International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 29 Issue 3 – APRIL 2022.

Automatic Paper Corrector Using NLP

MR.MK.NAGARAJAN
Department of Computer Science and Engineering,
Kalasalingam Academy of Research Education,
Virudhunagar, Tamil Nadu, India

R RAJA SUBRAMANIAN
Department of Computer Science and Engineering,
Kalasalingam Academy of Research Education,
Virudhunagar, Tamil Nadu, India

AKSHAYA S
Department of Computer Science and Engineering,
Kalasalingam Academy of Research Education,
Virudhunagar, Tamil Nadu, India

BASIREDDY SREYA
Department of Computer Science and Engineering,
Kalasalingam Academy of Research Education,
Virudhunagar, Tamil Nadu, India

ABSTRACT:*This paper proposes the idea of a "automated paper corrector."This is designed to check subjective responses in an online exam and assign marks to the user after the answer has been verified. The mechanism necessitates that you save the original response. To reduce manual labour, an automatic paper corrector has been developed. Everything has gone virtual as a result of the Covid19 epidemic. It is difficult to virtually correct papers during online exams. So, using this Automatic paper corrector, students can submit their answers online, and it will correct the paper and display the necessary marks. This paper will explain the design and features of automatic paper correcting.*

I. INTRODUCTION

It is made up of statistical natural language processing libraries and programmes. The libraries support tokenization, classification, parsing, stemming, tagging, and sematic reasoning. A machine can understand human language using this set of techniques.The proposed solution can eliminate all of the flaws in the current system while still providing acceptable security and reducing human labour. Regular language handling (NLP) is an area of semantics, software engineering, and man-made reasoning that looks at how PCs associate with human language, explicitly how to build PCs that can cycle and examines a lot of normal language data.The objective is to foster a PC that can "get" the items in archives, remembering language subtleties for setting. From that point onward, the framework can separate right information and bits of knowledge from the articles, as well as order and sort out them.

II.PROBLEM STATEMENT

In the existing system the exams are done manually. If the system is not provided the right inputs, it will produce incorrect outputs. Data security is lacking. More manpower is necessary to complete the task. Time can be saved by using an automatic paper corrector. Under the current system, exams are currently completed manually.

III.PROPOSED WORK

It is made of libraries and programmes for statistical natural language processing. Tokenization, classification, parsing, stemming, tagging, and sematic reasoning are all possible with the libraries. A machine can understand human language using this set of techniques. The proposed solution can eliminate all of the flaws in the current system while still providing acceptable security and reducing human labour.

3.1 TOKENIZING

Tokenizing text allows you to rapidly break it down into individual words or sentences.This will enable you to work with smaller chunks of text that are still cohesive and clear when removed from their context.On a word-by-word basis, tokenization:The building blocks of natural language are words.They're the tiniest unit of meaning that can nonetheless be comprehended on their own.It's the first stage in converting unstructured data into structured data that can be more easily analysed.You'll tokenize words by word and sentences by sentence while analysing text.The following are the advantages of both types of tokenization:You can detect common words by tokenizing your text word by word. If you look at a bunch of job ads, for example, you'll find that the word "Python" appears quite frequently. This could suggest a significant demand for Python.

3.2BLOCK DIAGRAM:

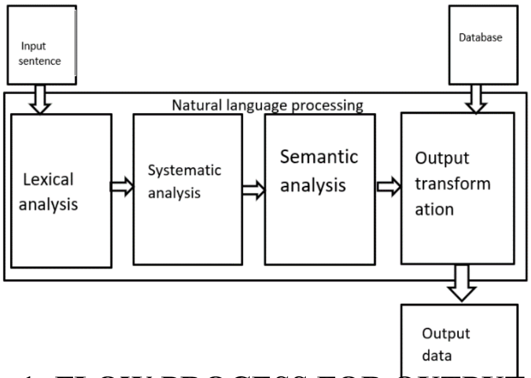

Fig 1: FLOW PROCESS FOR OUTPUT

3.3 FILTERING STOP WORDS, STEMMING

Stop words are terms that you don't want to read, therefore they're removed from your text as it's being processed. Stop words like 'in,' 'is,' and 'an' are frequently employed as stop words because they offer little to the text on their own. Stemming is a text processing technique that involves words to their root, or the most fundamental piece of the word. The names "making a difference" and "aide," for instance, have similar starting points. Stemming permits you to focus on the fundamental importance of a word as opposed to the complexities.

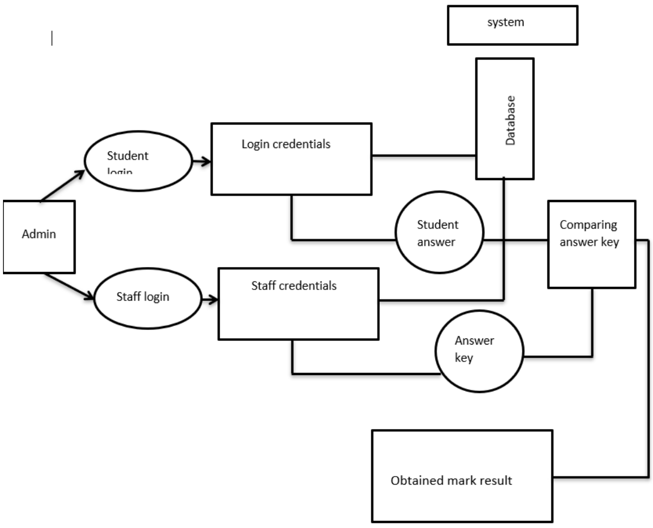

Fig 2.USE CASE DIAGRAM

3.4 TKINTER PROGRAMMING

Tkinter is Python's standard GUI library. Whenever Python is used associated with Tkinter, it's fast and straightforward to create graphical UIs. Tkinter accompanies a helpful thing. Tk GUI tool stash's arranged connection point. Making a GUI application with Tkinter's straightforward. you must simply follow the means beneath. • In your task, incorporate the Tkinter module. • For the GUI programming, make the essential window. • Add a minimum of one amongst the previously mentioned gadgets to the GUI program. • To react to each event that the user has triggered, enter the main event loop.

IV HARDWARE REQUIREMENT

Processor – i3,
Hard Disk – 5 GB,
Memory – 4GB RAM 4.3.

4.1 SOFTWARE REQUIREMENT

1. Software tools related to NLP: Python Tensor Flow ,PyTorch ,Hugging Faces Transformers, Frontend ,Basic GUI tools, Spacy Streamlit or Gradio.

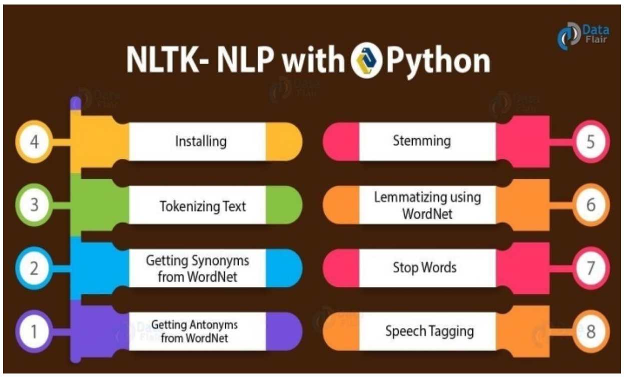

Fig 3:OVERVIEW OF NLP[process of tokenizing]

V. ADVANTAGES

1. Examiners become bored when they have to check a large number of answer sheets; therefore, the system reduces their workload by accurately automating the manual checking process.

2.The system computes the score and displays the findings immediately.

3. It eliminates human errors that are frequent during manual checks.

4.The technology generates a completely unbiased outcome.

5.As a result, the technology eliminates human effort while saving time and resources.

CONCLUSION

This project might be expanded by scanning the text and then converting it to legible text using optical recognition. OCR (Optical Character Recognition) is a software technology that recognises text (written or printed) within an image file or physical document (such as a scanned document) and turns it into machine-readable text for data processing. Text recognition is another term for it. optical The most important element of OCR is connecting to a scanner and scanning the document. The amount of variables to consider while building OCR software is minimised because scanning the paper standardises the inputs. This phase also improves the efficiency of the overall process by ensuring that the document is perfectly aligned and sized. Students can use this Automatic paper corrector to enter their answers online, and the application will correct the paper and provide the necessary marks. A computer programme that assesses and scores written responses in the same way that a human would. This software application verifies subjective responses in an online exam and assigns marks to the user when the user's answer is validated.

REFERENCES:

[1] Song, Shengli, Haitao Huang, and Tongxiao Ruan. "Abstractive text summarization using LSTM-CNN based deep learning." Multimedia Tools and Applications 78.1 (2019):857-875.

[2] R. Mar´ın, P. Sanz, O. Coltell, J. Inesta, F. Barber, and D. Corella, “Student-teacher communication directed to computer-based learning environments,” Displays, vol. 17, no. 3, pp. 167–178, 2019.

[3] .Shirwandkar, Nikhil S., and Samidha Kulkarni. "Extractive text summarization usingdeep learning." 2018 Fourth International Conference on Computing Communication

Control and Automation (ICCUBEA). IEEE, 2018.

[4] Standards for Web Applications on Mobile:current state androadmap https://www.w3.org/2015/01/mobile-web-app-state/

[5] .R. M. McHugh, C. G. Horner, J. B. Colditz, and T. L. Wallace, “Bridges and barriers adolescent perceptions of student–teacher relationships,” Urban Education, vol. 48, no. 1, pp. 9– 43, 2013.

[6] Nenkova, Ani, and Kathleen McKeown. "A survey of text summarization techniques."Mining text data. Springer, Boston, MA, 2012. 43-76.

[7] Mihalcea, Rada, Courtney Corley, and Carlo Strapparava. "Corpus-based and knowledgebased measures of text semantic similarity." Aaai. Vol. 6. No. 2006. 2006.

[8] Proceedings of IEEE 24th International Symposium on Fault- Tolerant Computing, J.D. BrightComputer Science Department, Johns Hopkins University, Baltimore, MD, USA-G.P. Sullivanhttp://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=315646&queryText%3D Ans wer+Checker.

[9] Proceedings., Fourteenth Annual International Computer Software and Applications Conference, S. Ramanna,J.F. Peters Kansas State University, Manhattan, KS, USA http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=139342&queryText%3DAns wer+Checker