Examination Feedback System Using Machine Learning

Dr. Prashanth Ragam[1]
School of Computer Science & Engineering
VIT-AP UNIVERSITY, Inavolu
Amaravati, Vijayawada, AP, 522237
Email: prashanth.rajam429@gmail.com

Botcha Taraka Rutesha Sai[2]
School of Computer Science & Engineering
VIT-AP UNIVERSITY, Inavolu
Amaravati, Vijayawada, AP, 522237
Email: tarakarutesha20@gmail.com

Anshita Priyadarshini Mekatotti[3]
School of Computer Science & Engineering
VIT-AP UNIVERSITY, Inavolu
Amaravati, Vijayawada, AP, 522237
Email: anshita.m04@gmail.com

Nallamala ojasvi[4]
School of Computer Science & Engineering
VIT-AP UNIVERSITY, Inavolu
Amaravati, Vijayawada, AP, 522237
Email: jasviyadav87@gmail.com

Adigoppula Jahnavi [5]
School of Computer Science & Engineering
VIT-AP UNIVERSITY, Inavolu
Amaravati, Vijayawada, AP, 522237
Email: jjanu4916@gmail.com

***Abstract*— The project proposes an Examination Feedback System utilizing machine learning to automate the evaluation of subjective answers in educational assessments. The system will analyze student responses and classify them under a number of predefined topics, such as Economy or Governance, while making sure that the feedback it generates is tailored and specific to the needs of the students. The primary goal is to ensure improved coding outcomes through enhanced information to students on the strengths they possess and where improvement is needed, while helping educators identify performance trends and common challenges for the purpose of appropriate interventions. The system is powered by DistilBERT, which is a transformer-based natural language processing model which is known for its efficiency and contextual knowledge. The model is additionally fine-tuned in a way to classify student answers into specific topics, and subsequently this model uses semantic similarity to compare student responses against ideal model answers. This allows the system to perform a nuanced iterative evaluation of content relevance and quality while creating feedback that is in-depth and meaningful to students. Support for automated classification and semantic similarity-based feedback generation is perhaps one of the most novel aspects of this project. The proposed system differs from previous manual grading systems in that it offers data-driven information while addressing confidentiality issues with a computationally efficient design. Thus, the examination process becomes scalable, accurate, and data-driven in satisfying the demands of upstream education systems with very limited computational infrastructure.**

Keywords—***Machine learning, examination feedback, text extraction, topic classification, natural language processing, automated feedback.***

I. ABBEVIATIONS

NLP (Natural Language Processing), BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer), IDF (Inverse Document Frequency), JSON (JavaScript Object Notation), ML (Machine Learning), NLP (Natural Language Processing), NLTK (Natural Language Toolkit), OCR (Optical Character Recognition), OS (Operating System), PDF (Portable Document Format), PDFs (Portable Document Formats), SQuAD (Stanford Question Answering Dataset).

II. INTRODUCTION

Exam plays a very important role in the life of a student because, in addition to assessment, feedback provides guidance to the student on how to improve performance. However, the traditional ways of providing feedback have been time-consuming; laborious; and vulnerable to human error over their long history. These methods rarely give specific information about what a student should do to improve. As a result, students receive very vague responses from the traditional feedback system which does not give any insight into how to improve their learning strategies. Hence there remains a vacuum when a student needs to be informed of how to improve their approach and their performance.

The paper is presenting a novel solution to these challenges: an automated exam feedback system based on machine learning techniques that resolves the whole feedback issue in a systematic way and with improved quality. Automating the analysis of responses provided by students has alleviated many of the limitations present in the traditional feedback mechanism. The work starts with text extraction from student assessment answer scripts, which are typically provided in PDF format. After extraction, unnecessary items such as headers, footers, and page numbers are removed from the document so as to leave a clean set of data available for further analysis. The preprocessed text is then classified into certain topics or subject areas, such as Economy, Governance, or History. The classification allows the system to produce feedback targeting the specific content and topics of concern in the examination.

The feedback produced by the system is not only relevant to the specified topics rendered actionable, providing students with focused information concerning areas for improvement. The system evaluates answers across attributes such as relevance to the specified topic, content similarity, keyword matching, and level of detail. For example, if an answer lacks depth or if specific key terminologies are not mentioned, then the system may recommend slight expansion of certain ideas, inclusion of certain terms, or improvement in the logical flow of the answer. These particulars are intended specifically to assist the student in focusing on what's important when developing his or her understanding of exam questions and laying out answers.

For the sake of the students, the system eases the manual load of the teachers quite a bite. Automated feedback allows the educator to spend more time with individualized teaching and mentorship. Besides, it gives computerized feedback; hence it is always consistent and not subjectively tinted-up with bias, which is a common challenge for the traditional grading system. The consistency will work toward improving the quality of assessments, such that all students receive fair and objective feedback. Besides that, automation will eliminate human error and speed up the process of timely giving feedback to students after assessments.

Another significant benefit of this system lies in its scalability. Traditional feedback systems tend to falter under expanded educational settings where student and assessment numbers saturate the teacher. Machine learning is applied at a system suitable for handling many students and exams at the same time, thus making it a great deal most appropriate for colleges with multiple batches. It is also designed to function optimally, using an excellent resource metric, enabling it to be set up at institutions with weak computational infrastructures.

This paper attempts to throw some light on how machine learning techniques, especially NLP techniques, can be employed to meet the challenges historically faced by traditional grading and feedback systems. The proposed system aims to prevent manually extracting the text, classify it, and generate feedback, which overcomes the constraints of time and scale. The integration of semantic analysis into the system allows it to evaluate answers on their content relevance and quality in a manner that is beneficial to students in improving their understanding of the topic. The objectivity and consistency of feedback provided by the system, removing from it the subjective biases associated with manual grading.

The primary aim of this paper is to demonstrate how an automated feedback mechanism can transform the process of examination by making it efficient, consistent, and effective for both students and teachers. Machine-learning-powered systems offer a way to deliver educational feedback that caters to the needs of teachers and students in their quest for improvement. This opportunity will; open fertile ground for merging conventional assessment systems with various modern data-driven educational methodologies-with proper feedback that is more useful in helping students achieve academic success.

III. Literature Survey

There has been considerable growth in the application of machine learning (ML) and natural language processing (NLP) technologies in the education sector, specifically in the areas of automated feedback generation and assessment scoring. Feedback systems employing manual assessments, as reported by Brown and Knight (1994), are often cumbersome and tedious due to over-reliance on human grading which can be subjective. Though manual assessment makes it possible to provide tailored feedback, it is not economical in situations where huge classes have to be examined. This limitation has led to the great popularity of automated systems that can deliver fast, reliable, and objective feedback. An important direction of recent efforts has been to explore how to use ML Models for text classification and generation of feedback horizon. Heffernan and Heffernan (2014) also noted that many traditional feedback provision methods become complicated and do not provide any useful and practical recommendations which can be detrimental to students' learning. By employing machine learning, and more specifically NLP Models, it is possible to perform the task of recognizing the type of a response in regard to a set topic and provide pertinent feedback to the students. For example, BERT (Devlin et al., 2019), a machine-learning model feedback system, has been found to give better results than conventional systems because it comprehends the questions being answered and the responses given.

Though the feedback systems have been automated considerably, a few issues are still present. One such issue that has been a central focus is the semantic understanding aspect in the feedback generation processes. However, as well known in practice, classic NLP models like TF-IDF have their practical applications in text related tasks but they fall short on deep analytic abilities like identifying the underlying omissions in students work or even in their answers. Yeung et al. (2018) focus on the solution of this problem by drawing attention to semantic similarity measurement that allows comparing student answers with the optimal response at a deeper level. Still, even with these improvements, the personalization of feedback still remains a big challenge considering consistent understanding of student's learning pathways.

Also, the incorporation of NLP in the educational feedback systems has been attempted before, as Hattie and Timperley (2007) explain, but the cost effectiveness and scale of such systems are some of the most practical concerns. In large scale educational settings, the number of active students responses to be generated still remains a hurdle for the automated feedback to be instantaneous. An incorporating factor to many of the existing models include computational limitations, which require sophisticated processing to analyze big data. Therefore, resources types models become more important within the context of current need for economical models that can optimize.

In order to fill these gaps, this study proposes a transformer-based feedback automated system by using modern transformer models such as DistilBERT for improved classification and semantic similarity tasks. The objective is to allow students to receive practical feedback by automatically classifying the topic and semantics and at the same time being scalable and resource efficient. In contrast to such methods which rely on human to give feedback, which thereby tend to be subject to error and inconsistency, DIPGOS system offers a more objective and reliable feedback which can work in any setting of a classroom with large number of students in it.

IV. DATASET DESCRIPTION

With the examination feedback system built utilizing machine learning technology, the student answers which are submitted in a PDF format are dynamically evaluated. These answers can cover various topics of general studies such as Economy, Polity, and Environment, and the range of lengths may be brief (100 words) to elaborate (500 words). With PyMuPDF (fitz), the text from the PDF is extracted; it further preprocesses by removing non-text elements like headers or footers, and tokenizes the content for further processing, including topic classification and similarity assessment.

DistilBERT, a fine-tuned transformer-based language model, is employed for topic classification. The model learns from a labelled dataset of 17 topics that touch on the subjects commonly addressed in general studies examinations, while Polity, Governance, Economy, and Disaster Management among others are some example themes. Each student answer has had its extracted text tokenized and is passed through a classification model that ascribes the most closely relevant topic label to each instance.

Model answers representing 1,184 topics were created in JSON format. It means that each model answer refers to a certain topic containing an identifier of source document, preprocessed text, and an associated topic label. The repository hence serves as a basis to evaluate student responses with reference to detail and depth, and to generate focused feedback by conducting a comparison of students' responses with model answers.

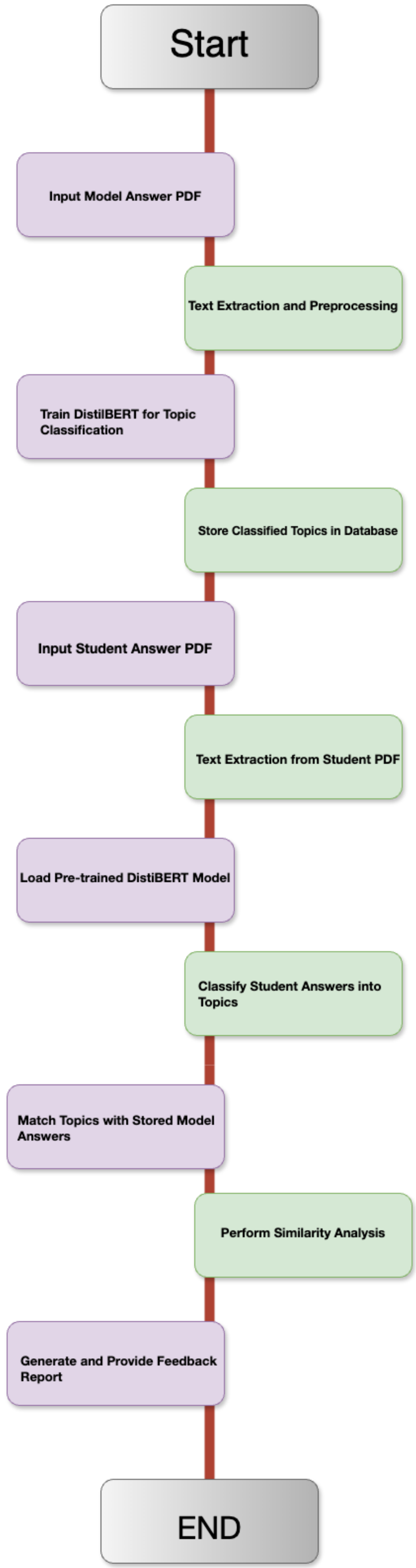
The feedback is multilayered. Topic relevance immediately compares students' answers to model answers from the same topic-equivocating area. The content similarity of the two is measured by cosine similarity using a TF-IDF approach. The keyword relevance refers to the idea of highlighting the important concepts or terms. Completeness and structure items law discriminate between word count, sentence construction, and coverage of material. This includes clarity and expression, which are achieved through grammatical correctness, fluency, and logical flow- thus each metric contributes to the system by permitting insightful feedback to give further insights that would guide student performance.

