolled in an introductory Java programming course. The results indicate that the optimal programming learning sequence is as follows: knowledge → comprehension → application → analysis → synthesis → evaluation, as illustrated by the dark lines in Figure 3. Any alternative routes delineated with indistinct lines are notably unsuitable for the purpose of computer programming education. The utilisation of PLS-SEM to test the proposed hypothesis stated that Bloom's taxonomy is an effective instrument for educating and evaluating computer programming.

5. CONCLUSION

In summary, this investigation provides evidence for the efficacy of combining data mining methods with Bloom's Taxonomy in order to forecast student achievement and provide insights for educational policy-making. Through the integration of sophisticated data mining algorithms and Bloom's Taxonomy, which serves as a systematic framework to comprehend cognitive processes, educators are able to acquire invaluable insights pertaining to student learning outcomes. By utilising the predictive models produced by this fusion, educators are able to better forecast the performance of individual students, thereby facilitating the implementation of focused interventions and individualised instructional strategies. Moreover, by applying the insights gained from this framework to curriculum development and instructional design, the quality of education as a whole can be ultimately improved. Developing further, the amalgamation of data mining and Bloom's Taxonomy exhibits considerable potential in enhancing pedagogical methodologies and fostering academic achievement among pupils.

REFERENCES

[1] Zareen, A., et al. (2023). Enhancing student performance prediction via educational data mining on academic data. *Informatics in Education*. https://doi.org/10.15388/infedu.2024.04
[2] Areej, F., et al. (2023). Early detection of student degree-level academic performance using educational data mining. *PeerJ Computer Science*, *9*, e1294. https://doi.org/10.7717/peerj-cs.1294

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

[3] Nehal, E., et al. (2022). Predicting student performance using educational data mining. *2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS)* (pp. 1–7). IEEE. https://doi.org/10.1109/SNAMS58071.2022.10062500

[4] Jayasree, R., & Sheela Selvakumari, N. A. (2022). A recent exploration on student performance analysis using educational data mining methods. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, *11*(9), 6–10. https://doi.org/10.35940/ijitee.g9188.0811922

[5] Banda, S., et al. (2023). Application of Bloom's taxonomy in categorization of cognitive process development in colleges. *Journal of Education and Practice*, *14*(4), 6–13.

[6] Banujan, K., et al. (2023). Revolutionising educational assessment: Automated question classification using Bloom's taxonomy and deep learning techniques—A case study on undergraduate examination questions. *International Journal of Education & Development Using Information & Communication Technology*, *19*(3).

[7] Leon, M. (2023). Educational data mining: A predictive model to reduce student dropout. In *Proceedings* (pp. 713–721). Springer. https://doi.org/10.1007/978-981-19-9331-2_61

[8] Dhankhar, A., & Solanki, K. (2022). Educational data mining tools and techniques used for prediction of student's performance: A study. *2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICRITO56286.2022.9965023

[9] Vaidehi, B., & Arunesh, K. (2022). A comprehensive assessment on predictive analytics using data mining techniques. *2022 6th International Conference on Electronics, Communication and Aerospace Technology (ICESCA)* (pp. 709–715). IEEE. https://doi.org/10.1109/ICESC54411.2022.9885580

[10] Safira, N., Safitri, H., Setiadi, D. R. I. M., & Suryani, E. (2022). Educational data mining using cluster analysis methods and decision trees based on log mining. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, *6*(3), 448–456. https://doi.org/10.29207/resti.v6i3.3935

[11] Dutta, A., et al. (2023). Cognitive evaluation of examinees by dynamic question set generation based on Bloom's taxonomy. *IETE Journal of Research*, 1–13.

[12] Cheng, Y., & Chen, Y. (2021). A cognitive level evaluation method based on a deep neural network for online learning: From a Bloom's taxonomy of cognition objectives perspective. *Frontiers in Psychology*, *12*, 661235.

[13] Omer, U., Muhammad, S. F., & Adnan, A. (2020). Cognitive learning analytics using assessment data and concept map: A framework-based approach for sustainability of programming courses. *Sustainability*, *12*(17), 6990.

[14] Makhlouf, K., et al. (2020). Exam questions classification based on Bloom's taxonomy: Approaches and techniques. *2020 2nd International Conference on Computer and Information Sciences (ICCIS)* (pp. 1–6). IEEE.

[15] Ullah, Z., et al. (2019). A rule-based method for cognitive competency assessment in computer programming using Bloom's taxonomy. *IEEE Access*, *7*, 64663–64675.

[16] Uma, D., et al. (2017). Analysis on cognitive thinking of an assessment system using revised Bloom's taxonomy. *2017 5th IEEE International Conference on MOOCs, Innovation and Technology in Education (MITE)* (pp. 152–159). IEEE.

[17] Parkavi, A., & Lakshmi, K. (2017). Predicting the course knowledge level of students using data mining techniques. *2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM)*. IEEE.