SMART ASSESSMENT TOOL FOR DESCRIPTIVE ANSWERS

A report submitted in partial fulfillment of the requirements

Of

Mini-Project (IS65)

In

Sixth Semester

By

1MS20IS036 Deepa C

1MS20IS044 Diya James

1MS20IS143 Anushka Gupta

1MS20IS146 Anagha Kalyani

Under the guidance of

Rajeshwari S B
Assistant Professor
Dept. of ISE, RIT

RAMAIAH
Institute of Technology

DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING
RAMAIAH INSTITUTE OF TECHNOLOGY
(AUTONOMOUS INSTITUTE AFFILIATED TO VTU)
M. S. RAMAIAH NAGAR, M. S. R. I. T. POST, BANGALORE – 560054

2022-2023

RAMAIAH INSTITUTE OF TECHNOLOGY

(Autonomous Institute Affiliated to VTU)

M. S. Ramaiah Nagar, M. S. R. I. T. Post, Bangalore – 560054

DEPARTMENT OF INFORMATION SCIENCE AND ENGINEERING

RAMAIAH
Institute of Technology

CERTIFICATE

This is to certify that the project work entitled **"Smart Assessment Tool for Descriptive Answers"** is a bonafide work carried out by **Deepa C, Diya James, Anushka Gupta** and **Anagha Kalyani** bearing **USN: 1MS20IS036, 1MS20IS044, 1MS20IS143** and **1MS20IS146** respectively in partial fulfillment of requirements of Mini-Project (IS65) of Sixth Semester B.E. It is certified that all corrections/suggestions indicated for internal assessment has been incorporated in the report. The project has been approved as it satisfies the academic requirements in respect of project work prescribed by the above said course.

_____
Signature of the Guide
Rajeshwari S B
Assistant Professor
Dept. of ISE, RIT,
Bangalore-54

_____
Signature of the HOD
Dr. Sanjay H A
Professor and Head,
Dept. of ISE, RIT
Bangalore-54

ACKNOWLEDGEMENT

Any achievement, be it scholastic or otherwise does not depend solely on the individual efforts but on the guidance, encouragement and cooperation of intellectuals, elders and friends. A number of personalities, in their own capacities have helped us in carrying out this project work. We would like to take this opportunity to thank them all.

First and foremost we would like to thank Dr. N.V.R Naidu, Principal, M.S.R.I.T, Bengaluru, for his moral support towards completing our project work.

We would like to thank Dr. Sanjay H A, The Head of Department, Information Science & Engineering, M.S.R.I.T, Bengaluru, for his valuable suggestions and expert advice.

We deeply express our sincere gratitude to our guide Rajeshwari S B**,** Assistant Professor**,** Department of ISE, M.S.R.I.T, Bengaluru, for her able guidance, regular source of encouragement and assistance throughout this project.

Our sincere gratitude to our Parents, and brothers for their continuous support in all aspects of my life. Without their support we wouldn't have completed our course.

Our thanks to all the Faculty members of the Department of Information Science & Engineering for their constant support and encouragement.

Last, but not the least, we would like to thank our peers and friends who provided us with valuable suggestions to improve our project.

ABSTRACT

Automated evaluation of handwritten subjective exam papers is an important area of research with potential applications in the education sector. The project aimed to develop a system that used OCR technology and string matching algorithms to automate the evaluation of handwritten subjective exam papers, with the objective of reducing the workload of examiners and providing faster and more accurate evaluation.

The proposed methodology involved using OCR technology to convert scanned documents into machine-readable text, followed by keyword extraction and string matching algorithms for evaluation. The system was developed using Python and some libraries, including NLTK and FuzzyWuzzy. The limitations of existing models and algorithms provided the motivation for this project, which aimed to improve the accuracy and efficiency of handwritten subjective exam paper evaluation.

The expected outcome of this project was a system that could automate the evaluation of handwritten subjective exam papers accurately and efficiently. Future work could involve the use of deep learning models, such as Recurrent Neural Networks (RNNs), for further improving the accuracy of the system. Additionally, the system could be extended to evaluate other types of responses, such as flow charts and diagrams.

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

Evaluating subjective exam papers is a challenging task for educators, as it requires them to evaluate the students' responses based on various factors, such as the clarity of ideas, logical reasoning, and language proficiency. The evaluation process is time-consuming and subject to human errors, such as biased grading and inconsistent marking, which can be detrimental to the students' academic progress.

