Automatic Assessment of Quadratic Equation Solutions Using MathBERT and RoBERTa Embeddings

Subhiksha Suresh Rao
Department of Computer Science & Engineering,
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
bl.en.u4cse21198@bl.students.amrita.edu

Sahil Mishra
Department of Computer Science & Engineering,
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
bl.en.u4cse21179@bl.students.amrita.edu

S Akhilesh
Department of Computer Science & Engineering,
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
bl.en.u4cse21175@bl.students.amrita.edu

Roshni M Balakrishnan
Department of Computer Science & Engineering,
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
m_roshni@blr.amrita.edu

Peeta Basa Pati
Department of Computer Science & Engineering,
Amrita School of Computing, Bengaluru
Amrita Vishwa Vidyapeetham, India
ORC ID: 0000-0003-2376-4591
bp_peeta@blr.amrita.edu

***Abstract*—This project aims to create a machine learning classification system that classifies solutions to quadratic equations into three different groups: "Incorrect," "Partially Correct," and "Correct." MathBERT and RoBERTa embeddings are utilized in this system. A multi-class classification technique is used by the system, with an emphasis on feature engineering, data preparation, and model selection. To improve model performance, methods like Principal Component Analysis (PCA), LIME (Local Interpretable Model-Agnostic Explanations), and ensemble meta models with Boosting, Voting, and Stacking are used. The efficacy of many classification models in correctly classifying quadratic solutions is examined in this work. To evaluate the performance of the model, evaluation measures including F1-score, accuracy, precision, and recall are used. The results show that MathBERT embeddings outperform RoBERTa embeddings. Ensemble models, especially those that include SVM and MLP classifiers, achieve the best performance, with the stacking approach yielding an accuracy of approximately 65%. This model accurately classifies solutions into the correct categories, demonstrating the potential for practical applications in mathematics education.**

***Index Terms*—Quadratic Equations, Machine Learning, Classification, Solution Accuracy, Automated Evaluation, Educational Technology, MathBERT, RoBERTa**

I. Introduction

The fundamental building blocks of mathematics, quadratic equations are used in many different fields' theoretical and practical applications. Their solutions, which are marked by complex patterns and a range of results, provide a significant obstacle to proper evaluation and feedback for both teachers and students. The manual grading process used in traditional techniques for evaluating quadratic solutions can be tedious, subjective, and prone to inconsistencies. There is a strong need for automated classification systems that can reliably and precisely classify quadratic solutions due to the growing use of technology in education. This study gives a thorough research into the creation of a machine learning-based categorization system for solutions to quadratic equations, driven by this necessity. Our strategy is to automate the assessment process and give instructors and students quick, relevant feedback by utilising recent developments in machine learning and natural language processing techniques. The use of MathBERT [15] embeddings, a BERT language model variation designed for mathematical content, is a key component of our research. Rich semantic information may be extracted with the help of MathBERT [15], which permits the encoding of quadratic solutions into high-dimensional vector representations, allowing for more complex classification choices. In addition to MathBERT [15] embeddings, our classification method includes many crucial components that improve its efficacy and resilience. To increase the openness and reliability of the system, interpretable and transparent insights into the categorization judgements are provided using Local Interpretable Model-Agnostic Explanations (LIME) [19]. Dimensionality reduction using Principal Component Analysis (PCA) [18] improves computing performance and allows for more effective handling of high-dimensional data. Additionally, in order to aggregate the predictions of several base classifiers and enhance overall classification performance, we investigate the use of ensemble meta models, which include Stacking, Voting [20], and Boosting strategies [22]. Through the utilisation of the combined intelligence of many classification models, ensemble meta models provide enhanced generalisation and resilience, ultimately improving the classification system's accuracy and dependability. Overall, by providing a creative and practical answer to the problem of automating the assessment of quadratic equation solutions, this study significantly advances the fields of mathematics education and educational technology. Our goals are to improve learning outcomes, encourage a deeper comprehension of quadratic ideas, and eventually advance the status of mathematics education in the digital era by giving instructors and students access to an automated and trustworthy categorization system.

