Enhancing the Admissions Process for Engineering Baccalaureate Programs: A Machine Learning Approach to Developing a Valid and Reliable Examination

Marben S. Ramos, John Raymond B. Barajas, Pee Jay N. Gealone, Nico O. Aspra, Oliver M. Padua,
Arpon T. Lucero, Jr.

Abstract— **This paper details an approach to identify multiple-choice questions that are most effective in discriminating deficiencies in mathematics competencies of incoming first-year engineering students who are graduates of the K-12 program that was recently implemented in the Philippines. To achieve this objective, machine learning algorithms such as the k-Nearest Neighbors (kNN), Logistic Regression (LR), Random Forest (RF), Decision Tree (DT), and Gradient Boosting Machines (GBM) were implemented. From a question bank containing 1,300 questions covering Algebra (A), Advanced Algebra (AA), Plane and Spherical Trigonometry (T), Analytic Geometry (AG), and Solid Mensuration (SM), five domain experts identified the suitability of the questions as part of a diagnostic examination in mathematics. Specifically, using a 5-point Likert scale (5 being the highest), the experts rated the suitability of each question to test the proficiency of a student in 23 mathematics competencies based on what is prescribed by the Commission on Higher Education (CHED). The collected survey data were then used to train the machine learning models, which extracted patterns to identify the questions that would be most suitable to test the mathematics competencies of incoming first-year engineering students. With a precision recall score of 99.90%, the LR model was selected as the best performing model and analysis of how the LR model predicts the labels through the use of shap values revealed that the preference was given towards questions which test student proficiency in foundational mathematics competencies like that of Algebra and Analytical Geometry. Overall, these findings provided a better understanding of the questions that are most effective in discriminating student deficiencies in mathematics subjects.**

I. INTRODUCTION

The implementation of the Enhanced Basic Education Act of 2013, as outlined in [1], has prompted higher education institutions (HEIs) in the Philippines to revise their admission policies, especially in the engineering baccalaureate programs they offer. With the mandate to accept K-12 graduates regardless of track [2], non-STEM students who meet HEI admission requirements can now pursue engineering baccalaureates even if their specialization is not aligned to the degree they are accepted in. This scenario is expected to result in a large disparity in basic mathematics proficiency among first-year engineering students. Although no evidence has been published yet to validate this scenario, prominent universities in the Bicol Region have shared experiences that suggest this scenario is already being experienced [3].

Ideally, to address this development, diagnostic tests are conducted to assess the current mathematics proficiencies of students who are pursuing a degree in engineering [4]. However, with the implementation of the K-12 curriculum, these tests are no longer employed since incoming first-year engineering students are already expected to be proficient in Algebra (A), Advance Algebra (AA), Plane and Spherical Trigonometry (T), Analytical Geometry (AG), and Solid Mensuration (SM) prerequisite competencies. This present practice is seen to be detrimental to entirety of incoming first-year engineering students since proficiencies of these students in prerequisite mathematics competencies are expected to be highly diversified. For instance, STEM students are in an advantageous position since these subjects are part of their graduation requirements, whereas non-STEM graduates would have little to no exposure to A, AA, T, AG, and SM subjects since these are not included in their curriculum [5]-[6]. It is therefore crucial to institutionalize a diagnostic tool that intend to capture the gaps in the mathematics competencies of students who wish to pursue an engineering degree.

To address this problem, this paper presents a systematic development of a diagnostic examination that can potentially be institutionalized with the help of machine learning classification algorithms (MLCA). Using prerequisite mathematics competencies identified in policies, standards, and guidelines (PSG) published by the Commission on Higher Education (CHED) [7]-[12], a pool of 1300 questions

M. Ramos is with the Divine Word College of Legazpi, Legazpi City, Albay, Philippines and is presently the Dean of the School of Engineering and Computer Studies. He also serves as a professional lecturer with Bicol University College of Engineering, Legazpi City, Albay, Philippines (email: marbz_ramos@dwc-legazpi.edu).

J. R. Barajas is with the Bicol University College of Engineering Chemical Engineering Department, Legazpi City, Albay, Philippines (email: jrbbarajas@bicol-u.edu.ph).

P. J. Gealone is with the Bicol University College of Engineering Electrical Engineering Department, Legazpi City, Albay, Philippines (email: pjngealone@bicol-u.edu.ph).

N. Aspra is with the Bicol University College of Industrial Technology Mechanical Technology Department, Legazpi City, Albay, Philippines (email: nico.aspra@bicol-u.edu.ph).

O. Padua is with the Bicol University College of Engineering Civil Engineering Department, Legazpi City, Albay, Philippines (email: ompadua@bicol-u.edu.ph).

A. Lucero, Jr. is with the Bicol University College of Engineering Chemical Engineering Department, Legazpi City, Albay, Philippines (email: ajdlucero@bicol-u.edu.ph).

covering A, AA, T, AG, and SM subjects was collected. A team of five domain experts was commissioned to label questions that would be effective in discriminating mathematics deficiencies among incoming first-year engineering students. The collected data from the assessment of the experts were then fed to five MLCA for further analysis. Interpretation of the best-performing MLCA using shap values revealed that questions testing proficiency in two mathematics subjects are most preferred to ensure effective identification of student deficiencies. By utilizing this substantiated finding, identifying gaps in mathematical proficiencies among incoming first-year engineering students that are critical to guarantee the successful completion of their engineering degree can be more efficiently achieved by designing a diagnostic assessment tailored to identify deficiencies in these domain expert-identified critical subjects.

II. METHODOLOGY

This paper utilized explainable machine learning to identify multiple choice questions that would be most discriminatory in identifying student deficiencies in A, AA, T, AG, and SM subjects. To achieve this objective, the following MLCA were employed: (1) k-Nearest Neighbors (kNN), (2) logistic regression (LR), (3) random forest, (RF) (4) decision tree (DT), and (5) gradient boosting machines (GBM). Interpretation of how the best performing MLCA predicts the label of a question was facilitated through the use of shap values, as in [13]-[16].

A. Collection of Multiple-Choice Questions

Fourteen (14) engineering faculty members willingly participated in creating a question bank that was initially developed through a shared cloud drive. Over the course of three months, they developed 3000 multiple choice questions covering A, AA, T, AG, and SM subjects, which were stored in the shared cloud drive. To assess these questions, five (5) domain experts were commissioned (see Table I for demographics of the domain experts). Due to limited resources, however, only 1300 questions (260 for each subject) were assessed as agreed upon during negotiations with the experts.

TABLE I. DEMOGRAPHICS OF TAPPED DOMAIN EXPERTS

Demographics	**Frequency**
Average Age (years)	51 ± 12.6
Average Years of Teaching Experience	17 ± 5.5
Courses Taught	
College Algebra	5
Plane and Spherical Trigonmetry	4
Advance Algebra	5
Analytical Geometry	4
Solid Mensuration	4
Differential Calculus	4
Integral Calculus	5
Differential Equations	3
Highest Degree Earned	
Master's Degree	4
Doctrate Degree	1

B. Development of Question Assessment Tool

The prerequisite competencies identified in the CHED PSGs common to engineering programs offered in the Philippines were used as the basis to assess the suitability of the pooled questions in identifying mathematics deficiencies of incoming first-year engineering students [7]-[12]. Specifically, the primary objective of this instrument was to seek expert opinion on which questions covering A, AA, T, AG, and SM are most suitable to include in a 50-item multiple choice diagnostic examination. For instance, the domain experts were asked to rate using a 5-point Likert scale (5 being the highest) the suitability of a question to test the proficiency of a student in the 23 mathematics competencies itemized under Table II. In addition, at the last column of the instrument, the experts were asked to label the questions either as 0 or 1, with 1 indicating that the question should be included in the diagnostic examination. Reliability analysis of the instrument revealed a Cronbach's alpha coefficient of 0.7485, indicating that the developed instrument is indeed internally consistent and reliable [17]. Data collected from this survey were then collated as the arithmetic average of the expert ratings, with the final label assignment determined as the majority.

TABLE II. IDENTIFIED PREREQUISITE COMPETENCIES

Competency Code[a]	**Description of Competency**[b]
A_1	Simplify and manipulate algebraic expressions
A_2	Determine the solution sets of algebraic equations, exponential and logarithmic equations; and inequalities
A_3	Apply algebraic concepts and techniques to solve word problems
A_4	Identify the domain and range of a given relation/function
AA_1	Determine the solution sets of inequalities
AA_2	Determine the solution sets of systems involving quadratics
AA_3	Apply advance algebra concepts and techniques to solve word problems
AA_4	Perform operations and manipulations on matrices and determinants
AA_5	Solve systems of linear equations using matrices and determinants
AA_6	Find the sum of elements in arithmetic and geometric sequences
T_1	Define angles and their measurements using degrees and radians
T_2	Evaluate the sine, cosine, tangent, cosecant, secant, and cotangent functions
T_3	Prove the properties and identities of trigonometric functions

Competency Code[a]	Description of Competency[b]
T_4	Define and evaluate inverse trigonometric functions
T_5	Solve trigonometric equations
T_6	Solve problems involving right triangles using trigonometric function definitions for acute angles
T_7	Solve problems involving oblique triangles by the use of the sine and cosine laws
AG_1	Set up equations given enough properties of lines and conics
AG_2	Draw the graph of the given equation of the line and the equation of the conic section
AG_3	Analyze and trace completely the curve, given their equations in both rectangular and polar coordinates, in two-dimensional space
SM_1	Calculate the area of plane figures using appropriate formulas
SM_2	Calculate the surface areas and volumes of various types of solids using appropriate formulas
SM_3	Determine the volumes and surface areas of solids using other methods such as the theorems of Pappus

[a] A – Algebra; AA – Advance Algebra; T – Plane and Spherical Trigonometry; AG – Analytic Geometry; SM – Solid Mensuration

[b] Competency descriptions were taken from approved CMOs [11]-[16].

C. Model Training and Selection Phase

Five MLCA were used to create a machine learning model that would extract patterns in the survey data (dataframe with 1300 records containing 24 features, inclusive of the assigned labels) which would identify questions most suitable to test the mathematics competencies of incoming first-year engineering students. Particularly, the kNN, LR, RF, DT, and GBM classification algorithms, as implemented in [18]-[19] at default parameters, were used to identify these patterns. To achieve this objective, 75% of the data were used as the training and validation set while the remaining were used as the holdout set. At a proportion chance criterion of 50.01%, the training data was substantially balanced, requiring only a 5-split stratified K-fold for the training and validation phase of the MLCA models. Finally, the precision-recall (PR) curve was used the criterion to identify which among the trained MLCAs performed best to predict the labels in the survey data.

D. Model Interpretation Phase

To gain insights into how the best performing MLCA predicts the label of a question based on the mathematics competencies listed in Table II, this study utilized shap values as implemented in [13]-[16]. Through the calculation and visualization of these values, this study was able to better understand the impact of the suitability of a question in testing the identified mathematics competencies on the label of the question given by the domain experts.

E. Model Validation Phase

As the final phase of this study, the best performing MLCA was validated through a pilot-test diagnostic examination deployed to 141 incoming first-year engineering students who expressed voluntary participation. Out of a pool of 641 expert-selected questions, 10 questions for each of the five identified subjects were randomly selected, resulting in a 50-item multiple choice diagnostic examination. The diagnostic scores obtained by the participants were then used to confirm the insights derived from the analysis of how the best performing MLCA predicts the label of a question.

From a pool of 641 expert selected questions, diagnostic scores garnered by the participants through a 50-item multiple choice diagnostic examination (10 questions taken randomly for each of the five identified subjects) were used to confirm the insights derived from the interpretation of how the best performing MLCA predicts the label of a question.

III. RESULTS AND DISCUSSION

A. Mean Expert Ratings of the Collated Data

Analysis of the mean expert ratings grouped by the expert assigned question labels revealed a clear preference primarily on Algebra and secondarily to Trigonometry questions. As shown in Figure 1, it is worth noting that the opinion of domain experts put emphasis to test A_1, A_2, A_3, and A_4 competencies of incoming first-year engineering students. This is consistent to what is observed in [20]-[25] and this finding further solidify the reliability of the data as basis for determining characteristics of a question that should be included in the diagnostic examination.

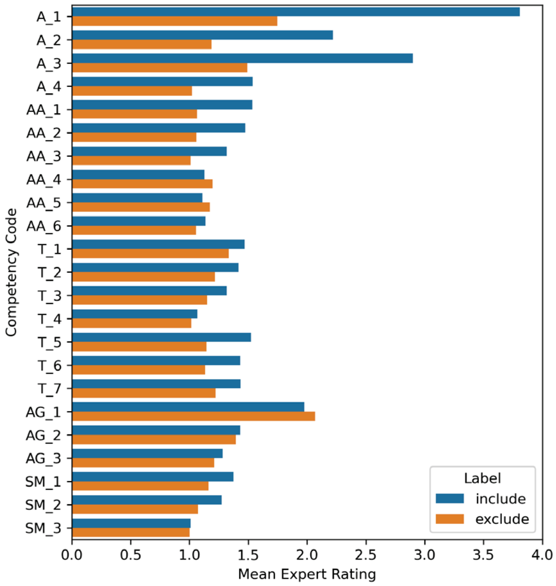

Figure 1. Barplot of Mean Expert Rating Grouped by Assigned Label.

B. Best Performing Model

Table III presents the precision-recall scores of the trained MCLA models. Comparison of the training, validation, and holdout scores revealed that both the Random Forest and Decision Tree models demonstrate overfitting on the data and that Logistic Regression was the best performing model. Tuning of the Logistic Regression model at a C value of 100 further increased its holdout precision-recall score to 99.90%. At this tuned parameter, this indicated that only 2 out 165 (or about 1.21%) questions from the holdout data were being misclassified as false-positives (no false-negatives were returned by the model). These results strongly support the use of the trained Logistic Regression model as the basis of understanding the patterns on how a question is predicted to be included in the diagnostic examination.

TABLE III. PERFORMANCE OF TRAINED MODELS

Model	Training PR Score (%)	Validation PR Score (%)	Holdout PR Score (%)
k-Nearest Neighbors	95.29	91.31	93.76
Logistic Regression	99.35	98.34	98.77
Random Forest	100.00	95.92	94.75
Decision Tree	100.00	93.42	94.00
Gradient Boosting Machnie	98.19	94.18	92.71

C. *Interpretation of Best Performing Model*

Interpretation of how the features affect the label prediction of the best performing model is shown in Figure 2 as a beeswarm plot. Parallel to the findings shown in Figure 1, A_1 and A_3 competencies were identified as the top predictors of the question labels. As previously observed in reviewed literature [20]-[25], proficiency in Algebra competencies is foundational to engineering programs and these even are identified as markers which predict the successful completion of an engineering program. Surprisingly, AG_1 was also identified as one of the top predictors of the best performing model. While this may not be conclusive, it may be speculated that proficiency in AG_1 would be highly dependent on the proficiency in A_1. As seen under Table II, AG_1 is defined as the competency for setting-up equations of lines and conics. It is therefore expected that if a student is proficient in AG_1, it would follow that this student would also be proficient in A_1. While interplay between these features is already beyond the scope of this work, it is highly likely that the interplay between these features could be understood better if analyzed in the perspective of latent variables.

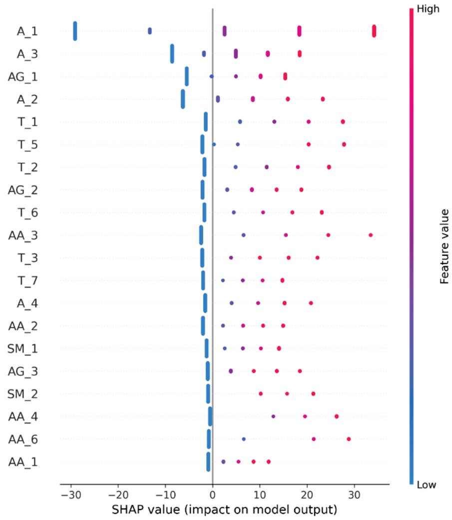

Figure 2. Beeswarm Plot of the Feature of the Best Performing Model.

D. *Validation of Best Performing Model*

To further solidify the insights derived from the interpretation of the best performing model, this study found it necessary to validate the applicability of the model selected questions as a diagnostic examination. As shown in Figure 3, results indicated that the majority of the participants garnered diagnostic scores that are less than 50%. Comparing the means of this result with the means of the collective mathematics grade the participants garnered in both their junior and senior high years (mean grade of 83.24 ± 4.8 and p-value that is less than 0.0001 using Welch's T-statistic), it can be inferred that the participants is expected to demonstrate gaps in the identified prerequisite competencies. It is also interesting to note that STEM and non-STEM participants have almost the same median scores. This finding may also indicate the need to implement intervention programs not only to non-STEM enrollees but also to STEM enrollees. While this work acknowledges that additional researches may be needed to further validate the arguments presented in this study, it is hoped that these would serve as a baseline or point of comparison to future researches in this area.

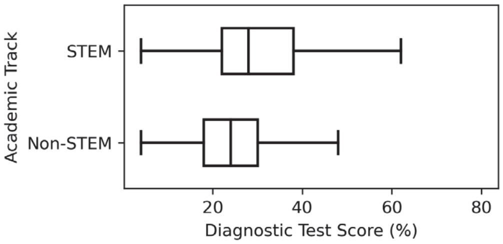

Figure 3. Results of Diagnostic Examination.

E. *Practical Implications of Findings*

The trained model in this study offers a valuable decision-making tool for engineering faculty that is aimed at simplifying the question selection process and providing a more informed basis for decisions regarding the inclusion or exclusion of a question based on its rating. This described value is demonstrated in Figure 4 and Figure 5, which show how the trained model determines the suitability of a question by including or excluding it based on its ratings. Specifically, the model identifies questions that test for proficiency in specific mathematics competencies, such as A_3, AG_1, and A_1, providing a clear guideline for question selection and revision. In addition, these force-plots provide a straightforward way for faculty to determine which questions should be included in diagnostic examinations, without the need for time-consuming item analysis. Overall, the trained model has the potential to enhance the effectiveness and impact of assessments of mathematics competencies in first-year engineering students.

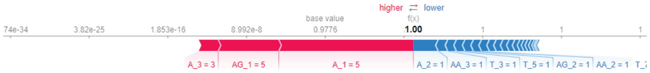

Figure 4. Force-plot of an Included Question.

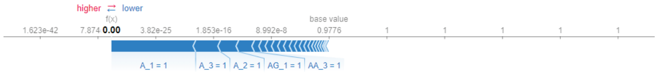

Figure 5. Force-plot of an Included Question.

IV. Future Work

As the field of technology continues to rapidly advance, including the emergence of revolutionary technologies such as ChatGPT [26], it is important to recognize that the machine learning model developed in this study must be transformed into an engaging and reliable data product in order to effectively support engineering faculty who are seeking to integrate artificial intelligence into their teaching methods. As a recommendation for future work, this study suggests packaging the question selection process as a GPT model problem to significantly enhance its usability and deployability to target users. However, like that of ChatGPT, it is highly crucial to thoroughly evaluate and test the efficacy of this approach prior to deploying this on a larger scale.

Acknowledgment

The authors gratefully acknowledge the generous support provided by Divine Word College of Legazpi which made this research possible. The authors would also like to extend their gratitude to Bicol University for the administrative support given during the conduct of this study. Moreover, the opinions expressed in this manuscript are solely those of the authors and do not reflect the position of Divine Word College of Legazpi and its respective offices, and Bicol University and its respective offices.

References

[1] "Republic Act No. 10533 (Enhanced Basic Education Act of 2013)," 2013. https://www.officialgazette.gov.ph/2013/05/15/republic-act-no-10533/. Accessed: February 24, 2023.

[2] "Commission on Higher Education Memorandum Order (CHED-CMO) No. 105, s. 2017, Policy on the admission of senior high school graduates to the higher education institutions effective academic year 2018-2019," 2017. https://ched.gov.ph/cmo-105-s-2017/. Accessed: February 24, 2023.

[3] J.R. Barajas, P.J. Gealone, O. Padua, N. Aspra, and A. L., Jr., "An informal interview on the status of K-12 students in BUCENG", unpublished.

[4] B. Eleri, J. Prior, S. Lloyd, and S. Thomas, "Engineering more engineers – bridging the mathematics and careers advice gap," Eng. Ed., vol. 2, no. 1, pp. 23-32, 2007.

[5] "Guidelines on the implementation of the senior high school (SHS) program in existing public junior high schools (JHSs) and integrated schools (ISs), establishment of stand-alone public SHSs, and conversion of existing public elementary and JHSs into stand-alone SHSs," https://www.deped.gov.ph/wp-content/uploads/2015/10/DO_s2015_51_0.pdf. Accessed: February 24, 2023.

[6] "Policy Guidelines on the K to 12 Basic Education Program," 2019. https://www.deped.gov.ph/wp-content/uploads/2019/08/DO_s2019_021.pdf. Accessed: February 24, 2023.

[7] "Policies and standards (PS) for the degree bachelor of science in civil engineering (BSCE)," pp. 1-50, 2007. https://ched.gov.ph/cmo-29-s-2007/. Accessed: August 12, 2022.

[8] "Policies and standards (PS) for the degree bachelor of science in chemical engineering (BSCHE)," pp. 1-50, 2008. https://ched.gov.ph/cmo-23-s-2008/. Accessed: August 12, 2022.

[9] "Policies and standards (PS) for the degree bachelor of science in electrical engineering (BSEE)," pp. 1-50, 2008. https://ched.gov.ph/cmo-34-s-2008/. Accessed: August 12, 2022.

[10] "Policies and standards (PS) for the degree bachelor of science in geodetic engineering (BSGE)," pp. 1-50, 2008. https://ched.gov.ph/cmo-12-s-2008/. Accessed: August 12, 2022.

[11] "Policies and standards (PS) for the degree bachelor of science in mechanical engineering (BSME)," pp. 1-50, 2008. https://ched.gov.ph/cmo-9-s-2008/. Accessed: August 12, 2022.

[12] "Policies and standards (PS) for the degree bachelor of science in mining engineering (BSME)," pp. 1-50, 2008. https://ched.gov.ph/cmo-10-s-2008/. Accessed: August 12, 2022.

[13] S. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions." arXiv, 2017. doi: 10.48550/ARXIV.1705.07874.

[14] S. M. Lundberg et al., "From local explanations to global understanding with explainable AI for trees," Nature Machine Intelligence, vol. 2, no. 1. Springer Science and Business Media LLC, pp. 56–67, Jan. 17, 2020. doi: 10.1038/s42256-019-0138-9.

[15] R. Mitchell, E. Frank, and G. Holmes, "GPUTreeShap: Massively Parallel Exact Calculation of SHAP Scores for Tree Ensembles." arXiv, 2020. doi: 10.48550/ARXIV.2010.13972.

[16] S. M. Lundberg et al., "Explainable machine-learning predictions for the prevention of hypoxaemia during surgery," Nature Biomedical Engineering, vol. 2, no. 10. Springer Science and Business Media LLC, pp. 749–760, Oct. 10, 2018. doi: 10.1038/s41551-018-0304-0.

[17] Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. Journal of Applied Psychology, 78(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

[18] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011. [Online]. Available: https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf.

[19] P. Virtanen et al., "SciPy 1.0: fundamental algorithms for scientific computing in Python," Nature Methods, vol. 17, no. 3. Springer Science and Business Media LLC, pp. 261–272, Feb. 03, 2020. doi: 10.1038/s41592-019-0686-2.

[20] B. Pepin, R. Biehler, and G. Gueudet, "Mathematics in Engineering Education: a Review of the Recent Literature with a View towards Innovative Practices," International Journal of Research in Undergraduate Mathematics Education, vol. 7, no. 2. Springer Science and Business Media LLC, pp. 163–188, Apr. 01, 2021. doi: 10.1007/s40753-021-00139-8.

[21] G. Greefrath, W. Koepf, and C. Neugebauer, "Is there a link between Preparatory Course Attendance and Academic Success? A Case Study of Degree Programmes in Electrical Engineering and Computer Science," International Journal of Research in Undergraduate Mathematics Education, vol. 3, no. 1. Springer Science and Business Media LLC, pp. 143–167, Oct. 26, 2016. doi: 10.1007/s40753-016-0047-9.

[22] R. Biehler, E. Lankeit, S. Neuhaus, R. Hochmuth, C. Kuklinski, E. Leis, M. Liebendörfer, N. Schaper, and M. Schürmann, "Different goals for pre-university mathematical bridging courses – comparative evaluations, instruments and selected results," in Proceedings of INDRUM 2018 Second Conference of the International Network for Didactic Research in University Mathematics, V. Durand-Guerrier, R. Hochmuth, S. Goodchild, and N. M. Hogstad, Eds. Kristiansand, Norway: Universitetet i Agder, 2018, pp. 467-476. [Online]. Available: https://indrum2018.sciencesconf.org/data/Indrum2018Proceedings.pdf

[23] S. Treffert-Thomas, "Getting into university: From foundation to first year engineering," in T. Dooley and G. Gueudet (Eds.), Proceedings of the Tenth Congress of the European Mathematical Society for

Research in Mathematics Education, Dublin, Ireland: DCU Institute of Education and ERME, 2017, pp. 2322-2323.

[24] M. Basitere, E. Rzyankina, and P. Le Roux, “Reflection on Experiences of First-Year Engineering Students with Blended Flipped Classroom Online Learning during the COVID-19 Pandemic: A Case Study of the Mathematics Course in the Extended Curriculum Program,” Sustainability, vol. 15, no. 6. MDPI AG, p. 5491, Mar. 21, 2023. doi: 10.3390/su15065491.

[25] S. Busto, M. Dumbser, and E. Gaburro, “A Simple but Efficient Concept of Blended Teaching of Mathematics for Engineering Students during the COVID-19 Pandemic,” Education Sciences, vol. 11, no. 2. MDPI AG, p. 56, Feb. 02, 2021. doi: 10.3390/educsci11020056.

[26] R. Peres, M. Schreier, D. Schweidel, and A. Sorescu, “On ChatGPT and Beyond: How Generative Artificial Intelligence May Affect Research, Teaching, and Practice,” International Journal of Research in Marketing. Elsevier BV, Apr. 2023. doi: 10.1016/j.ijresmar.2023.03.001.