To address these issues, this project aims to automate the assessment of handwritten subjective exam papers using Optical Character Recognition (OCR) technology, machine learning algorithms, and natural language processing techniques. The system's primary objective is to reduce the workload of educators and ensure uniformity and accuracy in the evaluation process. By automating the process, the system can provide an accurate and unbiased evaluation of the students' responses, which will ensure uniformity and consistency in the grading process.

The proposed system uses OCR technology to convert the scanned documents into machine-readable text and extract keywords. The system then uses WordNet, a lexical database, to check synonyms, hyponyms, and meronyms of the keywords to handle variations in student responses. String matching techniques and fuzzy logic are used for calculating scores based on the similarity between the students' responses and the model answers.

The expected outcome is a faster and easier evaluation process that allows for instant score calculation and reduces the workload of examiners. For assessment of descriptive answers, the machine needs to consider the important factors that a human assessor takes into account for manual evaluation. Handwritten documents can be scanned and fed as input to the system, and keyword matching can be used for the assessment of the learners' responses.

The proposed system will help make the evaluation process faster and easier. It will allow learners' responses to be assessed without instructors' manual interruption. The system recognizes handwritten

answers, converts them to machine-readable text, and calculates the scores instantly, which will help in reducing the workload of examiners by automating the manual checking process accurately.

Chapter 2

LITERATURE SURVEY

2.1 Handwritten Text Recognition using Deep Learning [1]

The project aims to develop a Handwritten Text Recognition (HTR) system that accurately recognizes handwritten text by collecting data, extracting features, and training a model using Deep Learning. The methodology of HTR includes pre-processing steps such as noise removal, segmentation, thinning, binarization, and normalization. The built model uses LSTM deep model and achieves high accuracy. The IAM offline handwriting dataset is used for training and testing. The HTR system can recognize words instead of characters to improve accuracy. The recognition process involves uploading an image that undergoes pre-processing steps and is normalized before being sent to the classifier. Some of the limitations are that this model is restricted to only the English language and doesn't recognize symbols.

2.2 An Intelligent System for Evaluation of Descriptive Answers [2]

The proposed system for evaluating descriptive answers of students uses four different strategies: keyword and similarity score, language score, fuzzy string matching score, and concept graph score. The question is classified into three categories: conceptual, specific, and analytical. The system automatically generates a concept graph, and the final score is computed using the weighted sum of scores obtained from each strategy. The language and fuzzy string matching modules are used to check grammar, spellings, and approximate string matching. Finally, the system uses the Levenshtein distance to calculate the difference between two strings.

2.3 Handwritten Text Recognition System Based on Neural Network [3]

It aims to develop a system that recognizes handwritten English characters by processing input, extracting features, training a neural network, and generating computerized output. The system is divided into two parts: training and testing of the artificial neural network using Resilient Back-propagation. The training part involves creating a dataset, preprocessing it, extracting features, training the ANN, and saving it for testing. The testing part involves pre-processing the input image, segmenting it into individual characters, and extracting features using a bit map version of the image, which serves as an input vector to the ANN. The system achieves high accuracy in recognizing handwritten characters without affecting the speed of recognition. The proposed system doesn't recognize symbols.

2.4 NLP-based Automatic Answer Script Evaluation [4]

This study aims to evaluate descriptive answer scripts and assign marks automatically using various similarity measures calculated after text extraction, summary generation, text preprocessing, and similarity measurement techniques. Python programming language is used, along with the Natural Language Processing (NLP) toolkit NLTK, to implement the algorithms. The study uses two approaches, one based on bag-of-words and the other on keywords. The study concludes that the keyword-based approach produces a higher F-score. Several similarity measures are used to evaluate the text. NLTK is used to preprocess the text by tokenizing, removing stop words, and lemmatizing. A weight value is assigned to each parameter manually by doing a survey.

2.5 Smart Online Assessment System for Descriptive Responses using Machine Learning Techniques [5]

The proposed system uses keyword matching and semantic analysis to assess learner response. It involves an instructor interface where questions and key answers are input, a learner interface for taking the test, and a text processing module for removing punctuation, tokenizing, and obtaining keywords from both key answers and learner responses. WordNet processing is used to obtain semantic meaning from the key answer keywords. Keywords from key answers and their semantic meaning are compared with those from the learner's response, and a decision tree is used to assign grades based on the number of matching keywords. The algorithm assigns five marks for five or more matching keywords, four marks for four matching keywords, and so on. It is assumed that the learner or instructor inputs the grammar and spelling. It requires answers to be typed, not handwritten.

Chapter 3

REQUIREMENTS SPECIFICATION

3.1 Functional Requirements

3.1.1 OCR and Keyword Extraction

- The system shall be able to accept scanned images of handwritten exam papers as input.
- The system shall use OCR technology to convert the scanned images into machine-readable text.
- The system shall extract keywords from the machine-readable text using natural language processing techniques.
- The system shall check synonyms, hyponyms, and meronyms of the keywords using WordNet to handle variations in student responses.

3.1.2 String Matching and Fuzzy Logic

- The system shall compare the extracted keywords with the correct answers to calculate a similarity score.
- The system shall use fuzzy logic to calculate a grammar score based on the extracted text.
- The system shall use string matching techniques to calculate a keyword accuracy score.

3.1.3 User Interface

- The system shall have a user-friendly graphical interface that allows examiners to easily upload and evaluate exam papers.
- The system shall display the calculated scores for each exam paper in an easy-to-understand format.

3.2. Non-Functional Requirements

- **Reliability**

The system shall be reliable and accurate in assessing handwritten papers.

- **Security**

The system shall be secure and protect exam papers and examiners' information.

- **Usability**

The system shall be easy to use and require minimal training for examiners.

3.3 Constraints

- The system shall run on a desktop computer with at least 4GB of RAM and a 2.0GHz processor.
- The system shall be developed using Python programming language and related libraries for natural language processing and fuzzy logic.
- The system shall be compatible with popular operating systems like Windows, Linux, and MacOS.

3.4 Assumptions

- The system assumes that the handwritten exam papers are of good quality and can be easily read by OCR technology.
- The system assumes that the correct answers for the exam papers are available for comparison and evaluation.

Chapter 4

FRAMEWORK AND SYSTEM DESIGN

4.1 User Interface

Python Tkinter - A standard Python library for creating graphical user interfaces (GUIs) with a set of built-in GUI components or widgets. Tkinter is based on the Tk GUI toolkit, which is developed for the Tcl programming language, but it has been adapted for use with Python.

4.2 Dataset

The IAM dataset is a widely used handwriting recognition dataset that contains handwritten text data along with corresponding ground truth transcriptions. It consists of forms from 657 writers, including 1,539 pages of text containing 5,826 paragraphs and 115,320 isolated text lines.

4.3 String Matching

FuzzyWuzzy - Provides a set of functions for comparing strings and determining the similarity between them based on different metrics such as Levenshtein distance, Jaro distance, and others. It is particularly useful when dealing with strings that may contain typos, misspellings, etc.

4.4 Reading input data

Xlrd - Used to extract data from Microsoft Excel spreadsheets. It is a module that is specifically designed to work with the older .xls format.It can also be used to extract data from specific cells, ranges of cells, or entire sheets.

4.5 Fuzzy logic

Fuzzy logic is a mathematical concept that deals with uncertainty and imprecision. It allows for values between true and false, rather than just binary values. Fuzzy logic is used in artificial intelligence and machine learning to model complex systems and decision-making processes. The Python library "fuzz" is a collection of fuzzy string matching algorithms, which can be used to compare strings for similarity even when they are not exact matches.

Chapter 5

IMPLEMENTATION

The evaluator must provide the answer key for the given questions. The handwritten responses of the learner must be scanned and uploaded. The system will convert the handwritten document to text using Optical Character Recognition (OCR). OCR is a technology that converts scanned text images into machine-readable text formats. The process involves-

a. **Preprocessing** - remove the background noise, enhance the region of interest in image and make a clear difference between foreground and background. In order to achieve these goals: noise filtering, conversion to binary and smoothing operations are performed on the input image
b. **Feature Extraction -** Feature vector is calculated from the processed image.
c. **Text recognition -** done by passing the obtained feature vector as input to machine learning algorithms to analyze and recognize the shapes and strokes of handwritten characters.

The model takes an image file as input, which could be a scanned document or a photo of text. The image is preprocessed to enhance its quality and make it easier to recognize the text. This includes operations like thresholding, smoothing, and binarization [1]. The preprocessed image is then passed to OCR engine, which analyzes the image and extracts the text using techniques such as character segmentation, feature extraction, and classification. The recognized text is returned to the Python program. The Python program then processes the recognized text further for evaluation. Extracting keywords from the learner's response involves -

- Tokenizing text into words.
- Counting frequency of each word and identifying keywords based on their frequency.
- Removing duplicates from the word list.

Since learners can explain answers in different ways, WordNet will be used to check synonyms, of the keywords from both learner's response and answer key [5]. The evaluation is done based on four factors namely, similarity factor, keyword factor, grammar factor, keyword accuracy

factor. Similarity score is calculated based on fuzzy logic and string matching [2]. Grammar is checked using language_tool_python library. The score is initially set to 0 and is incremented. The detailed report is displayed.