II. Literature Survey

Likforman-Sulem and colleagues [1] presented an approach that includes an iterative hypothesis validation strategy to extract text lines from various handwritten documents using both Hough domain and image domain information. Fei Yin and co-authors [2] used the MST clustering method with innovative distance measures to construct a tree of connected components that successfully extracted lines of text from unconstrained handwritten documents. Tamas Verga and team [3] improved word extraction by integrating a context-based and thresholding approach using a structured tree for systematic traversal. Sami Baral et al. [5] presents a model that uses sentence-level semantic representations to automate the evaluation of student responses, which outperforms current benchmarks on various metrics. They emphasize contextual information and its importance and promote techniques such as universal sentence encoders and sentence BERT. Gurpratap Singh et al. [7] presents an independent method for generating large image datasets that is device agnostic and suitable for real-time applications. Their approach, exemplified by the Punjabi handwritten image dataset, uses contour and edge detection to align shapes according to defined parameters, providing a standardized way to collect large datasets in various fields. Andrew S. Lan et al. [8] present a data-driven mathematical language processing (MLP) framework for open-response automatic scoring of mathematical questions using solution data for correctness assessment, partial credit assignment, and error detection. MLP groups solutions by separating correct, partially correct and incorrect, which simplifies decision making on large educational platforms. Mengxue Zhang et al. [9] propose an Automated Short Answer Assessment (ASAG) using MathBERT [15], a math-focused variant of BERT. The framework, which demonstrates improved score prediction, incorporates learning in context that enables efficient generalization to previously unseen questions, highlighting the importance of fine-tuning for optimal performance. Ansong Ni et al. [10] discuss the limitations of pre-trained language models in handling multilevel formal reasoning tasks in mathematics tasks. To overcome the challenge of fine-tuning datasets that provide only one reference solution, they propose a method that includes self-selected, fully correct, and partially correct solutions, which improves the adaptability of the model to unseen cases. The study examines different training objectives and examines their significant impact on performance. Experimental results on mathematical reasoning materials highlight the effectiveness of the approach compared to learning from a single reference solution. In addition, the authors of [11] emphasize the importance of self-selection strategies of pre-trained language models for elementary school mathematics tasks. They show excellent performance by exploring a wider solution space and using different learning signals. Sami Baral [11] deals with the problems of automatic evaluation of open-ended mathematical questions by introducing the "Math Term Frequency" (MTF) model combined with SBERT-Canberra to answer challenges using mathematical words. Future research aims to explore state-of-the-art techniques and extend the strategy to different open-ended response categories. The overall goal is to improve automated assessment and feedback for both math teachers and students. In a related work [12], Baral investigates the use of OpenAI/CLIP in combination with optical character recognition (OCR) to improve the automatic evaluation of student responses containing both text and image elements. Although the proposed image representation methods do not outperform the current leading automatic scoring methods, they show comparable performance, suggesting possible improvements in image-based response processing by integrating established text and image embedding techniques. Yidan Tang [13] improves machine learning pattern recognition by combining traditional and structural features in PDF files. The detection module uses ensemble learning with NB, RF, and DT as base learners and logistic regression as a meta-learner, achieving an impressive 98.70% accuracy on the test set, outperforming Adaboost and deep learning DNN models. The method proves effective in detecting hidden malicious PDF files and has significant improvements over other detection models. Future studies will investigate the relationships between the structural features of malicious PDF files with the aim of improving the model and the ability to detect such files. Zhigang Li [14] presents EL-MS, a mobile traffic forecasting model that uses stacked ensemble learning with distributed MLP as the base learner and SSVR as the meta-learner. EL-MS outperforms complex network models by showing better prediction performance and robustness to training data uncertainty. In particular, it exceeds the predictive accuracy of certain advanced machine learning models and proves its value in tasks such as mobile traffic segmentation and wireless network state planning.

III. Methodology

Classifying solutions to quadratic equations according to accuracy has important ramifications for mathematical analysis and educational evaluation. In order to overcome this difficulty, this study suggests an organised approach for automatic categorization. The methods used to classify quadratic solutions in an accurate and effective manner are

explained in detail in the following sections.
In order to Automate the assesment of Quadratic Equation Solutions this paper adresses the following research questions:
RQ1) How does the embeddings generated by MathBERT [15] perform compared to that of RoBERTa [16] in classifying the quadratic solutions
RQ2) Does application of SMOTE and PCA affect the results
RQ3) Which ensembling technique shows the best performance

A. Overview:

Figure 1 illustrates the methodology utilized in this study. The procedure includes importing the data preprocessing, representing the features using Roberta [16] and MathBert [15] vectorization, reducing the dimensionality using Principal Component Analysis (PCA) [18], training and assessing several classification models, and building an ensemble model. The methodology guarantees a methodical and thorough approach to the categorization of quadratic solutions.

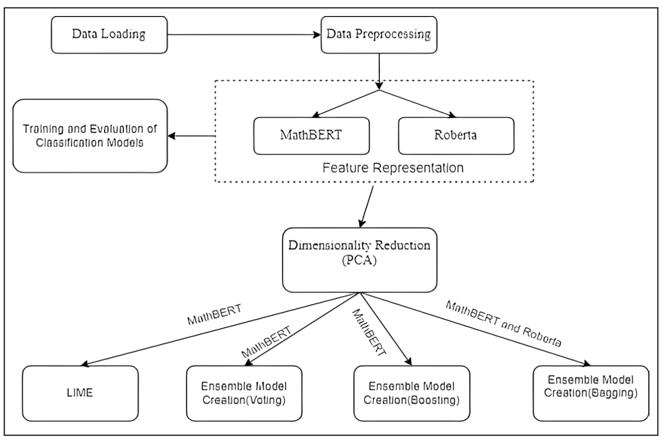

Fig. 1. Architecture Diagram of the Automatic Evaluation Process

B. Data Extraction and Preprocessing:

The dataset contains 1127 solutions to quadratic equation graded from 0 to 5. The criteria for grading is: The data is loaded and pre-processed to ensure consistency and quality. This includes steps such as missing data management, data cleaning, and data standardization. The database is collected from reliable sources and formatted for further analysis. Any missing values were imputed using appropriate methods, outliers were detected and corrected, and data were standardized to ensure uniformity across scales. Preprocessing ensures that the data set is ready for feature extraction and model training, laying a solid foundation for accurate classification.

C. Feature Representation:

Two different vectorization methods are used to numerically express the solution of quadratic equations:

1) MathBERT: MathBERT [15] is a variant of the BERT language model designed specifically for mathematical content. It allows encoding quadratic solutions into a high-dimensional vector representation, enabling a more accurate choice of classification. MathBERT [15] extracts rich semantic information and is suitable for capturing the nuances of mathematical language.
2) RoBERTa: RoBERTa [16] is a robust version of the BERT model that includes additional pretraining data and training methods. Although not specifically designed for mathematical content, RoBERTA [16] is effective in capturing the linguistic and semantic context, making it a suitable choice for numerical representation of quadratic solutions.

This method efficiently captures mathematical semantics and linguistic context, improving the classification process. The RoBERTa [16] and MathBERT [15] applications encode quadratic solutions into high-dimensional vector representations that enable more sophisticated classification options. Previous studies have shown that these methods perform well on comparable natural language processing problems, making them suitable for this classification problem.

D. Dimensionality Reduction:

Principal component analysis (PCA) [18] is used to measure the feature space of the data while retaining the required variance. This step aims to reduce measurement bias and improve model training efficiency. By reducing the number of features, PCA [18] helps capture the most important information by discarding excessive or less informative features. However, it is important to maintain the orthogonality and clarity of this process.

E. Ensemble Model Creation:

Ensemble models are built using various methods, including Voting [20], Boosting [22], and Bagging [21]. This ensemble method combines the prediction of several base classifiers to improve the overall prediction performance.

1) Voting:
   - Hard Voting: Combining the prediction of several state classifiers and the prediction of the class with the most votes. This approach is particularly useful when the base classifiers have different views of the data. [20]
   - Soft voting: It aggregates the predicted probabilities of each class from several base classifiers and predictions the class with the highest average probability. This method considers the confidence level of each classifier, which can lead to more accurate predictions. [20]
2) Boosting: Builds a strong classifier by sequentially training several weak classifiers, with each subsequent classifier paying more attention to cases misclassified by previous classifiers. This iterative process helps improve overall prediction accuracy by focusing on difficult cases. [22]
3) Bagging: It uses bootstrap sampling to independently train several base classifiers and then aggregate their

predictions by averaging or voting [20]. By reducing variance and overfitting, Bagging helps improve the stability and robustness of the final classifier. [21]

4) Stacking (SVM and MLP):
   - Base Models: As base models, the SVM and MLP classifiers were chosen. SVM is used because it is very effective in high-dimensional spaces and MLP because it can model complex non-linear relationships.
   - Integration Method: These models are trained independently, and their predictions are used to retrain the meta-model.
   - Meta-model: The other meta-model used is Logistic Regression. It used the prediction of both SVM and MLP to make a final prediction based on that, which effectively learned from the strengths and compensated for the weaknesses of both base models.
   - Why this stacking procedure: This method was chosen because it can synthesize various learning dynamics. The meta-model's ability to interpret the combined predictions led to a refined decision-making process that is more effective in dealing with complex classification tasks.

These ensemble methods exploit the complementary advantages of individual classifiers such that they bring about improved generalization and better performance compared to single-classifier systems. For example, a gain in classification accuracy and reliability was obtained by forming a stacked model based on SVMs and MLP analysis, which is included in our system.

F. Local Interpretable Model-Agnostic Explanations (LIME):

LIME is used to provide a clear and transparent understanding of the categorization judgments made by the model [19]. This helps to improve the reliability and validity of the classification process by explaining specific assumptions. LIME creates local explanations for model predictions by approximating the model's behavior around specific data points. This allows stakeholders to understand why certain assumptions were made and assess the reliability of the model. [19]

G. Hyperparameter Tuning

Hyperparameter tuning [23] is performed using GridSearch [17] to optimize the performance of each classification model. GridSearch [17] systematically explores the predefined hyperparameter space [23] and evaluates each combination through cross-sectional testing. This procedure ensures that the hyperparameters [23] of the model are adjusted to achieve the best performance in the database of quadratic equation solutions. By choosing the optimal combination of hyperparameters [23], GridSearch [17] improves the accuracy and reliability of the classification model, resulting in improved efficiency in accurately categorizing quadratic solutions. The specific hyperparameters [23] used for all the models in this research are mentioned in Table I.

TABLE I
HYPERPARAMETERS FOR TRADITIONAL CLASSIFICATION MODELS

Model	Hyperparameters
KNN	'algorithm': 'kd_tree', 'metric': 'euclidean'
SVM Classifier	'C': 0.1, 'degree': 4, 'gamma': 'auto'
Decision Tree	'criterion': 'gini', 'max_depth': 10
Naive Bayes	'var_smoothing': 1e-07
ADA Boost	'learning_rate': 0.1, 'n_estimators': 100
XG Boost	'learning_rate':0.001, 'max_depth': 17.0
CAT Boost	'depth': 10, 'min_data_in_leaf': 23
Random Forest	'max_depth': 11, 'max_features'= 'log2'
Logistic Regression	'fit_intercept': 'True', 'max_iter': 1000

H. Model Training and Evaluation:

A thorough training and assessment process is applied to ten different classification models: K-nearest neighbors, Support Vector Machine Classifier, Decision Tree, Naïve Bayes, ADA BOOST, XGBOOST, CAT BOOST, Multi-layer Perceptron, Random Forest, and Logistic Regression. For every model, cross-validation and hyperparameter tuning [23] was performed. The computation of performance indicators, such as accuracy, recall, and F1-score, serves as the foundation for further investigation.

I. Comparative Study:

A focused comparison analysis is carried out using two vectorization models, RoBERTa [16] and MathBERT [15], to assess the performance of the ensemble model and individual classifiers. Important measures like recall, accuracy, and precision are used to assess the two strategies' effectiveness in an unbiased manner. The objective is to find out the ways in which the ensemble models constructed using LIME [19], Soft Voting, Hard Voting [20], Bagging [21], and Boosting [22] approaches, makes use of the advantages of various classifiers and the influence of sophisticated vectorization techniques on classification result overall. The findings from a comparitive analysis provide light on the subtle performance distinctions between MathBERT [15] and RoBERTa [16]. They also help choose the best vectorization model for the quadratic solution classification problem.

J. Classification of Correctness:

The final phase involves classifying quadratic equation solutions into three distinct classes: partially correct, correct, and incorrect. This categorization makes it easier to provide users, teachers, or automatic grading systems with thoughtful input. To guarantee the openness and dependability of the categorization process, difficulties and factors to be taken into account when creating these categories are covered in detail.

K. Metrics and Justification:

The model assessment measures include accuracy, precision, recall, and F1-score. Recall is chosen to assess the model's capacity to identify all positive occurrences, precision is chosen to gauge the accuracy of positive predictions, and the F1-score is chosen as a balanced statistic that takes both precision and recall into account. The justification for using these measures is consistent with the research objective of precisely categorising quadratic solutions.

L. Technology Employed:

Python programming language,scikit-learn, LIME [19], and other technologies are used; their widespread use, community support, and efficacy in comparable research serve as justifications for their use. Their application guarantees the suggested methodology's scalability and repeatability.
In conclusion, the approach that has been provided offers a methodical and organised way to classify solutions to quadratic equations according to their accuracy. The suggested classification approach is more accurate and successful when complex vectorization techniques, dimensionality reduction, a variety of classification models, and ensemble methods are combined.

IV. RESULT ANALYSIS

Various traditional classification models were trained on the dataset and evaluated using K-Fold cross validation. The evaluation metrics used are Accuracy, Precision, Recall, and F1-Score. These models were used as base learners for meta models built using Stacking, Voting [20] and Boosting [22] techniques. The results of RoBERTa [16] and MathBERT [15] were compared against each other and that of MathBERT [15] showed better performance. The performance metrics values of RoBERTa embeddings are shown in Fig. 2.

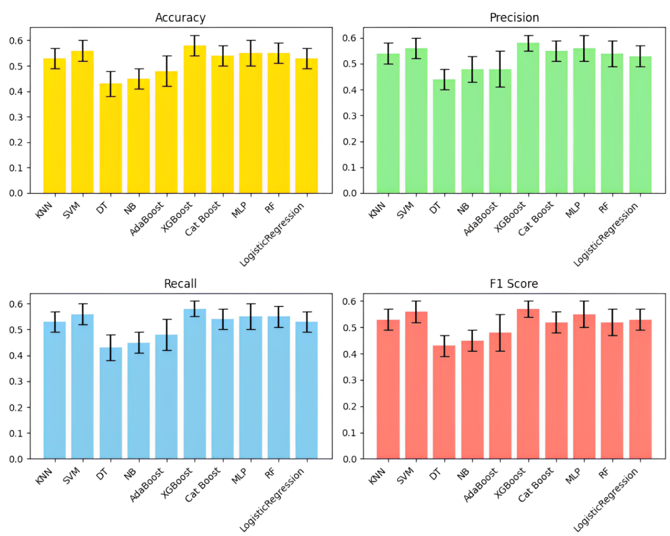

Fig. 2. Performance metrics values of RoBERTa embeddings

The results of MathBERT [15] embeddings are shown in Fig. 3.

A. Traditional Classification Models:

The text embeddings obtained from MathBERT [15] vectorization model showed better results in terms of evaluation metrics as compared to the RoBERTa [16] model as shown in Fig 4 .

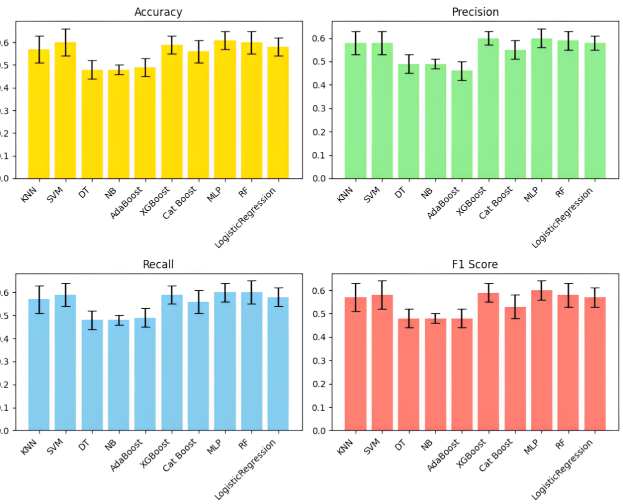

Fig. 3. Performance metrics values of MathBERT embeddings

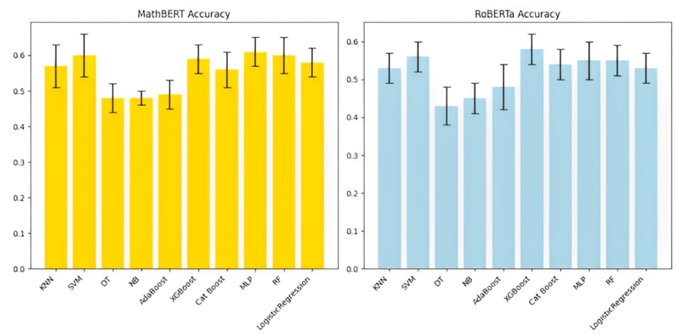

Fig. 4. Comparison of Roberta and Mathbert

B. Dimensionality Reduction:

As a result of applying Principal Component Analysis on both RoBERTa [16] embeddings and MathBERT [15] embeddings produced the following accuracy values shown in Fig 5 .

The values of accuracy, precision, recall and F1-Score of all the traditional models are illustrated in Table II.

TABLE II
PERFORMANCE OF INDIVIDUAL CLASSIFICATION MODELS

Base Learner	Accuracy	Precision	Recall	F1 Score
k-NN	0.57	0.58	0.57	0.57
SVM	0.60	0.58	0.59	0.58
Decision Tree	0.48	0.49	0.48	0.48
Naïve Bayes	0.48	0.49	0.48	0.48
ADA BOOST	0.49	0.46	0.49	0.48
XGBOOST	0.60	0.60	0.59	0.59
CAT BOOST	0.56	0.55	0.56	0.53
MLP	0.61	0.60	0.60	0.60
RF	0.60	0.59	0.60	0.58
Logistic Regression	0.58	0.58	0.58	0.57

C. Ensembled Model:

Various ensemble models were built using techniques such as Bagging [21], Boosting [22], Soft Voting, Hard Voting [20]

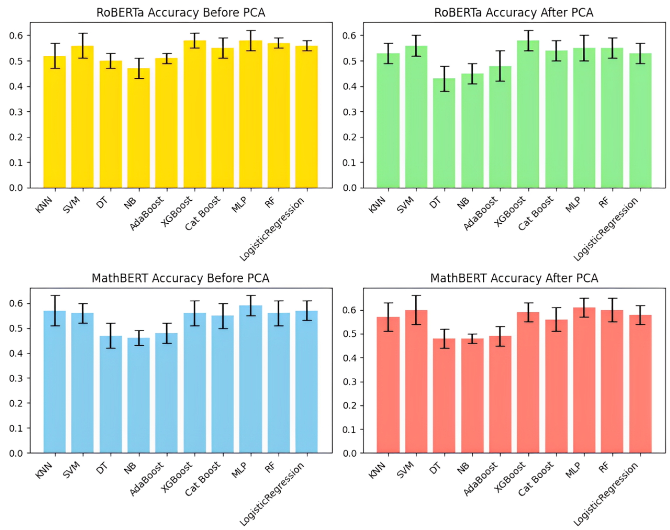

Fig. 5. Before and after applying PCA

and Stacking. The most promising results were produced by stacking SVM classifier and MLP classifier.
SVM and MLP were trained individually as base models. The predictions made by them was stacked and utilised to train the meta model. It showed an accuracy of approximately 65%. The metamodel shows an improvement over the existing models, whose performance is compared in Fig. 6

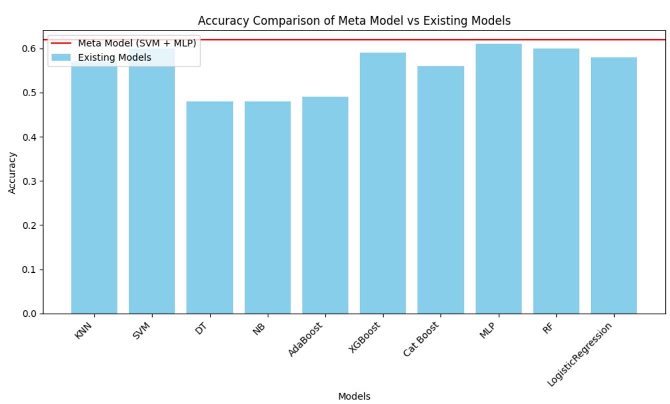

Fig. 6. Comparison of meta-model and the existing models

The confusion matrix shown in Fig. 7 reveals the class-wise distribution (Incorrect – Class 0, Partially Correct – Class 1, and Correct – Class 2) of the predictions. The model correctly classified solutions as Incorrect, Partially Correct, and Completely correct 112, 37, and 72 times, respectively. There are, however some misclassifications, mainly in the Partially Correct class, with confusion by the model of 23 instances as Incorrect and 30 cases as Correct. Moreover, 20 Correct solutions were misclassified as Partially Correct. Although these are the pitfalls, the overall performance metrics do reflect that the ensembled approach does harness the positives of both SVM and MLP classifiers, giving balanced performance across all classes. This model's strong performance illustrates the potential for practical automated assessment in educational settings.

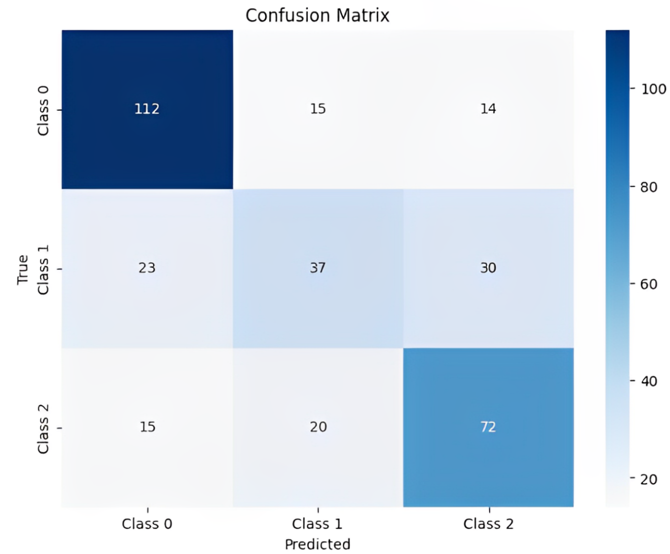

Fig. 7. Confusion Matrix of the Ensembled Model

V. Conclusion and Future Work

Upon analysing values of performance metrics, it is found that the classification models has better performance on the MathBERT [15] embeddings than on RoBERTa [16] embeddings. Categorising quadratic equation answers according to how right they are is a complex and important task. The significance of this categorization has been revealed by this literature analysis, not only in the context of mathematics education but also in practical applications where accuracy is crucial. We have studied numerous categorization approaches, evaluated examples of correct and erroneous results, and analysed the standards for accuracy. In addition to helping to improve problem-solving abilities, this categorization procedure also guarantees the accuracy of mathematical models and engineering designs. Quadratic equations may be learned and used more effectively, but there are obstacles to overcome, especially when adjusting to technology improvements.
As the future focus of this research, the following areas reveal promising opportunities:

- Techniques for Vectorization: Further research into other models for NLP could potentially reveal improvements in the accuracy and adaptability of the classification.
- Deployment in Edu-tech: Integrating this model into educational platforms that can be a facilitator for real-time feedback, making learning experiences much better.
- Adaptation to Other Mathematical Areas: Implementing this methodology in other mathematical topics could be a benchmark for standard and practical evaluation among various educational syllabi.

References

[1] Likforman-Sulem, L., Hanimyan, A. and Faure, C., 1995, August. A Hough based algorithm for extracting text lines in handwritten documents. In Proceedings of 3rd international conference on document analysis and recognition (Vol. 2, pp. 774-777). IEEE.

[2] Yin, F. and Liu, C.L., 2007, November. Handwritten text line extraction based on minimum spanning tree clustering. In 2007 international conference on wavelet analysis and pattern recognition (Vol. 3, pp. 1123-1128). IEEE.

[3] Varga, T. and Bunke, H., 2005, August. Tree structure for word extraction from handwritten text lines. In Eighth International Conference on Document Analysis and Recognition (ICDAR'05) (pp. 352-356). IEEE.

[4] Proceedings Volume 12174, International Conference on Internet of Things and Machine Learning (IoTML 2021); 121740X (2022) https://doi.org/10.1117/12.2628649 Event: International Conference on Internet of Things and Machine Learning (IoTML 2021), 2021, Shanghai, China

[5] Improving Automated Scoring of Student Open Responses in Mathematics-Sami Baral, Anthony F Botelho, John A Erickson, Priyanka Benachamardi, Neil T Heffernan.

[6] Li, Minli & Zhao, Peilin & Zhang, Yifan & Niu, Shuaicheng & Wu, Qingyao & Tan, Mingkui. (2021). Structure-aware Mathematical Expression Recognition with Sequence-Level Modeling. 5038-5046. 10.1145/3474085.3475578.

[7] A Robust Methodology for Creating Large Image Datasets Using a Universal Format Gurpartap Singh, S. Agrawal, B. S. Sohi

[8] Lan, A.S., Vats, D., Waters, A.E. and Baraniuk, R.G., 2015, March. Mathematical language processing: Automatic grading and feedback for open response mathematical questions. In Proceedings of the second (2015) ACM conference on learning@ scale (pp. 167-176).

[9] Zhang, M., Baral, S., Heffernan, N. and Lan, A., 2022. Automatic short math answer grading via in-context meta-learning. arXiv preprint arXiv:2205.15219.

[10] Ni, A., Inala, J.P., Wang, C., Polozov, A., Meek, C., Radev, D. and Gao, J., 2022, September. Learning math reasoning from self-sampled correct and partially-correct solutions. In The Eleventh International Conference on Learning Representations.

[11] Sami Baral, "Improving Automated Assessment and Feedback for Student Open-responses in Mathematics", in Proceedings of the 15th International Conference on Educational Data Mining, Durham, United Kingdom, Jul. 2022, pp. 795–798. doi: 10.5281/zenodo.6853036.

[12] Baral, S, Santhanam, A, Botelho, A, Gurung, A, and Heffernan, N. Automated Scoring of Image-based responses to Open-ended mathematics question

[13] Y. Tang, J. Dong, Y. Guo, Y. Zhou, F. Lu and B. Zhang, "A Malicious PDF File Detection Method Based on Improved Ensemble Learning Stacking," 2022 4th International Conference on Frontiers Technology of Information and Computer (ICFTIC), Qingdao, China, 2022, pp. 475-478, doi: 10.1109/ICFTIC57696.2022.10075332.

[14] Z. Li, D. Cai, J. Wang, J. Fu, L. Qin and D. Fu, "A Stacking Ensemble Learning Model for Mobile Traffic Prediction," 2020 IEEE/CIC International Conference on Communications in China (ICCC), Chongqing, China, 2020, pp. 542-547, doi: 10.1109/ICCC49849.2020.9238996.

[15] Peng, S., Yuan, K., Gao, L. and Tang, Z., 2021. Mathbert: A pre-trained model for mathematical formula understanding. arXiv preprint arXiv:2105.00377.

[16] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. and Stoyanov, V., 2019. Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.

[17] Xiao, T., Ren, D., Lei, S., Zhang, J. and Liu, X., 2014, June. Based on grid-search and PSO parameter optimization for Support Vector Machine. In Proceeding of the 11th World Congress on Intelligent Control and Automation (pp. 1529-1533). IEEE.

[18] Maćkiewicz, A. and Ratajczak, W., 1993. Principal components analysis (PCA). Computers & Geosciences, 19(3), pp.303-342.

[19] Mardaoui, D. and Garreau, D., 2021, March. An analysis of lime for text data. In International conference on artificial intelligence and statistics (pp. 3493-3501). PMLR.

[20] Burka, D., Puppe, C., Szepesváry, L. and Tasnádi, A., 2022. Voting: A machine learning approach. European Journal of Operational Research, 299(3), pp.1003-1017.

[21] Yu, N. and Haskins, T., 2021, July. Bagging machine learning algorithms: A generic computing framework based on machine-learning methods for regional rainfall forecasting in upstate New York. In Informatics (Vol. 8, No. 3, p. 47). MDPI.

[22] Schapire, R.E., 2003. The boosting approach to machine learning: An overview. Nonlinear estimation and classification, pp.149-171.

[23] Drăgulescu, B. and Bucos, M., 2020, October. Hyperparameter tuning using automated methods to improve models for predicting student success. In International Conference on Information and Software Technologies (pp. 309-320). Cham: Springer International Publishing.

[24] T. Palanisamy, Adithya Narayan Rajesh, greeshma Giri et al. Robust Regression for interval valued data, 07 September 2023, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-3277211/v1]

[25] T. Palanisamy, Adithya Narayan Rajesh, greeshma Giri et al. Robust Regression for interval valued data, 07 September 2023, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-3277211/v1]

[26] A. Nandakumar, P. P. Amritha, K. V. Lakshmy, and V. S. Talluri, "Non Linear Secret Sharing For Gray Scale Images," Procedia Engineering, vol. 30. Elsevier BV, pp. 945–952, 2012. doi: 10.1016/j.proeng.2012.01.949.

[27] R. M. Balakrishnan, P. B. Pati, R. P. Singh, S. S, and P. Kumar, "Fine-Tuned T5 For Auto-Grading Of Quadratic Equation Problems," Procedia Computer Science, vol. 235. Elsevier BV, pp. 2178–2186, 2024. doi: 10.1016/j.procs.2024.04.206.

[28] K. K S, A. Mahale and S. R. N, "Equation Detection in the Camera Captured Handwritten Document," 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023, pp. 1043-1049, doi: 10.1109/ICAAIC56838.2023.10141166. keywords: Training;Image segmentation;Recurrent neural networks;Text analysis;Symbols;Feature extraction;Cameras;Recurrent Neural Network (RNN);Support Vector Regression (SVR);Encoder-Decoder;Deep Learning,

[29] T. Palanisamy and J. Ravichandran, "Estimation of Variance Function in Heteroscedastic Regression Models by Generalized Coiflets," Communications in Statistics - Simulation and Computation, vol. 43, no. 10. Informa UK Limited, pp. 2213–2224, May 13, 2014. doi: 10.1080/03610918.2012.749283.