

Automatic Question And Answer Generation By Using NLP And Flask Framework

S. Saravanan ¹, Dr.Bhuvaneshwari.M ²

Department of Computer Applications

^{1,2}PG Student, Dr. M.G.R. Educational and Research Institute, Chennai – 95

³Associate Professor, Dr. M.G.R. Educational and Research Institute, Chennai – 95

Abstract- *In the evolving landscape of Natural Language Processing (NLP), the demand for automated systems capable of generating relevant questions from textual data has grown substantially. This research introduces an innovative approach leveraging advanced NLP techniques combined with the Flask framework to develop a dynamic question and answer generation system. The proposed system demonstrates promising results in generating coherent and contextually relevant questions based on the input text, thereby enhancing the accessibility and comprehension of textual content.*

Keywords- Automatic Question Generation, Natural Language Processing, Flask Framework, Text Analysis, Question Relevance, Contextual Understanding.

I. INTRODUCTION

Demand for intelligent systems capable of understanding and generating human-like language has surged. Natural Language Processing (NLP) has emerged as a pivotal technology, empowering machines to comprehend, interpret, and generate human language. One intriguing application of NLP is automatic question and answer generation, where machines analyze textual content and generate relevant questions along with corresponding answers [1].

To leverage the power of NLP alongside the Flask framework to develop a system capable of automatically generating questions and answers from textual inputs. By harnessing state-of-the-art NLP techniques, our system will analyze text passages, identify key concepts, and formulate meaningful questions. Additionally, it will generate accurate answers based on the content of the input text [2].

Through this system, we intend to demonstrate the potential of combining NLP and web development frameworks like Flask to create intelligent applications that enhance human-computer interaction. By enabling users to input text and receive automatically generated questions and answers, this system can find utility in various domains, including education, content creation, and information retrieval [3].

An automatic query generator using natural language processing (NLP) generates meaningful, syntactically and semantically accurate queries based on various input formats such as text, structured database or databases. NLP is a subfield of artificial intelligence (AI) that aims to facilitate human-computer interaction using natural language that recognizes and understands spoken and written human language. The ultimate goal of NLP is to read, analyze, understand and comprehend written or textual information in natural languages in a meaningful way. This system aims to use NLP to set questions for a computer-based exam for teachers and students preparing for competitive exams[4].

As a result, you ensure that the MCQ contained appropriate questions and options. These questions are consistent with the learning objectives and meaning of the topics covered in the study material. The automated question generator can be used in many domains, including massive open online courses, objective surveys, search engines, automated help systems, chatbots (eg for customer interaction), and treatment for mental health analysis. However, it takes time and effort to manually create meaningful and relevant questions [5].

II.LITERATURE SURVEY

According to N Hedberg.et al., 2020 The amount of information available and consumed by people around the world is increasing. To reduce mental fatigue and increase understanding of general complex texts or documents, we developed an application for this task. The application allows users to download documents and ask domain-specific questions about them using our website. A condensed version of each document is presented to the user, which can further facilitate understanding[6].

According to Deeksha.et al., 2021 world where technology has grown into a mega-race that increasingly affects the lives of humanity, the common man still has great difficulty in immediately adapting existing technological solutions. There are so many basic, laborious and time-consuming tasks in people's daily lives that can be automated through technology[7].

According to **Wenyan Yang**.et al., 2022 Based on natural language processing and methodology developed in Python Web's Flask framework, this paper completes the design and implementation of an automated grade evaluation system. To complete the similarity analysis, essential algorithms such as lexical analysis, syntactic dependency analysis and information extraction using the natural language processing technology of the system are mentioned[8].

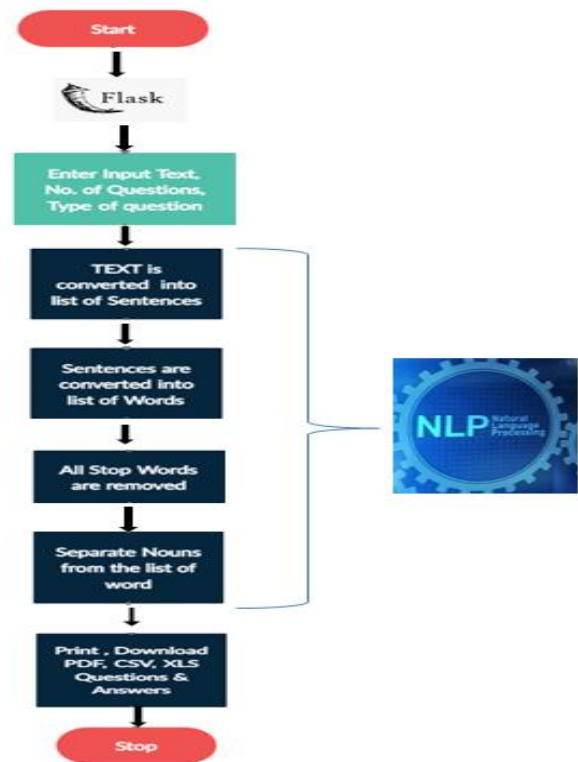
According to **Akanksha Sharma**. et al., 2023 As the world is digitized and changing rapidly, preparing questions for exams or studying is a time-consuming process and requires a lot of critical thinking. The questions we solve in exams like school and college level exams are similar to previous year papers and contain repeated questions with little or no paraphrasing or modification[9].

According to **A Saranya**.et al., 2023 we introduce a system for automatic assessment of question paper quality. The question paper is an important part of academic assessment. The quality of the questionnaire is crucial to achieve the goal of the evaluation. Question papers are prepared manually in many fields of education. A preliminary analysis of the questionnaire can help identify errors in the question and better achieve the assessment objectives[10].

III.PROPOSED SYSTEM

The proposed system integrates state-of-the-art NLP algorithms for tasks such as sentence segmentation, named entity recognition (NER), and semantic analysis. Upon receiving input text, the system preprocesses and analyzes the content to identify key information, important phrases, and significant entities. Subsequently, utilizing the extracted information, the system formulates questions that are contextually relevant and coherent with the input text. The Flask framework facilitates seamless integration and deployment of the system, providing an intuitive user interface for input and output interactions.

ARCHITECTURE DIAGRAM:



Explanation:

1.User Interface (Frontend): The user interacts with the system through a web-based interface developed using HTML, CSS, and JavaScript. This interface allows users to input text passages and submit them for question and answer generation.

2.Flask Framework (Backend): Flask, a lightweight Python web framework, serves as the backend of the system. It handles user requests, processes input data, and orchestrates the NLP-based question and answer generation process.

3.NLP Pipeline:Text Preprocessing: Input text passages undergo preprocessing steps such as tokenization, stop word removal, and stemming to enhance the quality of analysis.

Feature Extraction: NLP techniques are employed to extract relevant features from the preprocessed text, including word embeddings or contextualized representations.

Question Generation: Using the extracted features, questions are generated based on key concepts, entities, and relations identified within the text. Techniques such as rule-based approaches or neural models may be utilized for question formulation.

Answer Generation: Once questions are formulated, the system generates corresponding answers by analyzing the text passages. This process involves identifying relevant

information within the text and synthesizing concise and accurate responses.

Quality Assessment:

To ensure the quality of generated questions and answers, the system may incorporate evaluation metrics or human-in-the-loop validation mechanisms.

4.Model Integration:State-of-the-art NLP models such as BERT (Bidirectional Encoder Representations from Transformers) or GPT (Generative Pre-trained Transformer) may be integrated to enhance the accuracy and robustness of question and answer generation.

Pre-trained language models are fine-tuned on specific tasks related to question and answer generation using domain-specific datasets to improve performance.

5.Data Storage and Management: The system may utilize databases or file systems to store and manage input text passages, generated questions, and corresponding answers. Persistent storage mechanisms ensure that generated questions and answers can be retrieved and accessed efficiently.

6.Deployment: The system can be deployed on cloud platforms such as AWS (Amazon Web Services), Google Cloud Platform, or Azure for scalability and accessibility. Docker containers or virtual environments may be employed to encapsulate the application and its dependencies, facilitating seamless deployment and reproducibility.

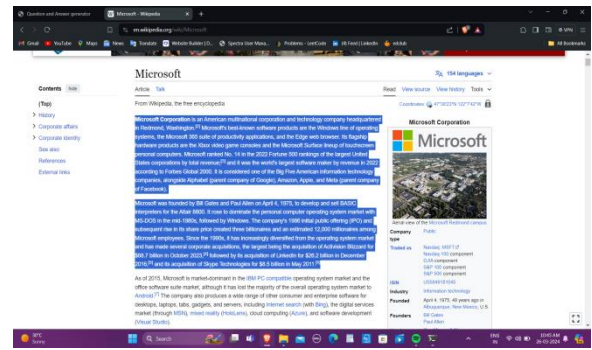


FIG2. Search any topic and copy any paragraph

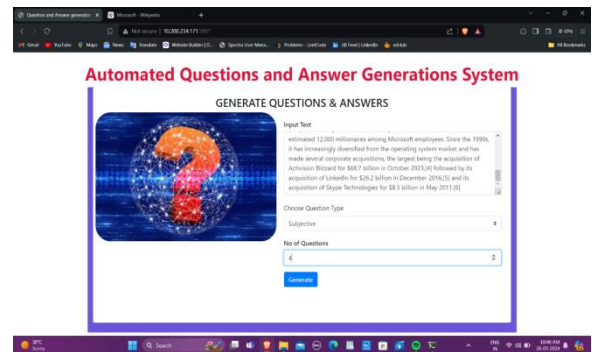


Fig3. Paste the paragraph from input text box and type number for no of questions.

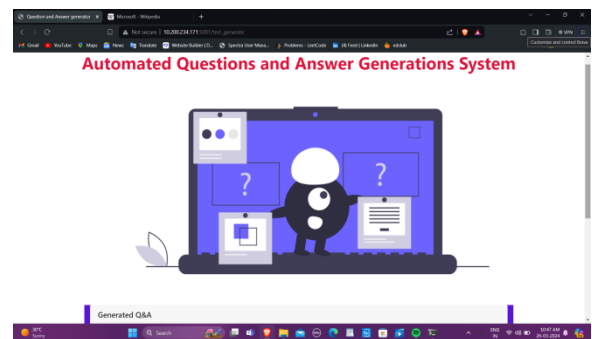


Fig4. You got output home page and you can scroll down for output.

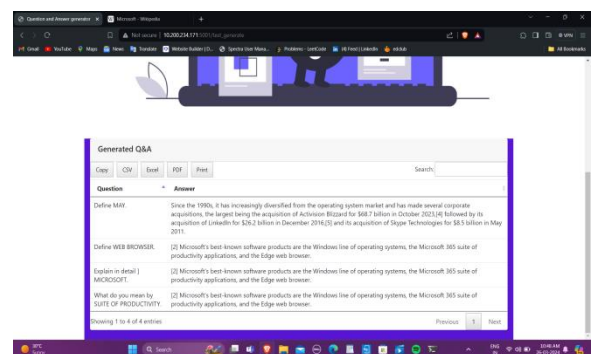


Fig5. You got the output and click any option and I can click pdf option to download the file

IV.RESULT AND DISCUSSION

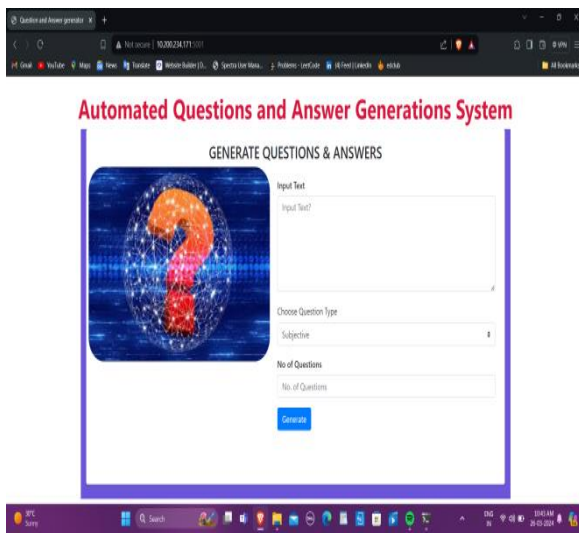


Fig1.HOME PAGE:The home page of SecureGuard offers a user-friendly interface for easy navigation, providing essential features like search, navigation menu, and announcements for system updates.

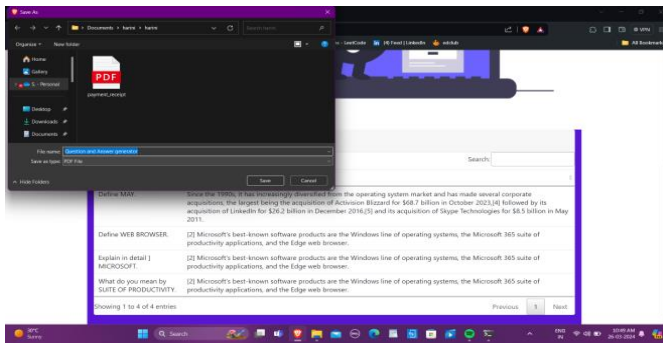


Fig6. You can see the display to save the file.

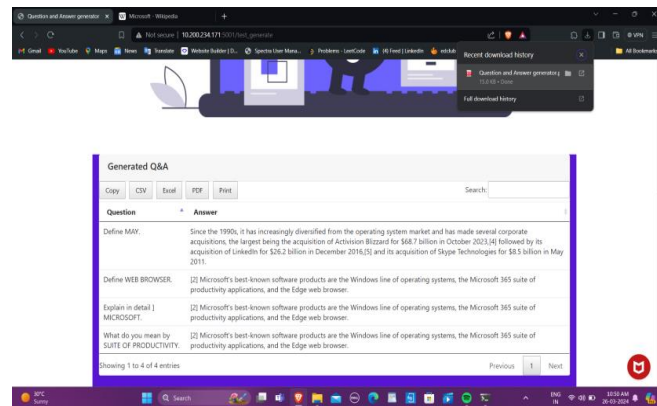


Fig7. Downloaded the file.

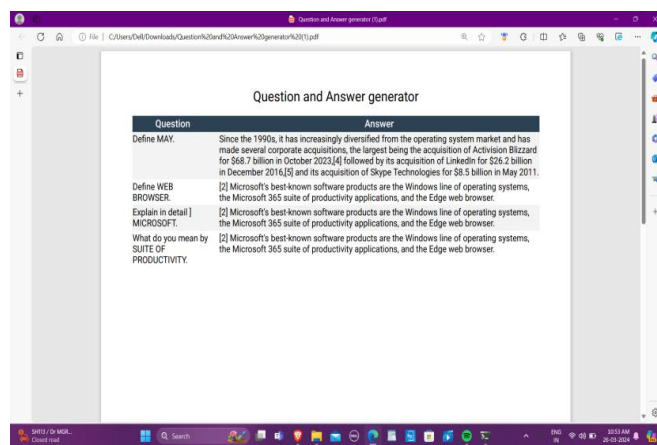


Fig8. And finally you got the output and you can read the file.

V.CONCLUSION

The development and evaluation of the automatic question and answer generation system signify a significant advancement in leveraging NLP capabilities for enhancing text comprehension and accessibility. The integration of advanced NLP techniques with the Flask framework has enabled the creation of a robust and efficient system capable of generating contextually relevant questions from diverse textual data. Future research and enhancements will focus on optimizing the system's

accuracy, scalability, and adaptability to cater to broader applications and domains in the realm of NLP.

REFERENCES

- [1] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.
- [2] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems* (pp. 5998-6008).
- [3] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
- [4] T. B. Shahi and C. Sitaula, "Natural language processing for Nepali text: a review," *Artif. Intell. Rev.*, vol. 55, no. 4, pp. 3401– 3429, 2022, doi: 10.1007/s10462-021-10093-1.
- [5] N. Patil, K. Kumari, D. Ingale, P. Patil, and A. R. Uttarkar, "A Survey on Automatic Multiple Choice Questions Generation from Text," *Int.J. Sci. Res. Eng. Trends*, vol.7, no. 3, pp. 1997–1999, 2021.
- [6] A Bergkvist, N Hedberg, S Rollino, M Sagen - 2020, *Sumrize: An Online NLP System for Close-Domain Question-Answering and Summarization* - diva-portal.org.
- [7] Ekta Tyagi, Deeksha, Lokesh Chouhan 2021, *Design of automatic answer checker, Congress on Intelligent Systems: Proceedings of CIS 2020, Volume 2, 237-252, 2021*
- [8] Wenyan Yang 2022 , *Design and Implementation of Automatic Examination Scoring System Based on Natural Language Processing, 3rd International Conference on Big Data and Informatization Education (ICBDIE 2022), 1048-1058.*
- [9] Akanksha Sharma, Saket Savarn, Adarsh Anand, Sunil Kumar S Manvi 2023, *Automated Questions Unique Arrangement (AQUA), Journal of Advanced Zoology 44.*
- [10] A Saranya, M Monisha, P Nandhini, M Rakshana, R Tejashree 2023, *Student Automatic Correction Using Natural Language Processing, International Journal of Modern Developments in Engineering and Science 2 (3), 27-31.*