The flowchart for the algorithm followed is as follows:

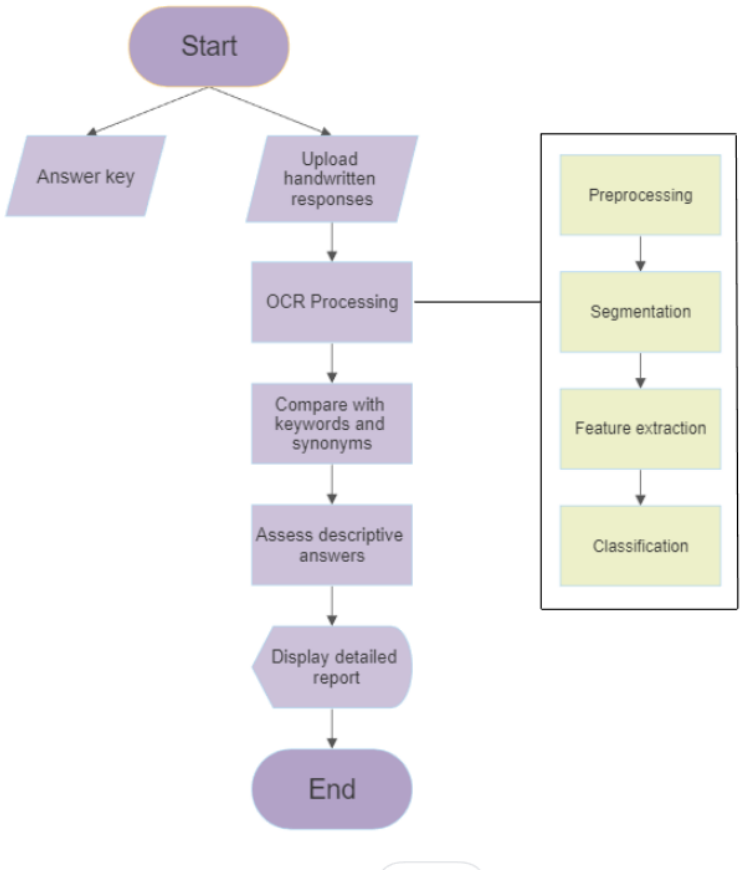

Figure 5.1 - Algorithm flowchart

The model uses Convolutional Neural Network (CNN) for text detection and a Long Short-Term Memory (LSTM) network for text recognition. The feature vector obtained after feature extraction is fed as input for the LSTM model [1].

Convolutional Neural Network (CNN) - The main idea behind CNNs is to learn local patterns or features from the input image by applying convolutional filters over the input. These filters detect specific visual patterns such as edges, corners, and textures, which are then combined to

form higher-level features. The convolutional filters are learned through backpropagation during training, allowing the network to learn the most relevant features for the given task.

LSTM - It is a special kind of recurrent neural network that is capable of learning long term dependencies in data. This is achieved because the recurring module of the model has a combination of four layers interacting with each other.

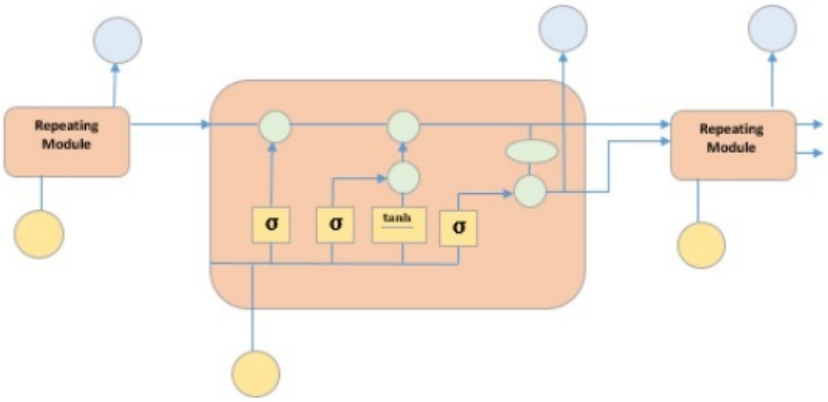

Figure 5.2 - LSTM Model Architecture

In fig. 5.1 four neural network layers are shown as yellow boxes, pointwise operators are shown as green circles, input is shown as yellow circles, and cell state is shown as blue circles. A cell state, three gates, and an LSTM module provide them the ability to selectively learn, unlearn, or retain information from each of the units. By allowing only a small number of linear interactions, the cell state in LSTM aids in the uninterrupted flow of information across the units. Each component contains an input, an output, and a forget gate that can add or remove data from the cell state. The forget gate utilizes a sigmoid function to determine which information from the previous cell state should be ignored.Using pointwise multiplication of "sigmoid" and "tanh," the input gate regulates the information flow to the active cell state. The output gate ultimately determines which information should be transferred to the following concealed state.

Chapter 6

TESTING

The model was tested with several test cases, some of which are discussed in this section.

6.1 Test case 1

In this test case, the answer was written by the student in lower case letters.

This was for a question based on “E-Commerce”. The text from the uploaded answer key was correctly recognised as “electronic way to transfer data” and fed to the evaluation portion of the code. The answer provided by the student had 3 keywords out of the 9 which were extracted from the answer key, showing a keyword match of 33.33%. The average similarity score of the answer was calculated as 84% after checking with all the sample answers given by the evaluator. Since the sentence was grammatically correct, the grammar score was 100%. The answer was also checked for the order in which the keywords occurred, based on which a keyword accuracy score of 66.66% was given. Overall, since the answer was relevant and passed all the 4 criteria based on the weightage provided by the evaluator, a total of 10 marks was given to the student.

Table 6.1 - Results for test case 1

Criteria	Weightage provided by evaluator	Score assigned by Smart Assessment Tool
Keyword match	0.5	33.33%
Similarity score	1.8	84%
Grammar score	0.2	100%
Keyword accuracy score	0.1	66.67%

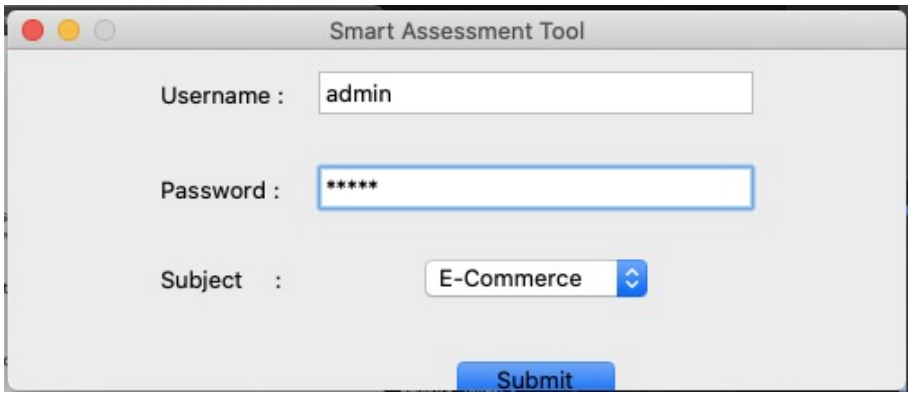

Figure 6.1 - Login Window

	Values	Similiarity factor	Keyword Factor	Grammer Factor	Keyword Accuracy Factor
2		1.8	0.5	0.2	0.1
3	Question:				
4			Q1: What is EDI?		
5					
6	Keywords:				
7	Electronic Data Interchange is structured transmission of data between organizations by electronic means. It was previously done by paper but now it is done electronically.				
8	EDI is a means of transfer of electronic documents or business data from computer system to another computer system electronically.				
9	It is a method of transfer of documents from one computer to another through online.				
10	Electonic Data exchange is a method to deliver data between computers effectively through the digital media.				
11	EDI stands for Electronic data exchange which enables the user to send data over the internet electronically				
12					
13					

Figure 6.2 - Answer key for E-commerce question (Test Case 1)

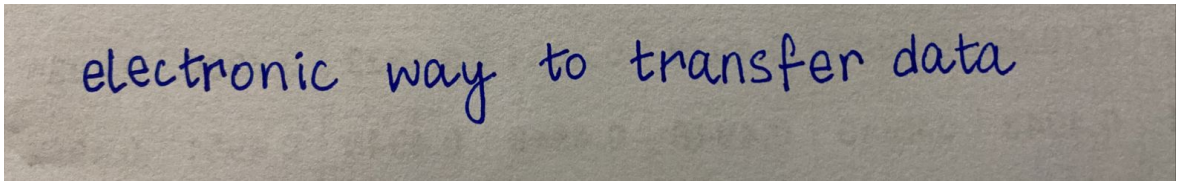

Figure 6.3 - Answer Script submitted by student (Test Case 1)

Figure 6.4 - