A Comparative Analysis of Machine Learning Algorithms for Assessing Engineering Students' Problem-Solving

Mr. Rifatul Islam Himel, Utah State University

Rifatul Islam Himel is a doctoral candidate in Engineering Education at Utah State University (USU), Logan, Utah, US. His doctoral research focuses on understanding human cognition in engineering problem-solving. He holds a Bachelor's degree in Electrical and Electronics Engineering and a Master's degree in Electronics and Telecommunication Engineering. With over six years of professional experience in the Information Technology industry and a strong background in teaching and research, Rifatul brings a unique interdisciplinary perspective to his work. His research interests include Human and Machine Cognition, Applied Artificial Intelligence in Engineering Education, Information Processing, and the study of Consciousness.

Dr. Oenardi Lawanto, Utah State University

Dr. Oenardi Lawanto is a professor in the Department of Engineering Education at Utah State University, USA. He received his B.S.E.E. from Iowa State University, his M.S.E.E. from the University of Dayton, and his Ph.D. from the University of Illinois at Urbana-Champaign. Dr. Lawanto has a combination of expertise in engineering and education and has more than 30 and 14 years of experience teaching engineering and cognitive-related topics courses for his doctoral students, respectively. He also has extensive experience in working collaboratively with several universities in Asia, the World Bank Institute, and USAID to design and conduct workshops promoting active-learning and life-long learning that is sustainable and scalable. Dr. Lawanto's research interests include cognition, learning, and instruction, and online learning.

Dr. Yashin Brijmohan, Utah State University

Yashin Brijmohan is a registered professional engineer who is currently appointed as Chairman of Engineering Education Standing Technical Committee of the Federation of African Engineering Organizations, Executive committee member of the Commonwealth Engineers Council, Board Member of the UNESCO International Centre for Engineering Education, and Co-Chair of the Africa Asia Pacific Engineering Council.

He was the founding Executve Dean of Business, Engineering and Technology at Monash South Africa, former Vice President of the World Federation of Engineering Organizations, and led several committees in the engineering profession.

Yashin has both leadership and specialist experience within the engineering power industry and education sectors and is known for his thought leadership in capacity building and engineering education.

Dr. Wade H Goodridge, Utah State University

Wade Goodridge is a tenured Associate Professor in the Department of Engineering Education at Utah State University. His research often has a focus on what drives us to learn, how we learn and how well we do it, and what can improve that learning. This work is often directed with a spatial thinking emphasis. Dr. Goodridge has investigated different learning modalities (fully online, hybrid, blended courses), cognitive and metacognitive learning processes (self-regulation, engineering problems solving, neural efficiency in engineering problem solving) learning motivators (engineering self-efficacy), and effective and personalized engineering education interventions (tactile engineering curriculum, physical teaching models, XR environments and AI informed systems of delivering and monitoring engineering learning). Dr. Goodridge has created a valid and reliable instrument to assess spatial ability in blind and low vision (BLV) learners thus opening new areas of research centered on tactile spatial ability. He has created tactile and visual learning tools and has offered professional development experiences on how to improve the education of tomorrows engineers.

Dr. Zain ul Abideen, Utah State University

Zain ul Abideen holds a Ph.D. in Engineering Education from Utah State University, where he also served as a Graduate Researcher. He is a computer engineer and earned his bachelor's degree in computer engineering and master's degree in engineering, and brings over a decade of experience in teaching and research with undergraduate engineering students. He actively contributes to several technical committees and serves as a reviewer for leading journals and international conferences. His doctoral research focused on cognitive engagement, AI in education and engineering design, self-regulated learning, and mobile learning. His work seeks to deepen our understanding of how these elements influence learning and success in engineering education. With a strong commitment to educational advancement and a robust academic background, Zain is recognized as a dedicated and emerging scholar in the field of engineering education.

Dr. Talha Naqash, Dickinson State University

Dr. Talha Naqash completed his doctoral studies in Engineering Education at Utah State University. With a diverse academic foundation, he also holds a Master's degree in Telecommunication and Networking. He is currently serving as a faculty member at Dickinson State University in the School of Applied Sciences. Dr. Naqash's research focuses on cognitive engagement analysis, assessment methods in engineering education, and the role of facial expressions and emotions in the learning process. His scholarly contributions include numerous publications and presentations at prestigious venues such as IEEE, ASEE conferences, and leading Wiley and Springer journals. He is an active member of several technical committees and serves as a reviewer for reputable journals and international conferences, including the ASEE and IEEE Transactions on Education. Dr. Naqash's dedication to advancing educational research, combined with his broad academic and professional experience, underscores his commitment to fostering innovation and excellence in engineering education.

Dr. Assad Iqbal, Ohio State University, Columbus Ohio

Assad Iqbal (Sid) is a Senior Lecturer at the Department of Engineering Education in the College of Engineering, the Ohio State University. Assad has been working with the National Science Foundation (NSF) funded national collaboration research project, "Engineering For Us All (e4usa) since October 2022. The core mission of this project is to increase student and teacher access to engineering with a focus on reaching traditionally underrepresented populations in this field.

With a BS in Computer Information Systems Engineering, a Master of Science in Engineering Management, and a Ph.D. in Engineering Education, Assad's long-term research goal is to converge his varied research experiences and endeavors to explore the possibilities of an inclusive instructional design to promote self-directed, self-regulated, life-long learning among pre-college and undergraduate engineering students.

A Comparative Analysis of Machine Learning Algorithms for Assessing Engineering Students’ Problem-Solving

Abstract

Assessing undergraduate and graduate students’ problem-solving is essential for engineering education. Existing problem-solving systems focus on the correctness of the final answers rather than reasoning or processes used to get there; the use of Artificial Intelligence (AI) as a tool to help analyze problem solving processes poses a new, interesting, and under researched gap. AI may be a beneficial tool capable of analyzing the problem-solving process itself and reducing the time required for a student to receive formative feedback. Conventional automated assessment systems integrated into Learning Management Platforms (e.g., Canvas, Moodle, Blackboard Learn) predominantly emphasize the correctness of final answers, often overlooking the underlying reasoning and solution processes. This oversight can result in inaccurate evaluations, where correct solutions may arise from flawed logic, or well-structured approaches may yield minor computational errors.

To investigate this area deeper, a hybrid human-machine dual assessment solution for LMS platforms is proposed in this study and a comparative analysis of supervised machine learning algorithms (Logistic Regression, Naïve Bayes, K-Nearest Neighbors, Decision Tree, Random Forest, etc.) are performed to predict student performance based on their problem-solving processes in the context of electrical circuit.

A dataset comprising 363 solution events was collected from the course *Fundamental Electronics for Engineers* at a mid-sized land-grant university in the Western United States. Since the dataset showed some imbalance across problem types and performance categories, an over-sampling method (SMOTEN) was used to balance the data. Each solution was qualitatively evaluated by two independent researchers using five process-oriented constructs adapted from Docktor et al.'s (2016) rubric: Useful Description, Engineering Approach, Specific Application of Engineering Principles, Mathematical Procedures, and Logical Progression are the construct involved.

Solutions were classified into three performance categories—Correct, Partially Correct, and Incorrect—based on expert comparisons. The machine learning (ML) models were trained and evaluated using key performance metrics: accuracy, precision, recall, F1-score, specificity, AUROC, mean cross validation scores, etc. The findings provide evidence of AI utility in assessing cognitive processes associated with engineering problem-solving. This study also has several implications across STEM disciplines, workforce management and development.

Keywords: Machine Learning, Problem-Solving, Engineering Education, Assessment Technology, Learning Management System

1. Introduction

Engineering problem-solving activities involve complex cognitive processing associated with knowledge articulation, management, implementation, meaning-making, planning, multimodal information processing (i.e., drawing, linguistic expression, and mathematical symbols), cognitive engagement, goal-orientations, as well as abstract reasoning [1] – [8] . The evolution of complex cognitive processes does not occur instantaneously; rather, cognitive development emerges through interaction among social, cultural, and environmental factors [9] – [11]. These factors also influence cognitive and mental model evolutions in engineering educational settings where learning, reasoning, problem-solving, and decision-making are paramount. Multiple studies have found that existing cognitive or mental models have an impact on how humans perceive, and approach problem-solving tasks, control cognitive and metacognitive processes, practice self-regulation, and generate solutions [2], [12] – [15].

In engineering education, problem-solving, and cognitive processes associated with procedural reasoning are the central learning outcomes. The assessment practices, therefore, must capture not only the correctness of the solution but also the quality of students' cognitive processes associated with the solution. Traditional automated scoring tools on learning management system (LMS) platforms only score solutions based on the final response. They are unable to assess the quality of cognitive processes associated with problem-solving. This results in a major flaw in the assessment process. While showing evaluation, the system sorely lacks the ability to support engineering students in cognitive and metacognitive processes associated with problem-solving and this can mean we do not realize the full advantage of process feedback that results in a solid foundation of problem solving [16]. Lawanto and colleagues found students to be overconfident (also known as metacognitive illusions [17]) in evaluating their solutions [18], foreshadowing that engineering students may not always possess the ability to self-evaluate their cognitive and metacognitive processes during problem-solving [19]. These metacognitive illusions necessitate interventions from human assessors (instructors or mentors) who possess domain expertise. This lack of cognitive and metacognitive skill development can be further exacerbated through the prevalent usage of Generative AI tools (such as ChatGPT, Gemini, Claude, etc.) in engineering courses [15] [16]. The cognitive sophistication or processes associated with these problem-solving activities, within a controlled academic environment, is the core foundation needed by students for tackling real-world, open-ended, and ill-structured engineering challenges and societal and scientific advancements. Therefore, engineering educators and their evaluation resources should focus on the quality of cognitive and metacognitive processes associated with problem-solving rather than the correctness of the solutions only. As Albert Eienstein famously said, "*Education is not the learning the facts, but it is training the mind to think*" [20].

This study is a part of a larger investigation to understand and assess cognitive processes associated with problem-solving. One often overlooked approach by LMS platform developers is their lack of use of a problem-solving assessment rubric (i.e. Polya's [21], Docktor et. al.'s [22], AACU's [23]) for evaluating solutions of engineering problems. These problem-solving assessment rubrics frame the dimensions of success in problem solving with qualitative distinctions of cognitive processes (the indicatives of cognitive sophistication) easily visible for the interpretation of a good solution.

The researchers of this study wove an understanding of human cognition, problem-solving, and artificial intelligence (AI) in a pre-automated or Human-Machine Dual Problem-Solving Assessment (HHMDPSA) solution for the CANVAS© LMS platform. For automating the whole process, the researchers may explore unsupervised, or natural language processing algorithms. However, the correct algorithms for automating the whole processes are still under investigation. The HHMDPSA solution of this study is expected to function independent of problem-solving rubric. This HHMDPSA solution is also expected to reduce the human effort in scoring problem solutions based on cognitive sophistication utilizations.

2. The Study

This study was designed based on the students' solutions submitted on an LMS platform for homework problems in a general foundational engineering course, *Fundamental Electronics for Engineers*. The homework problem-statements of this course are domain-specific, academic, and situated in constrained and finite rules [24]. However, the solutions to the problems of this course may vary in terms of cognitive processing quality. The course contains nearly 100 homework problems in total. Four of these problems (varied in terms of cognitive efforts) are carefully chosen for this study. Students who attended this course are from multiple engineering and computation disciplines (Mechanical, Civil, Biological, etc.), ensuring a broad diversity in likely cognitive processes engrained in the problem-solving data. This attribute makes this study more enticing for studies regarding the qualitative characterization of a score from the evaluation rubric used and allows an investigation into, understanding cognitive sophistication and proposing a Machine Learning (ML) or AI-based HHMDPSA solution for the LMS platform.

The proposed HHMDPSA solution is illustrated in **Figure 1**. This HHMDPSA solution for the LMS platform consists of human judgment and machine analysis sections. This hybrid assessment solution is designed to combine both qualitative human judgments, and mechanistic AI algorithms in indicating cognitive sophistication evident in the problem-solution. HHMDPSA is a flexible solution and can facilitate absolute and relative human judgment of cognitive sophistication or processes based on the users' choice of assessment and of the data. This HHMDPSA solution is expected to be easily synched with the current LMS platform with little modification in the system architecture. **Figure 1** shows a representation of the HHMDPSA solution architecture. As shown in **Figure 1**, the setup requires the CANVAS© LMS platform, where the proposed HHMDPSA solution needs to be integrated. The development of this HHMDPSA solution requires the submission page, rubric selection options, and scoring pages containing the submitted solutions. In the machine section, the HHMDPSA solution requires a central physical database that links to the training, development and testing (TDT) segment, a running classifier in the production environment that can take data from the central database, and the students' gradebook page. Depending on the HHMDPSA system requirements by the users, an additional physical database may or may not be needed in the TDT segment. All the pages and classifiers needed to be coded and maintained as similar as any LMS platform.

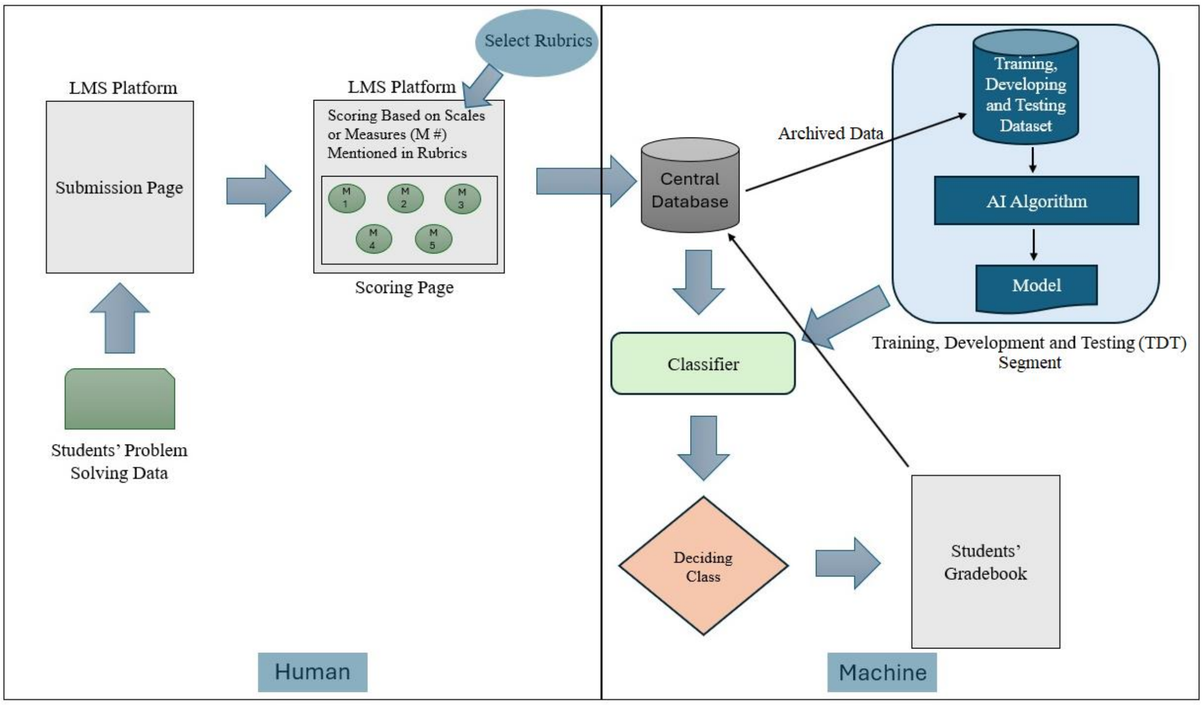

Figure 1: Proposed Hybrid Human-Machine Dual Problem-Solving Assessment (HHMDPSA) for LMS Platform

The human judgment section of HHMDPSA assessment system has two sections. One is for the students, and another is for the instructors. In the students' section, the LMS provides facilities for students to submit their solutions to the given problems. The students are able to upload their files in different formats (pdf, .docx, .doc, .jpg, .pages). After uploading, these files will be segmented into their individual pages for instructors to proceed with assessment activities.

The instructors' interface has to be modified for developing the HHMDPSA solution. First, the LMS platform requires instructors to choose a problem-solving rubric, such as, Polya's, Docktor's, AACU's etc. Based on the selection of the problem-solving rubric, the platform provides the scoring page, containing (1) the students' information (name and identification number), (2) individual solution page to the given problem in their course, and (3) some blank boxes (i.e., M1 to M5 in green oval in Figure 1) for instructors to score based on measures and scales mentioned in the Docktor et al.'s (2016) problem-solving rubric. The letter 'M' in the green oval in Figure 1 represents the selected rubric measure and 1 to 5 are the serial numbers of the measures. The measures are indicative of cognitive processes applied in the solutions. For instance, M1 is the 'useful description' measures, and a score of 5 is the appropriate, and the best possible useful description shown in the solution in the form of written text, drawing, and symbols. The instructors then score or rank individual students' solutions. These scores or ranks illustrate how students utilized their cognitive processes associated with the measures mentioned in the rubric. These humans produced problem-solving scores, then feed into the machine analysis section, allowing it to develop a foundation to analyze problem-solving process from that is attuned to human pilot data.

The machine analysis section contains central database, the training, development, and testing (TDT) segment, and the classifiers or production environment. First, the instructor generated problem-solving scores are stored in the central database. This central database is for one or multiple engineering courses associated with problem-solving activities. It is important to note that this HHMDPSA is an engineering domain-agnostic assessment solution, as the problem-solving processes (mentioned in the rubric) can be similar for many engineering courses. Moreover, the domain expertise of any problem-solving activities in any engineering domain is imputed through the human judgment section.

The Training, Development, and Testing (TDT) segment (colored in blue on the right) takes the archived data from the central database to train, develop, and test the AI algorithms, and prepare the final model. This archived data represents the previous recorded scores (cognitive processes associated with problem-solving data) based on the selected problem-solving rubric. First, this data is used for training the AI algorithms, and developing models, as shown in Figure 1. The term 'AI Algorithms' is seen in the diagram instead of 'Machine Learning (ML) Algorithms' in an effort to make this block more generalizable for neural networks (NN) models, multimodal generative AIs or for AI agents. It is recommended to train the AI models with larger relevant datasets to receive better performance as they evaluate solutions. The detailed workflow process of training AI or Machine Learning (ML) models (**Figure 2**) is illustrated and discussed in the data collection and analysis section of this paper. The final trained model works as a second brain in this HHMDPSA solution for assessment. The analysis to identify the best performing ML model is the key discussion of this paper.

After developing the final model, it needs to be set to the production environment, illustrated as Classifier in **Figure 1**. This classifier consumes the data from the central database, containing scores or ranked (ordinal or categorical) information of students' problem-solving activities based on the rubrics. The classifier receives the data and runs the model to classify or decide which problem-solving activities they are in and sends that group score to students' gradebook. This combined HHMDPSA solution can ensure efficient and expedited grading based on students' cognitive efforts rather than just their outcomes. As a result, a focus can develop and remain on the solution process and not just a solution answer and students will not be rewarded for guesses or results may not be treated as effortful and successful solutions.

3. Study Context

The study was conducted at a Carnegie R1 university in the western part of the United States in a sophomore-level engineering class. The problem-solvers mostly belong to similar socio-cultural and academic groups. These problem-solvers have enough experience in abstract and logical reasoning through their K-12 and other STEM-related classes.

The course contained DC and AC networks and digital electronics concepts. These concepts were taught during the class. Moreover, additional learning materials were provided on the LMS platform for the students to understand the concepts. Problem-solvers were required to solve homework assignment problems associated with those concepts in each week. Overall, the problem statements were not conceptually and procedurally unfamiliar to them.

4. Research Question and Objectives

This study is guided by a single research question: Which machine learning algorithm demonstrates superior performance in predicting the engineering students' cognitive sophistication utilized in problem-solving? The major objective of this study is to (1) propose a hybrid Human-Machine Dual Assessment method linked to a LMS platform and used for analyzing and categorizing cognitive sophistication in engineering problem solutions; (2) the effectiveness in implementing ML Algorithms for problem-solving assessment; (3) providing a direction to develop an automated AI-driven problem-solving assessment solution for future learning management systems.

5. Data Collection and Analysis

The data of this research was collected from students' submissions of the solutions to the selected homework assignment problems in Fundamental Electronics for Engineers class. These problem-solving data on LMS platforms contained identifying information and were protected by Family Education Rights and Privacy Act (FERPA). Therefore, the researchers collected the deidentified problem-solving data from the data governing body at the university, while maintaining Institutional Review Board (IRB) and FERPA regulations.

A total of 396 problem-solving data files were collected for analysis and for entry into the HHMDPSA analysis for the LMS platform. Due to the nature of the study, no additional think-a-loud or interview sessions are conducted. After collecting the problem-solving data, the researchers followed step-by-step procedures shown in Figure 2. Each step of the data analysis process is further described in detail following the workflow data analysis diagram.

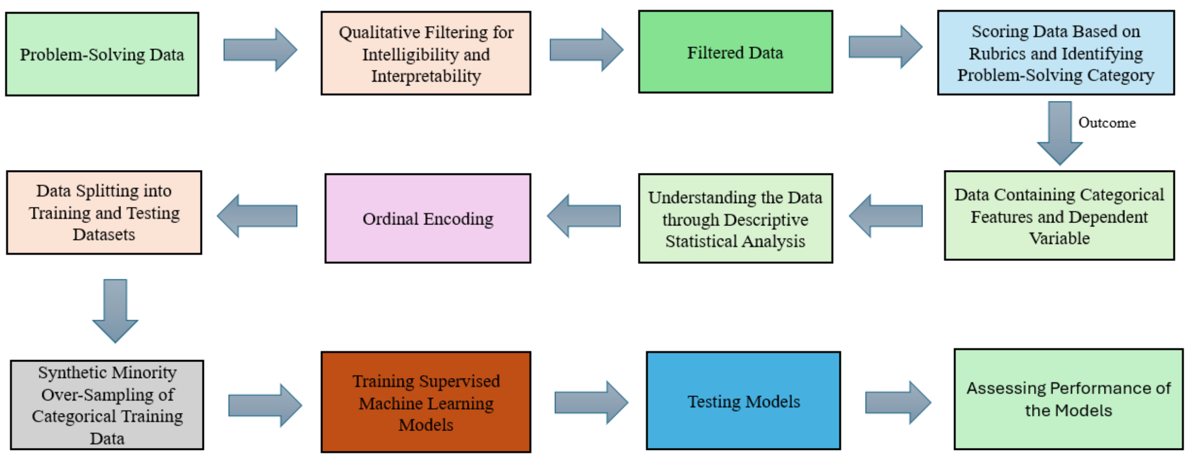

Figure 2: Workflow Diagram of Data Analysis

First, the collected problem-solving data were qualitatively filtered for their intelligibility and interpretability by two independent researchers. They found that 33 problem-solving data were either unintelligibility or non-interpretable (i.e., blank, or griffonage). The researchers, therefore, left with 363 problem-solving filtered data.

Then, two independent researchers, possessing similar expertise in cognition and problem-solving, tagged each problem-solving event with unique identifiers (such as, event001, event002, event003, etc.). Then, they studied and scored individual problem solutions of the qualitatively filtered data based on five measures mentioned in the adapted Docktor et. al.'s rubric [22]. The rubric is adapted by slight modification from the description of 'Physics' to the description 'Engineering'. The details of the rubric are presented in the **Appendix** section. The researchers discussed their scoring process among themselves, refined any disagreements, and reached 100% consensus of their scores [25]. The individual problem-solving scores are stored in a separate Excel© file with unique identifiers. The scoring is done in 0-5 Likert scale, where 0 means no cognitive effort, and 5 is a highly sophisticated cognitive effort associated with problem-solving [22]. The overall qualitative description of the modified rubrics (measures and scales) is provided in the **Appendix** section of this paper. The specific adapted mastery levels of this rubric are useful description, engineering approach, specific application of engineering, mathematical procedures, and logical progression. The description or definition of the measures of the problem-solving rubrics is shown in Table 1.

Table 1: Definition of Modified Measures (or, Independent Variables) of Problem-Solving Rubrics

Measures	**Definition**
Useful Description	Problem-solver's written records of the task analysis, including redrawn diagrams, givens, findings, labeling, notations etc.
Engineering Approach	The written record of broader conceptual understanding of Engineering Fundamentals and their applications in a particular course.
Specific Application of Engineering	Specific Application of Engineering refers to contextual use or the applications of the problem-solving concepts in specific use cases.
Mathematical Procedures	Mathematical procedures are the use of rules and the appropriate mathematical processes or steps to solve a problem.
Logical Progression	Logical Progression is an organized and consistent reasoning pattern in the problem-solving process.

After scoring all problem-solving data, the researchers met, discussed their scoring, and resolved any disagreements. Then, they shared the final scoring with the additional two researchers who then used them for observing and conducting initial analysis on more data. Through this means, consistency in their scoring was assured based on the rubric. Furthermore, the descriptive statistical analysis (**Table 3**) based on the dependent variables (categories or classes) was found to be consistent. The analysis in **Table 3** assures rigor and qualitative precision in the problem solution scoring process. In the next step of data analysis, two researchers reexamined each problem-solving sample based on their analyzed quality. For this research purpose, they defined categories, or classes as k, when the number of classes is three (i.e. correct, partially correct, and incorrect) mentioned in **Table 2** and investigated and categorized each problem based on these categories. These categories are not absolute and can be expanded or shrunk for future application based on the instructor's choice of how to categorize students for a full-fledged HHMDPSA solution. The categorized problem-solving data, unique identifiers, and their rubric-based cognitive processes scores that were associated with the problem-solving events were also stored in the Excel© file. Through this scoring and problem-solving data categorization process, the researchers created a dataset containing both dependent and independent variables (i.e., Useful Description, Engineering Approach, Specific Application of Engineering Principles, Mathematical Procedures, and Logical Progression) for the next phase of analysis.

Table 2: Definition of Categories or Classes (Dependent Variable)

Category or Class (k = 3)	**Definition**
Correct (~ 100%)	The solution contains 'almost no' to 'very minimal' mistakes.
Partially Correct (~ 50%)	The solution possesses both consistent and inconsistent problem-relevant solution components.
Incorrect (~ 0%)	The solution is totally wrong or has major flaws.

After creating the dataset, the researchers explored the dataset and found that it contained 134 (37%) solutions that were correct, 76 (21%) partially correct solutions, and 153 (42%) incorrect categories, illustrated in Figure 3.

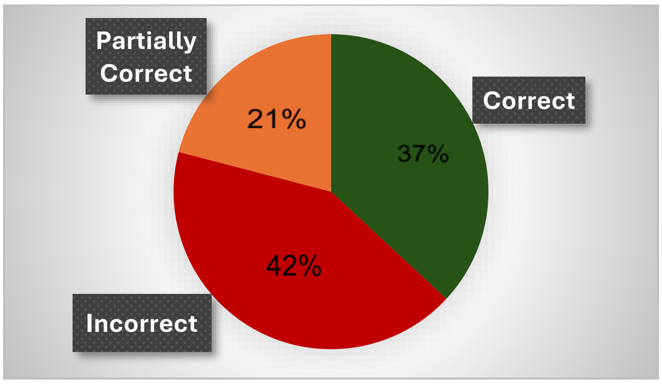

Figure 3: Percentage of Data in Multiple Outcome Classes (Categories)

The researchers, then, performed descriptive analysis (average and standard deviation) for total dataset and for individual classes, such as, correct, partially correct, and incorrect. The average (*M*) of total sample size (n=363) is 13.66 and the standard deviation (*SD*) is 6.94. From the descriptive analysis shown in **Table 3**, it is found that the average (*M*) of correct, partially correct, and incorrect categories is very different. Problem-solving data in correct categories are comparatively superior in all measures (i.e., in describing the task, applying engineering and mathematical concepts, demonstration of better logical thinking) than partially correct and incorrect categories. The significant differences in student solutions are found when we look through a lens of Specific Applications of Engineering (i.e., Mechanical, Electrical, etc.). The average score (*M*) of 'Correct' category in Specific Applications of Engineering (or, contextual application) is 4.51. For 'Partially Correct', the average (*M*) value is 2.49 and 'Incorrect' category has received only 0.94 on average (*M*). This finding confirms the existing literature of the cognitive evolution of experts, which states that the experts tend to have comparatively superior performance at subordinate levels or domain-specific (i.e., Mechanical, Electrical, etc.) contextual tasks [26] [27]. Evidence also shows a significant difference in conceptual understanding between the incorrect and correct categories. This demonstrates the comparative deficiency in general engineering knowledge among incorrect (no part of the solution corrects) categories than correct (almost no or minimal mistakes) ones. Table 3 also reveals another profound finding. For instance, standard deviations (*SD*) of correct categories in three measures, such as, engineering approach, mathematical procedure and logical progression, are much lower (0.83, 0.74 and 0.87 respectively). These lower values indicate that cognitive processes of 'Correct' class are more consistent (closer to mean) in those measures than their counterparts. Overall, **Table 3** provides researchers with fundamental characteristics of the dataset in each class (category). The descriptive analysis also assures that the dataset (scoring and categories) developed by the researchers in this study are reducing bias in their qualitative analysis. Rather the dataset is generated based on strictly followed procedures and rubric adopted for this study.

Table 3: Descriptive Statistics of Problem-Solving Sample

	Total Problem-Solving Sample (n=363, *M*=13.66, *SD*=6.94)					
	Correct (~ 100%)		**Partially Correct (~ 50%)**		**Incorrect (~ 0%)**	
	M	*SD*	*M*	*SD*	*M*	*SD*
Useful Description	2.37	1.56	2.25	1.47	1.88	1.50
Engineering Approach	4.69	0.83	3.00	1.15	1.69	1.14
Specific Application of Engineering	4.51	1.01	2.49	1.16	0.94	1.13
Mathematical Procedure	4.67	0.74	3.25	1.14	2.09	1.22
Logical Progression	4.07	0.87	2.58	1.15	1.28	1.17

n: Sample Size*; M*: Mean or Average; *SD*: Standard Deviation

Based on the dataset generation process discussed above, it is apparent that both independent and dependent variables are not numerical but categorical in nature. Therefore, the researchers cannot send the ordinal (a type of categorical variables) data directly to ML or AI algorithms for comparative analysis as shown in previous studies [28] – [30]. To prepare data for comparative analysis, an ordinal encoding technique was applied during data preprocessing to convert categorical features into numerical ranks while preserving order [28]

– [30]. Then, this converted data was distributed into the system to train and test datasets. For this study, 70% of the data was in the training dataset, and 30% was in the test set. During train-test splitting, stratifying methods are used to preserve category or class proportionality in both the training and test sets. This 70:30 distribution is followed to ensure availability of comparatively enough datasets for testing the algorithms within this research context. Moreover, 70% of total sample size of this study is comparatively smaller for training the ML models. Therefore, a standard oversample technique, Synthetic Minority Over-Sampling for Categorical Data (SMOTEN), was applied to generate similar synthetic data and reduce some imbalance for our study purposes [31]. However, the proposed final HHMDPSA solution used hereafter will likely contain more problem-solving samples and associated human produced judgment scores from one or more engineering courses. Therefore, a fully functional HHMDPSA solution can omit this oversampling procedure in practical settings.

Finally, the synthetically over-sampled training dataset was fitted to the supervised machine learning algorithms or models, such as, Logistic Regression, Naïve Bayes, K-Nearest Neighbors (kNN), Decision Tree, and Random Forest. These algorithms are called supervised models, because they are trained on labeled datasets (Independent variables and their associated dependent variables are seen in the training data). After training these supervised ML models, each model was tested with an unseen or test dataset. Then, the models predicted the category or class of test data. The predicted categories or classes are compared with the actual category identified in the test data. The overall model performance and the performance in classifying individual categories are studied and compared in this paper. For a fully developed HHMDPSA solution, the ML model may have to go through three datasets (training, development, and testing) before sending it to the production environment.

5.1 Performance Evaluation Metrics

In analyzing, evaluating, and comparing the ML model performance, the researchers used several performance metrics in this study. The performance evaluation metrics are: True Positive Rate (TPR) or Recall, True Negative Rate (TNR) or specificity, False Positive Rate (FPR) or Type I Error, False Negative Rate (FNR) or Type II Error, precision (P_k), F1-score ($F1_k$), and classification accuracy (ACC_k) [32] – [34]. All these evaluation metrics in this study are calculated based on the confusion matrix ($C_{i,j}$) generated from the individual machine learning (ML) model [32], [33].

A confusion matrix, $C_{i,j}$, from machine learning models is an N × N matrix, which is used for representing and comparing actual labels with predicted classes (classes, $k \in \{1,2,\ldots,N\}$) [32], [33]. The row, i, of the matrix represents actual (labeled) categories, and the column, j, represents predicted categories of the models. It is important to note that row, i, and column, j, of the confusion matrix, $C_{i,j}$ is always the same (i=j). TPR or Recall is the proportion of actual k solution class that are also predicted as k by an ML model [32] – [34]. TNR or Specificity is the proportion of actual 'not k' solution class that are also classified as 'not k' [32] – [34]. FPR is the proportion of actual 'not k' solution class, but the classifier predicted otherwise, and FNR is the proportion of actual k solution class but incorrectly predicted as 'not k' class [32] – [34]. FPR and FNR are also known as Type I Error, and Type II Error respectively. TPR, TNR, FPR and FNR for each ML model are calculated based on the following generalized equations for k classes (where, i and j are the rows and columns of the confusion matrix respectively) [32] – [34]:

$$TPR_k = \frac{C_{k,k}}{\sum_{j=0}^{N-1} C_{k,j}} \quad (1)$$

$$TNR_k = \frac{\sum_{i=0,i\neq k}^{N-1} \sum_{j=0,j\neq k}^{N-1} C_{i,j}}{\sum_{i=0,i\neq k}^{N-1} \sum_{j=0,j\neq k}^{N-1} C_{i,j} + \sum_{i=0,i\neq k}^{N-1} C_{i,k}} \quad (2)$$

$$FPR_k = 1 - TNR_k \quad (3)$$

$$FNR_k = 1 - TPR_k \quad (4)$$

The same N × N confusion matrix, $C_{i,j}$, are used in implementing precision (P_k), F1-score ($F1_k$), and classification accuracy (ACC_k). These performance metrics are calculated using the following generalized equations [32] – [34]:

$$P_k = \frac{C_{k,k}}{\sum_{i=0}^{N-1} C_{i,k}} \quad (5)$$

$$F1_k = 2.P_k.\frac{TPR_k}{P_k + TPR_k} \quad (6)$$

$$ACC_k = \frac{\sum_{i=0}^{N-1} C_{k,k}}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C_{i,j}} \quad (7)$$

6. Results

The performance metrics for classifying individual categories (i.e., correct, partially correct and incorrect) are primarily compared based the calculation performed using equations **(1) – (4)** for each ML model. After calculating TPR, TNR, FPR and FNR for all models and for all k classes using equations **(1) - (4)**, it is found that Logistic Regression, Naïve Bayes and Random Forest classifiers are performing comparatively better than other algorithms in all performance measures. **Figure 4** illustrates all four measures (TPR, TNR, FPR, and FNR) of the performance metric for five ML algorithms used for this study in line charts. The green line represents the **'Correct'** class, the orange line is **'Partially Correct'**, and the red is **'Incorrect'** class.

From **equations (1)** - **(4)** and **Figure 4** (below), it is apparent that TPR and TNR are mathematically complementary parts of FNR and FPR respectively. Therefore, FNR and FPR will not be discussed in the following few paragraphs to avoid redundancy.

From the TPR graph (Top Left in Figure 4), it is found that Naïve Bayes classifier could categorize 'Correct' solution to be correct more often than other algorithms. Its performance score is about 0.93, which is comparatively better than Logistic Regression (0.9) and Random Forest (0.88). On the other hand, Logistic Regression and Naïve Bayes classifiers correctly identified 'Partially Correct' class effectively than others. These classifiers' TPR score is 0.61. The Decision Tree classifier has the lowest score, only 0.67. Decision Tree, Random

Forrest, and K-nearest Neighbors (kNN) are performing 0.57, 0.48 and 0.4 respectively. For classifying ‘Incorrect’ solution to be ‘Incorrect’, Logistic Regression Algorithm is performing better. The TPR score for Logistic Regression Algorithm is around 0.8.

From the TNR graph (Top right in Figure 4), Logistic Regression and K-nearest Neighbors (kNN) are classifying ‘not Correct’ more often than others. Their TNR score is 0.94. Logistic Regression also performed comparatively better in identifying ‘not Partially Correct’ class as shown its TNR score (0.88). However, Naïve Bayes surpasses other classifiers in finding ‘not Incorrect’ solutions. The TNR score of Naïve Bayes Classifiers is 0.89. K-nearest Neighbors’ (kNN) performance is very low in classifying ‘not Incorrect’ solutions, having a TNR score of only 0.79.

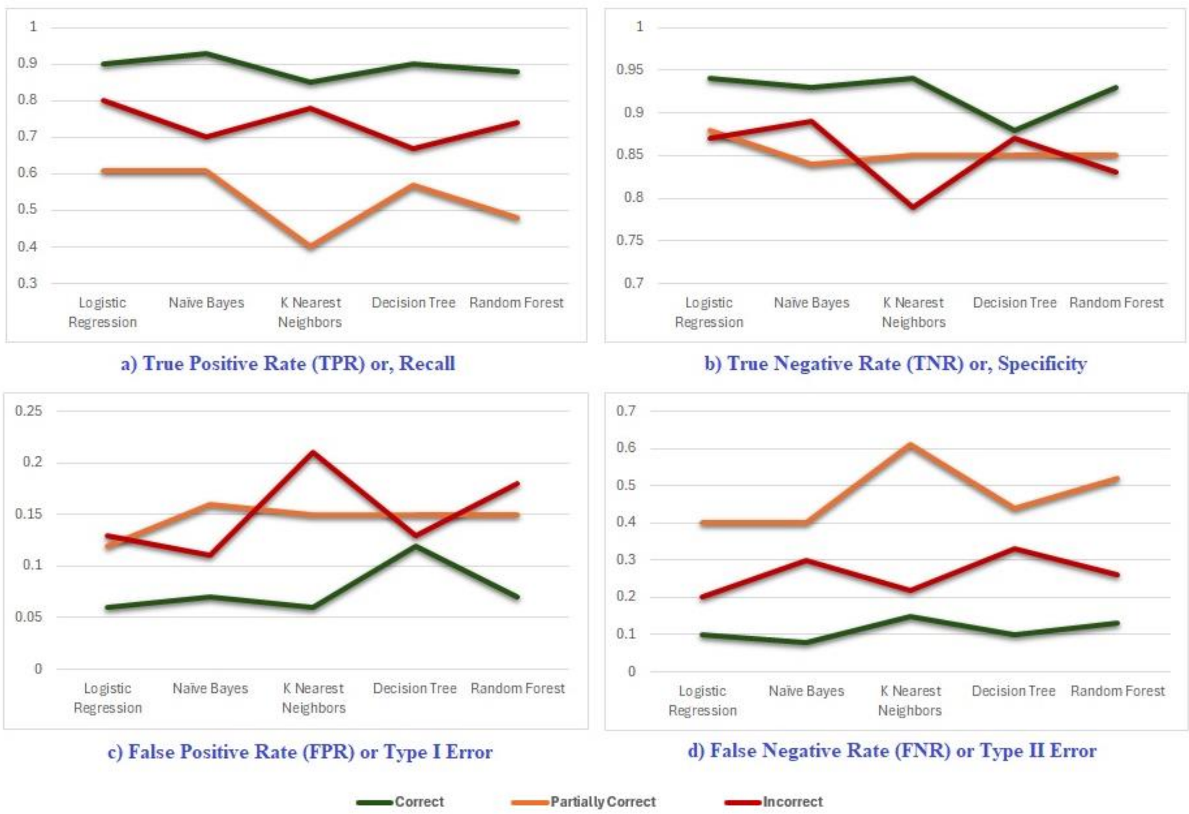

Figure 4: a) TPR/Recall, b) TNR/Specificity, c) FPR/Type I Error, and d)FNR/Type II Error of Machine Learning Models for Individual Category

The three prospective algorithms performed superiorly for classifying individual classes (correct, partially correct, and incorrect) than others, allowing the researchers to conduct further investigation into their application to observe performance consistency in other metrics of this study. Therefore, precision (P_k), F1-score ($F1_k$), and classification accuracy (ACC_k) are calculated for each machine learning model for individual categories.

After calculating precision (P_k), F1-score ($F1_k$), and classification accuracy (ACC_k) values from equations **(5) – (7)** for each ML model for individual classes (k=3), the calculated values of these metrics are plotted in Figure 5. Figure 5(a) represents precision scores for three classes for all models based on the calculation of equation (5). Similarly, 5(b) and 5(c) illustrate F1-Score and Accuracy respectively for three classes (correct, partially correct & incorrect) for all models based on equations (6) and (7). In Figure 5, green represents correct class, orange represents partially correct class and red represents incorrect class. In all cases, the Logistic Regression model performed comparatively better than any other algorithms in precision, F1-score and accuracy. From the following diagram, 5(a), the precision of the Logistic Regression algorithms is found: 0.9 for correct, 0.58 for partially correct and 0.82 is for incorrect classes. The F1-score of this algorithm is 0.9 for correct class and 0.6 and 0.81 for partially correct and incorrect classes respectively. Finally, the accuracy of the Logistic Regression classifier is 0.93 for correct, 0.83 for partially correct and 0.84 is for incorrect classes. In all cases, K-nearest Neighbors (kNN) and the Decision Tree classifier performed comparatively poorly.

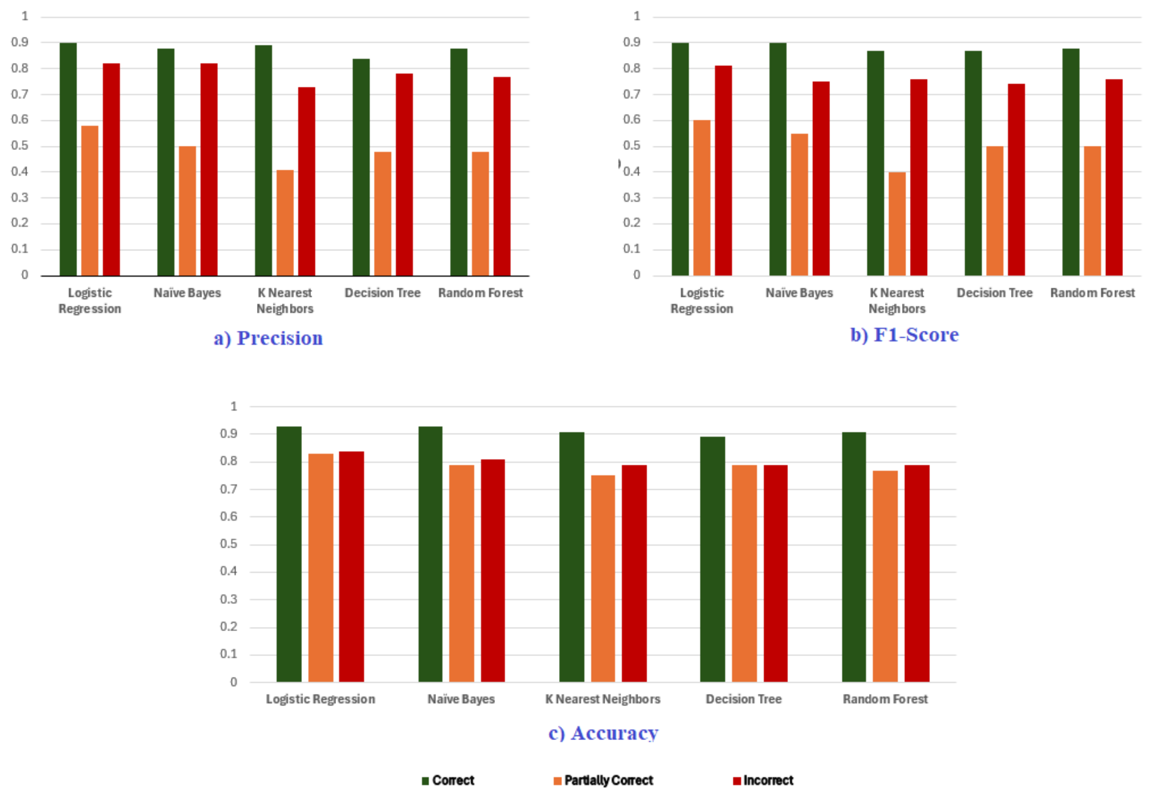

Figure 5: a) Precision, b)F1-Score, and c) Accuracy of Machine Learning Models for Individual (k=3) Classes

After analyzing the individual classes, the researchers investigated overall classifiers' performance based on key performance metrics, such as, accuracy, precision, recall, F1-score, specificity, AUROC or AUC Score, and mean cross validation score (where, k-Fold=5) [32] – [35]. The k-Fold value is selected as 5, because it is the most common k-Fold value according to the literature [35]. The ROC-AUC or AUROC or AUC scores of each model are the computed value of the area under the ROC curve [36]. The AUROC or AUC score provides a summary of the curve information in a single number [36]. The comparative performance analysis of the overall models for all classes (correct, partially correct and incorrect) is shown in Table 4.

In calculating and comparing overall ML model performances, the final calculation incorporated weighted averages to ensure that each class received relative importance. Therefore, this final outcome is different from the above analyses shown in Figure 4 and 5. Based on the overall analysis, the Logistic Regression model demonstrates comparatively superior performance than any other algorithm. The performance metric score of this algorithm is 0.8 in accuracy, precision, recall, and F1-score, 0.9 for specificity, 0.88 for AUROC score, and 0.76 is the mean cross validation score (k-Fold=5). Based on this comparative analysis shown in Table 4, K-nearest Neighbors (kNN) and Decision Tree models are comparatively low performing algorithms. The accuracy, precision, recall and F1-score of Decision Tree is 0.72, whereas the mean cross validation score for k-Fold=5 for Decision Tree classifier is only 0.68.

Table 4: Comparative Performance Analysis of Machine Learning Models for all Classes

Machine Learning/ AI Models	Performance Metrics						
	Accuracy	**Precision**	**Recall**	**F1-score**	**Specificity**	**AUROC, or AUC Score**	**Cross Validation Score, k-Fold=5 (mean)**
Logistic Regression	0.80	0.80	0.80	0.80	0.90	0.88	0.76
Naïve Bayes	0.76	0.78	0.76	0.76	0.89	0.88	0.76
K-Nearest Neighbors	0.72	0.72	0.72	0.72	0.86	0.84	0.74
Decision Tree	0.73	0.74	0.73	0.73	0.87	0.79	0.68
Random Forest	0.74	0.75	0.74	0.75	0.87	0.86	0.73

7. Discussion

This study is a comparative analysis of machine learning (ML) models and an exploration of observing model performance based on the metrics of accuracy, precision, recall, F1-score, etc. These machine learning models are tunable mathematical frameworks (i.e., log-linear, probabilistic, distance-measure, and tree-based) that determine the model parameters during training phase. There is no benchmark set for testing and model performance analysis for individual classes and for overall models in this study context.

From the above analysis, the researchers identified that the Logistic Regression algorithm or log-linear model performed comparatively better in all performance metrics of all problem-solving activities than the others tested. The Logistic Regression algorithm, therefore, is the best suited model for the proposed HHMDPSA solution for the LMS platform. The logistic regression algorithm may perform better in this study, because of its inherent mathematical representation of sigmoid function which typically works better for limited number of ordinal data.

The primary purpose of this study is to demonstrate a viable HHMDPSA solution for assessing engineering students' cognitive sophistication or processes utilized in problem-solving, and finding the best suited ML algorithms used to do this based on categorical data. To the authors' knowledge, this is one of only a few or even only attempts to integrate both qualitative human judgment and mechanistic machine learning algorithms in categorizing and assessing engineering students. Although this study is conducted based on sophomore level problem-solving data, the HHMDPSA solution can also be applied across all problem-based engineering courses.

Aside from assessing the sophistication of students' cognitive processes utilized in problem-solving activities, this HHMDPSA solution can also assist engineering educators in categorizing students into certain classes and designing appropriate instructional interventions in the form of course materials, instructions, or problem-solving assignments. For instance, the instructors can identify that students in 'incorrect' (where they did not have any part of the solution correct) categories may not spend appropriate time in understanding the problem-solving task or may have trouble with specific or contextual applications of engineering concepts [6], [37] – [39]. Therefore, some appropriate interventions for those groups or categories of students can be designed during tutoring and assisting sessions. By providing the practice materials in the weak areas, the instructors can design more group or cohort-oriented learning experience. Furthermore, the students in 'Correct' (minimal error) categories can be asked to submit more challenging or open-ended problem-solving tasks with different guidance or support. These pedagogical practices can enhance and personalize learning and problem-solving skills for engineering students.

8. Implications

This study has many implications in academics, industries, and in any training and assessment activities in education, finance, marketing or in other areas. This HHMDPSA solution can be used to understand and analyze training outcomes, performance, or social activities, such as leadership, creativity, critical thinking, role as team members, etc. The above analyses can be implemented to assess engineering students' solution processes and understand the specific learning needs of the students in both online and offline settings. Moreover, the transition from offline to online AI-integrated cognitive achievement based assessment tool (i.e., HHMDPSA) could enable the educators and students to cope up with both online and offline evaluation practices, resulting in reducing some assessment relevant anxiety and mental health issues identified by Lalin and colleagues (2024) [40].

The same inquiry can be performed for employees in the corporate settings. In working with such Human-Machine integrated solutions, the assessors (such as educators, cognitive engagement, human resources, management, trainers, etc.) need to design or select appropriate rubric and scales for their own needs of the assessment. The effectiveness of HHMDPSA solution depends on AI models and associated preferred data to train the models.

As stated earlier, HHMDPSA solution is very flexible. This categorical data driven AI solution can also integrate other data types, such as text data for training or exam experiences, images, videos, etc., into the system. The multimodal AI models can also be utilized to make better judgments and decisions.

9. Limitations and Future Direction

While this study demonstrates a method to utilize AI algorithms to study cognitive processing associated with problem-solving, this study, like many others, has its own limitations. First, the larger data set (containing more dimensions and observations) may provide a more accurate, robust and generalizable model. In that case, deep learning or neural network or complex non-linear models (e.g., XGBoost, kernel-based SVMs, etc.) could perform better than traditional machine or statistical learning algorithms. Second, the machine learning models implemented in this study were not tweaked or tuned, which is outside the scope of this study. The tweaking or tuning of the model can improve the overall performance of the model. These ML models could also be tuned using multiple encoding techniques, oversampling methods, hyperparameter tuning techniques, etc. Third, the HHMDPSA solution is not a fully automated system. For fully automated models, the researchers need to find algorithms that can study problem-solving data like text data in NLP models. The researchers will work on to overcome these three limitations in future studies by collecting more data (such as categorical, text, images, etc.), in more venues (such as other courses and at other age levels) studying and tuning more sophisticated AI models and finding appropriate algorithms to study students' solution submissions.

References

[1] K. VanLehn, “Problem solving and cognitive skill acquisition,” in Foundations of Cognitive Science, M. I. Posner, Ed., 1st ed. Cambridge, MA, USA: *MIT Press*, 1993, pp. 527–579.

[2] W. Rumite, P. Purwanto, I. N. Parta, and S. Rahardjo, “Unpacking mental models, strategies, and schemas pre-service mathematics teacher in solving maximum rectangular

areas," *Eurasia J. Math. Sci. Technol. Educ.,* vol. 19, no. 8, p. em2309, 2023, doi: 10.29333/ejmste/13430.

[3] Y. Wang and V. Chiew, "On the cognitive process of human problem solving," *Cogn. Syst. Res.,* vol. 11, no. 1, pp. 81–92, 2010, doi: 10.1016/j.cogsys.2008.08.003.

[4] J. G. Greeno, A. M. Collins, and L. B. Resnick, "Cognition and learning," in *Handbook of Educational Psychology*, D. C. Berliner and R. C. Calfee, Eds. New York, NY, USA: Macmillan; Prentice Hall International, 1996, pp. 15–46.

[5] Z. Abideen, S. Jabeen, O. Lawanto, A. Minichiello, and A. Iqbal, "Exploring the relationship between achievement goal orientation and cognitive engagement in engineering problem-solving," in *Proc. IEEE Frontiers in Education Conf. (FIE)*, 2025.

[6] Z. Abideen, S. Jabeen, O. Lawanto, A. Minichiello, and A. Iqbal, "Exploring the impact of the need for cognitive closure on undergraduate engineering students' cognitive engagement during problem solving," in *Proc. IEEE Frontiers in Education Conf. (FIE)*, 2025.

[7] Z. Abideen, S. Jabeen, and A. Iqbal, "A systematized literature review on the role of goal orientation in active learning and engagement during problem-solving in STEM education," in *Proc. IEEE Frontiers in Education Conf. (FIE)*, 2025.

[8] Z. Abideen, O. Lawanto, T. Naqash, and A. Minichiello, "A systematized literature review on problem-solving in STEM education exploring the impact of task complexity on cognitive factors and student engagement," in Proc. 2024 *ASEE Annu. Conf. & Exposition, 2024*, doi: 10.18260/1-2--46499.

[9] B. Thompson, S. Kirby, and K. Smith, "Culture shapes the evolution of cognition," *Proc. Natl. Acad. Sci.* USA, vol. 113, no. 16, pp. 4535–4542, 2016.

[10] R. E. Nisbett, *The Geography of Thought: How Asians and Westerners Think Differently—and Why*. New York, NY, USA: *Free Press*, 2003.

[11] B. R. Hergenhahn and M. H. Olson, *An Introduction to Theories of Learning*. Upper Saddle River, NJ, USA: *Pearson/Prentice Hall*, 2005.

[12] D. Gentner and D. R. Gentner, "Flowing waters or teeming crowds: Mental models of electricity," in Mental Models, D. Gentner and A. L. Stevens, Eds. Hillsdale, NJ, USA: *Erlbaum*, 1983, pp. 99–129.

[13] D. Gentner, "Psychology of mental models," in International Encyclopedia of the Social and Behavioral Sciences, N. J. Smelser and P. B. Bates, Eds. Amsterdam: *Elsevier*, 2002, pp. 9683–9687

[14] S. Scielzo, S. M. Fiore, H. M. Cuevas, and E. Salas, "The utility of mental model assessment in diagnosing cognitive and metacognitive processes for complex training," in *Proc. Hum. Factors Ergonomics Soc. Annu. Meeting*, vol. 46, no. 3, pp. 566–568, 2002, doi: 10.1177/154193120204600373.

[15] O. Lawanto, A. Minichiello, and Z. ul Abideen, “Impeccable learning: Deep understanding, self-evaluation, and monitoring for problem-solving success,” in *Proc. 2025 ASEE Annu. Conf. Expo.*, Montreal, QC, Canada, 2025.

[16] O. Lawanto, “Integrating metacognitive knowledge and self-regulation: Insights from problem-solving in engineering mathematics courses,” in Proc. 2025 *ASEE Annu. Conf. Expo.,* , Montreal, Quebec, Canada, June 22-25, 2026

[17] R. A. Bjork, J. Dunlosky, and N. Kornell, “Self-regulated learning: Beliefs, techniques, and illusions,” *Annu. Rev. Psychol.*, vol. 64, pp. 417–444, 2013, doi: 10.1146/annurev-psych-113011-143823.

[18] O. Lawanto, A. Minichiello, and R. Himel, “Students’ self-evaluative judgements in problem-solving: How accurate are they?,” in *ICERI2023 Proc.*, 2023, pp. 1253–1259.

[19] O. Lawanto and Z. Abideen, “Routine vs. non-adaptive learning: Examining the impact of metacognitive knowledge and self-regulation on problem-solving,” in Proc. 2025 *ASEE Annu. Conf. & Exposition, 2025*, Montreal, Quebec, Canada, June 22-25, 2026.

[20] P. Frank, *Einstein: His Life and Times*, G. Rosen, Trans. New York, NY, USA: *Alfred A. Knopf*, 1947.

[21] J. Salazar-Torres, A. Morales-Carruyo, and O. Y. Buitrago-Suescún, “The rubric as an assessment tool for solving problem situations in the physics and mathematics teaching context,” *J. Phys.: Conf. Ser.*, vol. 1981, no. 1, p. 012018, 2021, doi: 10.1088/1742-6596/1981/1/012018.

[22] J. L. Docktor et al., “Assessing student written problem solutions: A problem-solving rubric with application to introductory physics,” *Phys. Rev. Phys. Educ. Res.,* vol. 12, no. 1, p. 010130, 2016, doi: 10.1103/PhysRevPhysEducRes.12.010130.

[23] Association of American Colleges and Universities, “*Problem solving VALUE rubric,” 2009*. [Online]. Available: https://www.aacu.org/initiatives/value-initiative/value-rubrics/value-rubrics-problem-solving.

[24] D. H. Jonassen, “Toward a design theory of problem solving,” *Educ. Technol. Res. Dev.*, vol. 48, no. 4, pp. 63–85, 2000.

[25] Z. U. Abideen, *Cognitive Dynamics in Undergraduate Engineering Education: Effects of Need for Cognitive Closure and Achievement Goal Orientation on Cognitive Engagement*. Logan, UT, USA: *Utah State Univ.*, 2025, doi: 10.26076/52e3-0b05.

[26] K. A. Ericsson and A. C. Lehmann, “Expert and exceptional performance: Evidence of maximal adaptation to task constraints,” *Annu. Rev. Psychol.,* vol. 47, no. 1, pp. 273–305, 1996, doi: 10.1146/annurev.psych.47.1.273.

[27] K. E. Johnson and C. B. Mervis, “Effects of varying levels of expertise on the basic level of categorization,” *J. Exp. Psychol.:* Gen., vol. 126, no. 3, pp. 248–277, 1997, doi: 10.1037/0096-3445.126.3.248.

[28] M. K. Dahouda and I. Joe, “A deep-learned embedding technique for categorical features encoding,” *IEEE Access*, vol. 9, pp. 114381–114391, 2021, doi: 10.1109/ACCESS.2021.3104357.

[29] M. Bohanec, “Car Evaluation,” *UCI Machine Learning Repository*, 1988, doi: 10.24432/C5JP48.

[30] M. Bohanec and V. Rajkovic, “Knowledge acquisition and explanation for multiattribute decision making,” in *Proc. 8th Int. Workshop Expert Syst. Appl.,* Avignon, France, 1988, pp. 59–78.

[31] G. Lemaître, F. Nogueira, and C. K. Aridas, “*SMOTEN — Version 0.14.1*,” *imbalanced-learn*, 2024. [Online]. Available: https://imbalanced-learn.org/stable/references/generated/imblearn.over_sampling.SMOTEN.html

[32] M. Grandini, E. Bagli, and G. Visani, “Metrics for multi-class classification: An overview,” *arXiv*, 2020. [Online]. Available: https://arxiv.org/abs/2008.05756

[33] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification with unsaturated class distributions,” *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–444, 2009.

[34] T. Naqash, M. H. Tanveer, S. H. Shah, and M. Salman, “Statistical analysis-based intrusion detection for software defined network,” in *Smart Trends in Computing and Communications*, Y. D. Zhang, T. Senjyu, C. So-In, and A. Joshi, Eds. Singapore: Springer, 2022.

[35] J. Qiu, “An analysis of model evaluation with cross-validation: Techniques, applications, and recent advances,” *Adv. Econ., Manage. Polit. Sci.,* vol. 99, no. 1, pp. 69–72, 2024, doi: 10.54254/2754-1169/99/2024OX0213.

[36] F. S. Nahm, “Receiver operating characteristic curve: Overview and practical use for clinicians,” *Korean J. Anesthesiol.,* vol. 75, no. 1, pp. 25–36, 2022, doi: 10.4097/kja.21209.

[37] O. Lawanto, A. Minichiello, T. Naqash, and Z. Abideen, “Board 193: Adaptive v. faulty adaptive learning: The interplay between knowledge about task and self-regulation,” in Proc. 2024 *ASEE Annu. Conf. & Exposition, 2024*, Portland, Oregon, June 23-26, 2026, doi: 10.18260/1-2--46758.

[38] Z. Abideen and S. Jabeen, “The role of context in mathematical problem-solving: Bridging informal and formal learning: A systematized literature review,” in Proc. 2025 *ASEE Annu. Conf. & Exposition, 2025*, Montreal, Quebec, Canada, June 22-25, 2026

[39] R. Himel, O. Lawanto, Y. Brijmohan, W. Goodridge, Z. Abideen, F. Zinnah, S. Jabeen, (in press). “Exploring the Impact of Contextual Framing and Construct-Level Complexity on Engineering Problem-Solving Performance: A Study on Electrical Circuit Course” *2026 ASEE Annual Conference & Exposition*, Charlotte, North Carolina, June 21 - 24, 2026

[40] S. A. A. Lalin, M. N. Q. Ahmed, and S. M. A. Haq, "The effects of the COVID-19 pandemic on students' academic performance and mental health: An overview," *Reg. Sci. Policy Pract.*, vol. 16, no. 7, Art. no. 100046, 2024, doi: 10.1016/j.rspp.2024.100046.

Appendix

MODIFIED RUBRIC FOR ENGINEERING PROBLEM-SOLVING

	5	4	3	2	1	0
USEFUL DESCRIPTION	The description is useful, appropriate, and complete.	The description is useful but contains minor omissions or errors.	Parts of the description are not useful, missing, and/or contain errors.	Most of the description is not useful, missing, and/or contains errors.	The entire description is not useful and/or contains errors.	The solution does not include a description and it is necessary for this problem /solver.
ENGINEEERING APPROACH	The engineering approach is appropriate and complete.	The engineering approach contains minor omissions or errors.	Some concepts and principles of the engineering approach are missing and/or inappropriate.	Most of the engineering approach is missing and/or inappropriate.	All of the chosen concepts and principles are inappropriate.	The solution does not indicate an approach, and it is necessary for this problem/ solver.
SPECIFIC APPLICATION OF ENGINEEERING	The specific application of engineering is appropriate and complete.	The specific application of engineering contains minor omissions or errors.	Parts of the specific application of engineering are missing and/or contain errors.	Most of the specific application of engineering is missing and/or contains errors.	The entire specific application is inappropriate and/or contains errors.	The solution does not indicate an application of engineering and it is necessary.
MATHEMATICAL PROCEDURES	The mathematical procedures are appropriate and complete.	Appropriate mathematical procedures are used with minor omissions or errors.	Parts of the mathematical procedures are missing and/or contain errors.	Most of the mathematical procedures are missing and/or contain errors.	All mathematical procedures are inappropriate and/or contain errors.	There is no evidence of mathematical procedures, and they are necessary.
LOGICAL PROGRESSION	The entire problem solution is clear, focused, and logically connected.	The solution is clear and focused with minor inconsistencies	Parts of the solution are unclear, unfocused, and/ or inconsistent.	Most of the solution parts are unclear, unfocused, and/ or inconsistent.	The entire solution is unclear, unfocused, and/ or inconsistent.	There is no evidence of logical progression, and it is necessary.