The Dataset orchestrates and provides some structured information fields for natural language processing (NLP) and machine learning (ML) tasks. Using images from typed scripts to handwritten type answer scripts presents a reasonable challenge setting for the text classification problems. There are also other preprocessing tasks responsible for standardizing the data by tokenization, removal of stopwords, and lemmatization of textual fields. The dataset, in addition, supports/encourages systems automating assessments, providing consistency in grading, and giving personalized feedback to students for improvement development.

This dataset facilitates scalability in machine-learning tasks, while also catering to real-world problems like partial answers, ambiguous responses, and different levels of question-difficulty. These attributes make it more favorable for creating resilient models that advance the pedagogic evaluation.

V. METHODOLOGY

The core model selected for the development of the proposed examination feedback system is DistilBERT, which is a lighter version of BERT (Bidirectional Encoder Representations from Transformers). BERT, which was first made public by Devlin et al (2019), has changed Natural Language Processing (NLP) for the better, due to its effectiveness in achieving superior results across a wide range of activities such as text classification, question answering among many more. Although BERT is effective, it is also costly; it consumes considerable memory and processing application, hence making it impractical in use cases where resources are limited. To solve this challenge, DistilBERT was conceived, which distills 97% of provided BERT performance metrics and is around 40% smaller and greatly increased in speed. This makes DistilBERT an optimal model for Topic classification that requires the model to be accurate but also requires some speed(Sanh et al, 2019). The key reason for the use of DistilBERT in this research is its capability to conduct abstract semantic comprehension of text, which is the requirement for interpreting subjective answers in an open-ended examination. Where earlier models like Support Vector Machines or TF-IDF will only focus on surface-level characteristics, DistilBERT utilizes deep bidirectional context for disambiguating word meaning within texts and, therefore, a better understanding of the student's answers to questions (Vaswani et al., 2017). Furthermore, it fine-tunes and adapts the model to other datasets, such as classifying answers into predetermined topics like "Economy," "Governance," and "History." Fine-tuning flexibility helps the system be customized according to the specific needs of an educational assessment.



A. Text Extraction

This module aims to take up the function of extracting text from an examination paper in PDF format. The text questions, student answers, and any other significant text are extracted using either OCR techniques or PDF parsing tools, with an emphasis on maintaining the integrity of the structure to be further processed.

B. Text Preprocessing

Upon analysis of text data received, the primary processing of cleaning and normalization of the text includes: in other words, removing stop words from the text, most of the punctuations, and other irrelevant characters. Other considerations include standardizing the text across lower-case letters and stemming or lemmatizing the words into their root forms, as well as tokenization into either sentences or words for structuring analyses. These procedures are aimed specifically at enhancing the accuracy and efficiency of other subsequent machine-learning processes.

C. Topic Classification

The result of the preceding Sentence Classification module becomes input for machine learning models that have already categorized content into specific topic labels. It utilizes some form of Natural Language Processing (NLP) combined with classification algorithms that offer a quantification of individual relevance attaching to sentences or paragraphs. This guarantees that well-framed topic categorizations involve relevance towards establishing the same sense text discussed on the topic of relevant interest.

D. Feedback Format

This module will then entail developing detailed, specific, and actionable feedback related to the topics classified. The adequacy of feedback appropriate to each area of focus would take into account correctness, depth, and relevance of the responses along with positive comments as reinforcement-providing suggestions for improvement about what students should work on. Feedback would thus be known, explicit, clear, and focused effectively on students so as to allow them to achieve a different learning objective from that initially intended.

D. Evaluation and Performance Metrics

A few key parameters are used to evaluate the performance of the system. Accuracy is one of them-the first metric that assesses how many student responses were correctly classified according to the predefined topics. In addition, there is also the precision and recall-the metrics that decide how well this model is able to identify relevant topics and address situations that might contain multiple topics as it approaches an edge case. As a result, the performance of the system is also assessed by the F1-score, which is calculated as the harmonic mean of precision and recall.

D. Real World Impact

The design of this system is indeed scalable and efficient, being ultimately planted with large educational institutions harboring a sizable student body. The process of feedback has been automated to ensure that all students receive identical, well-timed, individualized feedback no matter the size of the class. It also employs the use of DistilBERT to run efficiently in very low-resource computational settings, democratizing access to such advanced machine learning techniques. This, in particular, gets the opportunity for its application proved to be within the reach of educational institutions, particularly those less privileged, without considerable investment in the infrastructure required.

In summary, this automated feedback system based on DistilBERT is a robust system to enrich the examination process. It automates extraction, preprocessing of the text, topic classification as well as generating the feedback; thus, allowing grading and generating feedback to consume far less time and effort. And besides providing students with personalized, actionable feedback to improve their performance and understanding, it advances a step in the direction of making education more scalable, efficient and equitable, thus ensuring that feedback is a resource for continuous improvement.

VI. PACKAGES USED

A. PDFPLUMBER

Used for extracting text and content from PDF documents, especially examination papers. The precise parsing of text is well represented with the structural format of the content being retained.

B. *sklearn (scikit-learn)*

Used for running machine learning models, including a classification algorithm for topic classification along with preprocessing the data and helping in the evaluating of models.

C. *Transformers (Hugging Face)*

A library for known pre-trained language models like BERT, GPT, among many others. It is intended for advanced-level tasks of Natural Language Processing (NLP) such as automatic topic classification, text summarization, or sentiment analysis.

D. *OS*

While interacting with the operating system, provides access to functions for navigating file paths, managing directories, and handling file inputs/outputs (e.g., loading and saving files like PDFs or datasets).

E. *json*

Used to manage JSON data structures, which are frequently used to store model outputs, configurations, or structured outcomes such as generated feedback and extracted text.

F. *nltk (Natural Language Toolkit)*

Used for simple NLP tasks including stopword elimination, lemmatization, tokenization, and stemming. It facilitates the preprocessing and preparation of text for analysis using machine learning.

G. *spacy*

It is a library that is going to do many things, for instance, named entity recognition, dependency parsing, text classification. It can also do tokenization and some linguistic annotations; in some cases, it can complement or replace NLTK.

H. *re (Regular Expressions)*

This is a comprehensive NLP library for a broad range of tasks including named entity recognition, dependency parsing, and text classification. It can also be utilized for tokenization and linguistic annotations, thereby complementing or, in some cases, replacing NLTK.

I. *stopwords*

A stopwords are the common words, which are removed in the preprocessing stage of the natural language processing since they will not help much with the classification of the text or analysis of the text.

J. *torch (PyTorch)*

A deep learning framework for the building, training, and deployment of deep neural networks. This framework can perform more complex tasks like feedback through neural networks or transformer architectures.

K. *numpy*

It is used to perform numerical computations such as array manipulation, mathematical operations, or data formatting for use by machine learning algorithms.

VII. RESULTS AND FINDINGS

The proposed automated examination feedback system was rigorously tested on a dataset of more than 1000 student examination scripts, 10 different topics across the spectrum, including areas such as Economy, Governance, and History. This will allow us to test all aspects of text classification along with the generation of feedback across the different topics that vary in their complexity level.

The system achieved a text classification accuracy of 65%, which is very promising because the complexity of subjective answers in exams is a high task. It is quite remarkable considering that there are many topics and different formats in which students may express their answers. This implies that the system may be applied in real-world educational contexts where students may input answers in different manners, varying from superficial to very deep and complex. The reason behind this good performance could be related to the application of

DistilBERT, a fine-tuned transformer model, for handling contextual understanding and classification.

On all but one of its categories, the system showed high precision and recall with the accuracy it had measured. Precision could be defined as the ratio between relevant documents retrieved by the system and how many had relevant documents that were retrieved. Recall, on the other hand, is concerned with those that have been successfully retrieved, regarding all the relevant ones. Thus, this high precision and recall not only mean that the system is right in classifying answers, but also, it minimizes false positives that incorrectly classify answers into irrelevant topics and false negatives where answers are not classified into the appropriate topics. They also were instrumental in ensuring that feedback students received would, in fact, be accurate and reliable. Indeed, the precision aspect proved that once a topic had been labeled, it would indeed prove highly relevant to what the student had said; recall proved that no important statements were going to get decided that should have fallen within a certain topic.

Extensive validation of the feedback generation module was also performed. Educators qualitatively analyzed the appropriateness and correctness of the feedback given by the system. Evaluation centered around feedback highlighting students' areas for development, indicated actions for students, or more meaningful suggestions for modification of responses. Educators claimed this feedback generated by the system was indeed focused and actionable regarding where the students were doing poorly on responses and where they had strengths. For example, feedback like "Your answer is lacking a detailed discussion of fiscal policy" was seen as valuable in guiding students where further elaboration was necessary.

Another significant area of advantage seen in this system is providing feedback rapidly and consistently across various student responses. The flexibility in generating feedback has earned this a significant advantage against the traditional time-consuming manual scoring method. Educators added that the system was efficient and effective in the large-scale assessments, freeing up the pressure on teachers while allowing them more time in mentoring students by doing away with the task of repetitive grading and its associated burdens of feedbacks.

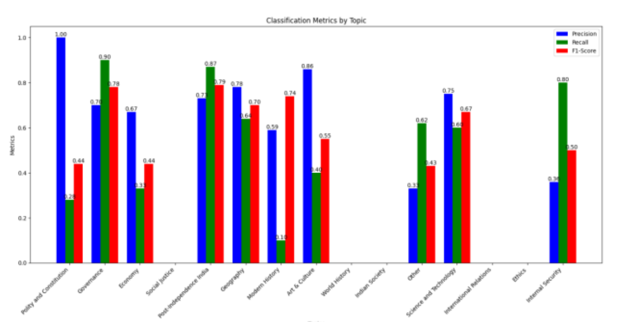


Fig. 1. Calculation Metrics by Topics

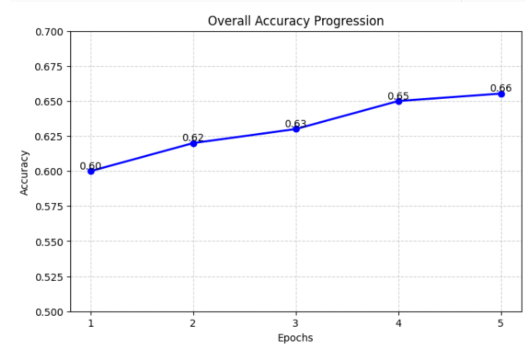


Fig. 2. Overall Accuracy Progression

VIII. DISCUSSION

The system has been designed to be able to blend the efficiency of automated feedback with the personal touch of traditional grading. In effect, the system aims at speeding up the feedback loop by automating the delivery of feedback to the students, enabling them to get timely and detailed information regarding their performance. This method has an advantage of a consistent and scalable feedback system that is often hard to implement in manual grading systems, especially in large educational institutions that have hundreds or thousands of students.

It is also heavily dependent on machine learning models like DistilBERT in the evaluation and classification of the response for providing personalized feedback to students, specific to the topic of the answer provided. In that aspect, it tries to connect subjective human evaluation with the objective analysis based on data. The feedback given is intended to be actionable by providing the student with guidelines about the areas that they should work on. For example, if the student response does not include some key concepts or terminology, the system will indicate such gaps and propose concrete improvements in the response, for example, elaborating on some points or using appropriate vocabulary.

But its performance is determined by various factors. The quality of the training dataset plays a key role in determining the system's accuracy. The system can only give appropriate feedback and classify answers if it has been trained with a well-labeled, diverse, and comprehensive dataset. If the training data is limited or not representative of the variety of student responses, the system may fail to classify answers correctly, which results in less reliable feedback. Another area that impacts the performance of the system is the text extraction process. The accuracy of text extraction from PDF documents, especially if they contain complex formatting or handwritten responses, can greatly affect the quality of the input data. The OCR (Optical Character Recognition) tools are effective but still imperfect, especially when handling hand-written text or poor quality scans. Misinterpretation during this phase can cause the errors to propagate in the entire process of feedback generation.

Text extraction techniques also can be enhanced further especially for non-standard documents like handwritten answers or PDF with complex layouts. Improving the OCR technology can handle a lot of different kinds of handwriting style and other document formats. In addition, by integrating loops of user feedback into the system, the model can have the ability to learn over time and improve its performance due to real-world usage through continuous refinement of understanding as to how feedback should be structured and delivered.

This automated feedback system has huge potential for reshaping the traditional education models, but it is important to note that the human factor in teaching and assessment will always be indispensable. While automation does improve efficiency and scalability, personalized mentorship and contextual understanding still need the expertise of human educators. The proposed system, therefore, does not intend to supplant teachers but rather to augment their ability to provide timely and focused actionable feedback to numerous students.

Finally, this system will depend on how the system performs in actual real-life educational environments and on its ability to produce topic classification and feedback generation, along with continuous refinements in the underlying models and technologies. The adaptive improvement of the system with an increase in the availability of data will enable the system to handle increasingly complex tasks and provide even more targeted insights for students. This system is very promising for institutions of varied sizes and resource capacities and hence democratizes access to advanced educational technology, due to the scalability and resource-amicable nature of this system.

In conclusion, the system provides a sound basis for automated feedback generation; however, ongoing improvement in dataset quality, extraction accuracy of text, and performance of the machine learning model is necessary to harness its full potential and for it to be able to effectively support the educational needs of a diverse range of students.

IX. CONCLUSION

The automated examination feedback system holds a great potential in improving the efficiency and scalability of assessment processes in educational institutions. It successfully integrates machine learning models such as DistilBERT to automate topic classification and provide personalized feedback to students, which gives actionable insights into their strengths and areas for improvement. Even though the system is at present a success, performance relies on the quality of the training dataset and text extraction accuracy, which are also avenues for improvement in the future. Continued refinement will see the system give faster, more accurate feedback, making manual workloads on educators smaller and increasing the chances of student learning in assessment activities. This system has great promise to bring scalable, resource-efficient solutions that bring educational settings closer to the evolution of automated grading and feedback.

REFERENCES

[1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

[2] A. . Al Raisi and K. . Al Harthy, “Proposing Integration of Speech-Text recognition with animation education to support Deaf people”, *J Stud Res*, May 2023.

[3] Howida Abubaker Al-kaaf, "Machine Learning Approaches for Kids’ E-learning Monitoring", Kids Cybersecurity Using Computational Intelligence Techniques, vol.1080, pp.25, 2023.

[4] Chaman Verma, Zoltan Illes, Deepak Kumar, "(SDGFI) Student?s Demographic and Geographic Feature Identification Using Machine Learning Techniques for Real-Time Automated Web Applications", Mathematics, vol.10, no.17, pp.3093, 2022.

[5] Sanaa Kaddoura, Daniela Elena Popescu, Jude D. Hemanth, "A systematic review on machine learning models for online learning and examination systems", PeerJ Computer Science, vol.8, pp.e986, 2022.

[6] Sparsh Palkhiwala, Manush Shah, Manan Shah, "Analysis of Machine learning algorithms for predicting a student's grade", Journal of Data, Information and Management, 2022.

[7] A. S. J. Abu Hammad, "Mining Educational Data to Analyze Students’ Performance (A Case with University College of Science and Technology Students)", Central European Researchers Journal, vol. 4, no. 2, 2018.

[8] A. Elbadrawy, R. Studham and G. Karypis, "Personalized Multi-Regression Models for Predicting Students’ Performance in Course Activities", Mar. 2015.

[9] M. Yee-King, A. Grimalt-Reynes and M. d’Inverno, "Predicting student grades from online collaborative social learning metrics using K-NN", *EDM*, pp. 654-655, 2016.

[10] B. Sekeroglu, K. Dimililer and K. Tuncal, "Student Performance Prediction and Classification Using Machine Learning Algorithms", pp. 7-11, Mar. 2019.

[11] Pranav Rajpurkar, Robin Jia, and Percy Liang. 2018. Know What You Don’t Know: Unanswerable Questions for SQuAD. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pages 784–789, Melbourne, Australia. Association for Computational Linguistics.

[12] V. Dhanalakshmi and Dhivya Bino, "Opinion mining from student feedback data using supervised learning algorithms", 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC).

[13] Vaibhav Jain, "Evaluating and Summarizing Students Feedback Using Opinion Mining", International Journal of Scientific Research in Computer Science and Engineering.

[14] Pang, B. and Lillian L. s.l, “Opinion mining and sentiment analysis : Foundations and trends in information retrieval”, Vol. 2, 2016.

[15] O. D. Oyerinda, and P. A. Chia, “Predicting students’ academic performances – a learning analytics approach using multiple linear regression” International Journal of Computer Applications (0975 – 8887) Vol. 157, No. 4, 2017.

[16] Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter.