

Study Support and Feedback System Using Natural Language Processing

Chandrasiri K.T.H.S.¹, Kumarasinghe S.², Umayangie M.K.A.³, Gunathilaka U.K.R.S.W.⁴, Anjalie Gamage⁵, Dhammika H. De Silva⁶

¹Undergraduate, Faculty of Computing, Sri Lanka Institute of Information Technology, Colombo, Sri Lanka

²Undergraduate, Faculty of Computing, Sri Lanka Institute of Information Technology, Colombo, Sri Lanka

³Undergraduate, Faculty of Computing, Sri Lanka Institute of Information Technology, Colombo, Sri Lanka

⁴Undergraduate, Faculty of Computing, Sri Lanka Institute of Information Technology, Colombo, Sri Lanka

⁵Senior Lecturer, Faculty of Computing, Sri Lanka Institute of Information Technology, Colombo, Sri Lanka

⁶Senior Lecturer, Faculty of Computing, Sri Lanka Institute of Information Technology, Colombo, Sri Lanka

Abstract - Study Support and Feedback System Using Natural Language Processing (NLP) aims to provide an abstraction layer for a system that assists in studying and providing feedback to users through the application of NLP techniques. The system utilizes NLP algorithms and methodologies to analyze user input in natural language, allowing for personalized study support and feedback. The abstraction encompasses receiving natural language input from users, which can include structured and essay questions targeting feedback on study materials. NLP techniques are applied to preprocess, tokenize, and parse the input, extracting relevant keyword information and determining the user's intent. NLP algorithms focus on providing a highly accurate score with constructive feedback which enlighten the students to reflect on their progress with time. The system has a significant impact to generate tailored responses to user queries or feedback, enabling better learning. Coherent and informative natural language responses are provided in a clear and concise manner which eases the work of teachers and lecturers. Students gain the opportunity to seek on areas which need improvement, and teachers can help them to have a remedy through the designed system, which generates an efficient and wholesome study environment with a positive impact for both students and teachers. Users can productively interact with the study support and feedback system through a convenient and user-friendly interface. By abstracting into these components, the Study Support and Feedback System Using NLP aims to enhance the studying experience, provide personalized support, deliver effective feedback to users and track progress of users effectively.

Key Words: Natural Language Processing, NLTK, Study Support and Feedback

1.INTRODUCTION

The current educational landscape has witnessed the rapid integration of technology, with a particular focus on enhancing learning experiences and delivering personalized support to students. In this context, Natural Language Processing (NLP), a prominent subfield of artificial

intelligence, has emerged as a powerful tool for comprehending and processing human language. This research aims to propose the development of a Study Support and Feedback System that utilizes NLP techniques to revolutionize the manner in which students receive assistance and feedback during their learning journey.

The Study Support and Feedback System will leverage NLP techniques to enable the analysis and comprehension of natural language input from students. By extracting the meaning and intent behind their queries, tailored study materials, explanations, and feedback can be provided, thereby adapting effectively to each student's unique needs and requirements. The primary objective of this research is the design and implementation of an abstraction layer for the Study Support and Feedback System, thereby establishing an intelligent and interactive platform for students to engage with educational content. Through the utilization of NLP, students will be empowered to seek guidance, clarification, and feedback by posing natural language queries, ultimately enhancing their understanding and knowledge retention [1].

Online education and exams have become increasingly popular due to their benefits, such as automatic grading, immediate feedback, and a reduction in administrative work. However, there are difficulties with taking exams online, particularly when it comes to evaluating subjective questions like essays and structured questions [2]. To provide more accurate and efficient grading of subjective questions, this study will examine the application of Natural Language Processing (NLP) techniques in online examination systems.

A suggested Study Help and Feedback System is also discussed [3]. The Study Support and Feedback System is a web-based system that uses NLP methods to enhance the caliber of the feedback given to students in online testing environments. The feedback module gives students specific feedback. The system is made to be malleable and versatile so that it may incorporate new NLP methods as they become available.

It is important to address the issues faced by students and teachers in the evaluation of student responses, particularly in practice tests and model questions [4]. The inability to obtain correct answers and insufficient analysis of incorrect responses can have a negative impact on students' grades and motivation, affecting their prospects. Manual evaluation of student responses is a time-consuming process that takes away teachers' ability to provide context-specific feedback and constructive criticism. Without timely intervention, these issues may have a lasting impact on students' academic and professional success. Therefore, it is essential to develop a better and more robust study environment that addresses these challenges.

Students who are motivated to test their skills by attempting past paper questions and model questions often face demotivation due to not being able to receive correct answers and proper analysis. This suffering can lead to lower grades and poor results, which could have been vehemently avoided if necessary steps were taken on time. On the other hand, teachers and lecturers face hardships when assessing the student answers manually, which is a tedious task [5]. These inefficient practices prevent them from providing their best and utilizing their maximum potential to create a better learning environment. Giving students timely and useful feedback is one of the most significant challenges confronting teachers [5].

Feedback is an important part of the learning process but can be time-consuming for teachers. Natural language processing and Machine Learning (NLP and ML) can be used to evaluate student writing and pinpoint weaknesses. Existing systems such as "netexam.sliit.lk" are only able to mark multiple choice questions. The proposed system will additionally allow students to receive personalized feedback on their academic performance and obtain recommendations for further study. NLP-based feedback systems can help teachers analyze student improvement following feedback and detect patterns and trends in students' progress.

2. LITERATURE REVIEW

Online exam question correction systems offer a quick and precise way to grade exams and give students feedback. The advantages of these systems include improved exam grading effectiveness, quick feedback to students, and more in-depth exam result analysis. Computer-assisted testing and automated grading systems can improve student performance and learning outcomes. Machine learning algorithms have also been used in online examination question correction systems to produce precise and trustworthy grading outcomes. A study has found that machine learning algorithms produced precise and trustworthy results for grading, lowering the effort of teachers and enhancing the grade of feedback given to students. It also found that the use of a self-adaptive question difficulty algorithm in an online exam system

increased grading system accuracy and lowered instructors' workloads [6]. These systems offer advantages to both students and instructors, such as improved exam grading effectiveness, quick feedback to students, and more in-depth exam result analysis.

Natural Language Processing (NLP) has advanced significantly in recent years, enabling the creation of intelligent computers that can comprehend and analyze input in natural language. As a result, support and feedback systems have been developed to aid students in their learning process by offering individualized help, feedback, and direction [7]. Multiple Choice Questions (MCQs) have been the primary focus of traditional assistance and feedback systems, while Structured Essay Questions (SEQs) have become increasingly well-liked due to their capacity to evaluate higher-order thinking abilities. NLP algorithms can analyze and grade these answers, giving both students and teachers useful feedback. Key phrase extraction and ontology mapping are used to identify the most relevant words or ideas in a text and offer individualized feedback [8].

The development of a support and feedback system for SEQs that makes use of key phrase extraction, NLP, and ML algorithms, as well as ontology mapping, has the potential to completely alter the way instructors evaluate and give feedback on students' learning. NLP algorithms, key phrase extraction, and ontology mapping have the potential to revolutionize the field of education by enhancing the process of SEQ evaluation and feedback. This system can analyze and assess student responses to SEQs with accuracy and provide individualized and substantive feedback. It can help to overcome the drawbacks of conventional support and feedback systems, which often rely on MCQs [9]. Additionally, it can be used to monitor student development over time and spot areas where they might require more assistance or resources. This approach has the potential to raise educational standards overall and improve student learning results.

Ontology creation is the process of developing an ontology for a particular domain, which involves identifying the relevant concepts, relationships, and constraints and defining them using formal language. There are various approaches to ontology creation, such as top-down approaches, bottom-up approaches, and ontology learning approaches. Top-down approaches involve refining a high-level ontology for a specific domain, while bottom-up approaches involve building an ontology from scratch based on domain-specific data and knowledge. Several approaches have been proposed, such as ontology learning and ontology merging, each with its strengths and limitations. Evaluation of ontologies is essential to ensure their quality and usefulness. Several metrics have been proposed for ontology evaluation, such as completeness, consistency, and coherence.

In recent years, NLP has gained popularity in the educational industry, particularly in the development of automated test scoring systems. Automated grading systems that assess essay type questions using NLP methods were also required in the study sector. The use of Natural Language Processing (NLP) for creating online question-correcting systems for exams has the potential to greatly enhance the efficacy and efficiency of the educational system. An NLP-based system has been created for assessing computer science short answers, which assessed the accuracy of the replies using syntactic and semantic similarities [10].

However, the system's effectiveness as a teaching tool was constrained since it did not give students feedback on their errors. Research [11] suggested an automated grading system that assesses essay-type questions using NLP methods. The method recognized several characteristics of an essay, including coherence, grammar, and substance. The essays were assessed by algorithm using word embeddings, part-of-speech tagging, and sentence similarity.

Azhar and Ullah's investigation highlights the significance of the exercise of NLP strategies to the assessment of mathematics test problems [12]. The system uses machine learning methods to predict the grades of new questions after being trained on a dataset of manually graded mathematics problems [13]. The outcomes demonstrated that the amount of the training dataset had an impact on the system's accuracy.

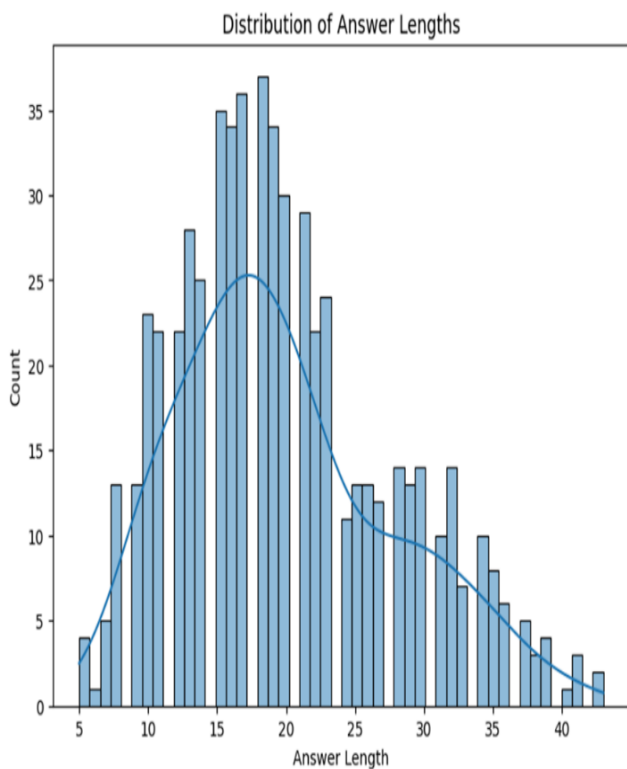


Fig -1: Distribution of Answer Lengths

The use of NLP in automated grading systems has been investigated in several studies. A method based on NLP was created by Ramana and Karthik [14] to automatically score brief responses for programming problems. The algorithm evaluated the students' responses based on how closely they matched the right answers using a deep learning model. The outcomes of several research demonstrate that the system's precision was on par with that of human graders [15].

NLP has been used in education and assessment in a number of studies, such as automated formative feedback in higher education [16], natural language processing for assessment in higher education [17] and enhancing automated assessment of open-ended questions through NLP. These studies show how NLP may enhance the precision and effectiveness of grading and feedback in academic contexts [18]. Question correction applications for online exams have several advantages for both students and teachers [19]. The capacity of machine learning algorithms to produce precise and trustworthy grading outcomes has led to an increase in their use in online examination question correction systems [20].

3. METHODOLOGY

A Google form was used to collect data for these tests.

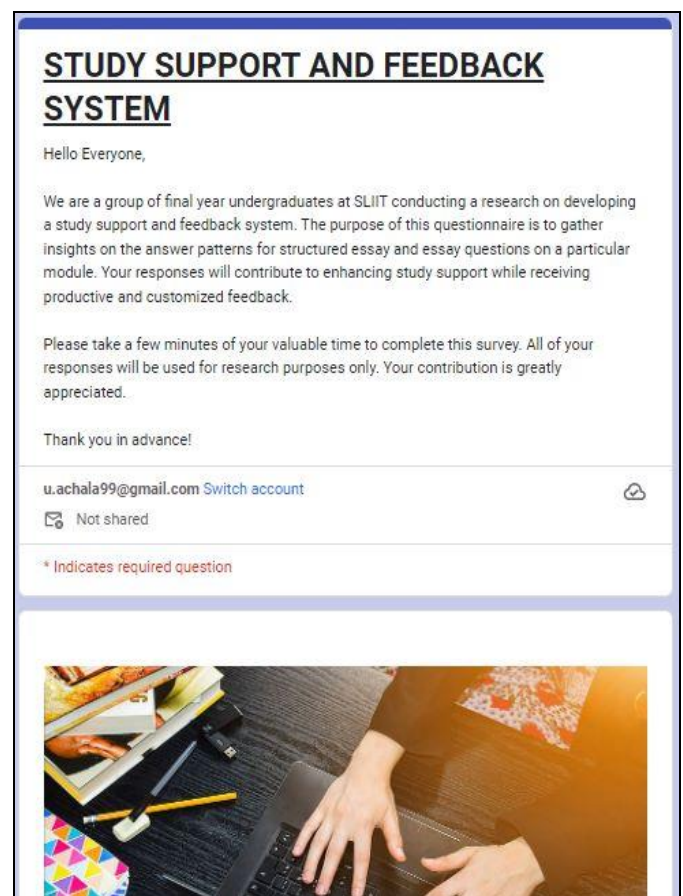


Fig -2: Created Google Form

The initial phase entails gathering a sizable and varied dataset of graded open-ended questions from prior online tests. In order to train machine learning algorithms, this dataset will be used. To ensure uniformity in the grading standards and to remove any redundant or extraneous information, the obtained dataset will then be preprocessed. The preprocessed dataset would then be used to extract the pertinent features. These features, which can be utilized to train machine learning algorithms, may comprise linguistic, semantic, and syntactic elements. Selection of the most appropriate machine learning algorithm for grading open-ended questions in online exams will include evaluating and comparing several different methods. Using the retrieved features from the preprocessed dataset, the chosen machine learning method would be trained. After that, the model would be optimized by changing its parameters in order to increase its precision and dependability. By contrasting the developed machine learning model's grading accuracy with that of human graders, its performance would be assessed. To make sure that the model can generalize to new queries, evaluation would be conducted on a separate test dataset. The resulting machine learning model would next be implemented into an online examination system to automatically grade open-ended questions.

The proposed machine learning-based online examination question-correcting system would be put to the test in a real-world online examination environment to confirm its usefulness. The system's correctness, dependability, and efficiency in grading open-ended questions will be evaluated. Future work and improvements: Based on the analysis of the data, it will be possible to determine what needs to be done in the future to improve the system's functionality and look into potential new applications. Data collection, preprocessing, feature extraction, machine learning algorithm selection, model training and optimization, model validation, system development and integration, system validation and testing, result analysis, and future work would all be included in the methodology for creating an online examination question-correcting system using machine learning.

The software for the study support and feedback system will include NLP techniques, ML algorithms, keyword extraction, and ontology mapping. It will be created using Python and other NLP libraries and will provide an intuitive interface for students to contribute essays and display criticism. The system will review the student's text input and performance data to give feedback and recommendations for further study.

This project intends to provide a study assistance and feedback system for SEQs utilizing NLP methods such as keyword extraction, ontology mapping, and ML algorithms. There will be numerous stages in this study's technique. At the Data Collection and Pre-processing phase, the project will gather a dataset of SEQs from various educational institutions, such as universities or high schools. Also, the dataset will go through pre-processing to get rid of any

duplicate or irrelevant replies and get it ready for analysis. Ontology Creation to extract important ideas and words from students' answers to SEQs, the study will create an ontology-based model. The ontology will be developed utilizing domain-specific knowledge sources and will be intended to extract the SEQs' most pertinent terms and concepts. At the Machine Learning Algorithm Development stage, the study will create an algorithm to assess student replies using the ontology-based model and deliver precise feedback based on the discovered ideas and words. ML methods like supervised learning will be used to train the algorithm on the pre-processed dataset of SEQs. In order to accurately and efficiently assess SEQs and give students individualized feedback, the research will assess how well the system was created. Using a sample of student users, the system's usability and efficacy will also be assessed. The data gathered during the assessment phase will be analyzed by the research using a variety of approaches. Metrics including precision, recall, and F1 score will be used to measure the system's efficacy and accuracy in analyzing SEQs. Metrics including ease of use, user happiness, and engagement will be used to assess the usability and efficiency of the user interface.

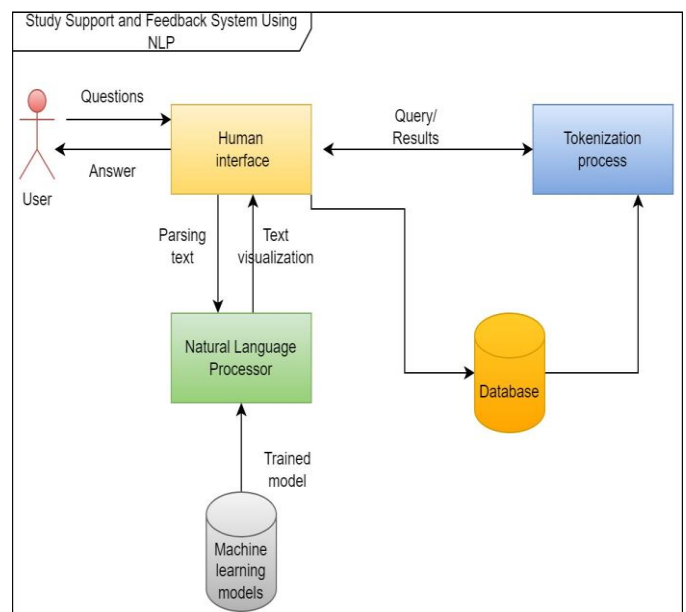


Fig -3: System Overview Diagram

Ontology creation is a critical process in knowledge engineering that aims to model the domain knowledge in a structured manner. This aims to outline a methodology for ontology creation using NLP modules. The proposed methodology consists of four stages namely data collection, data preprocessing, ontology creation, and ontology evaluation. In the first stage, data collection, a comprehensive dataset of domain-specific text is collected from various sources, such as academic papers, research reports, and online databases. The data collection process should be conducted carefully to ensure that the dataset is representative of the domain knowledge. In the second stage,

Ethical considerations were also critically emphasized in the research. The research on the study support and feedback system using NLP has showcased the potential of NLP techniques in revolutionizing education. The system's ability to provide personalized support, timely feedback, and efficiently track student progress has brought forward its accessibility to contribute to an enhanced learning experience. As further advancements are made and ethical considerations are prioritized, such systems have the potential to reshape education by providing effective and tailored support to students, ultimately improving their academic success and overall learning outcomes.

ACKNOWLEDGEMENT

We would like to pay our immense gratitude to Ms. Anjalie Gamage, who was our research supervisor and Mr. Dhammika H. De Silva, who was our research co-supervisor. Our heartfelt gratitude extends to our parents and colleagues who supported us significantly during the research project. We also appreciate the utmost support provided by our workplaces to handle the research effectively.

REFERENCES

- [1] Q. P. S. BS, "Natural Language Processing and Assessment of Resident Feedback Quality," 2021.
- [2] M. Khanbhai et al., "Applying natural language processing and machine learning techniques to patient experience feedback: A systematic review," *BMJ Health Care Informatics*, vol. 28, no. 1, 2021.
- [3] T. B. Shaik, "A Review of the Trends and Challenges in Adopting Natural Language Processing Methods for Education Feedback Analysis," 2022.
- [4] H. Qudaih, "Natural Language Processing in Customer Service: A Systematic Review," 2022.
- [5] Z. Kastrati, "Sentiment Analysis of Students' Feedback with NLP and Deep Learning: A Systematic Mapping Study," 2021.
- [6] A. Moharil, "Integrated Feedback Analysis And Moderation Platform Using Natural Language Processing," 2020.
- [7] F. K. Khaizer, "Sentiment Analysis of Students' Feedback on Institutional Facilities Using Text-Based Classification and Natural Language Processing (NLP)," 2023.
- [8] E. Ötleş, "Using Natural Language Processing to Automatically Assess Feedback Quality: Findings From Three Surgical Residencies," 2021.
- [9] L. Zhao, "Natural Language Processing (NLP) for Requirements Engineering: A Systematic Mapping Study," 2020.
- [10] S. J. J, "A prior case study of natural language processing on different domain," 2020.
- [11] S. Gupta, A. Kumar, and A. Prasad, "Automated grading system for essay-type questions using natural language processing," in *Proceedings of the IEEE International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, July 2020, pp. 639-642.
- [12] M. F. Azhar and R. Ullah, "Evaluation of mathematics exam questions using machine learning and natural language processing," in *Proceedings of the 6th International Conference on Advances in Computing and Data Sciences (ICACDS)*, Coimbatore, India, April 2021, pp. 31-38.
- [13] Y. Qiang, J. Xu, L. Li, X. Li, and J. Li, "Automatic feedback for English writing based on NLP technology," in *Proceedings of the 2020 9th International Conference on Educational and Information Technology (ICEIT)*, Tokyo, Japan, July 2020, pp. 264-268.
- [14] G. V. Ramana and M. Karthik, "Automated grading of programming questions using natural language processing," in *Proceedings of the IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, Canada, October 2021, pp. 1148-1153.
- [15] K. Huang, Z. Ren, H. Xiong, Y. Zhou, and Z. Hu, "Using natural language processing to provide automated feedback on student writing: A literature review," *British Journal of Educational Technology*, vol. 51, no. 5, pp. 1695-1710, 2020.
- [16] Nulty, D. Olson, and S. Chaudhuri, "Exploring the potential of natural language processing for automated formative feedback in higher education," *British Journal of Educational Technology*, vol. 52, no. 2, pp. 875-892, Mar. 2021.
- [17] V. Vukovic and C. Weir, "Can natural language processing improve assessment in higher education?" *Educational Research Review*, vol. 28, pp. 100288, Dec. 2019.
- [18] D. J. Nicol and D. Macfarlane-Dick, "Formative assessment and self-regulated learning: A model and seven principles of good feedback practice," *Studies in Higher Education*, vol. 31, no. 2, pp. 199-218, 2021.
- [19] A. S. Babatunde, O. J. Oyelade, O. O. Oluwagbemi and F. O. Akande, "Automatic grading system for programming assignments," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 12, pp. 25-35, 2021.

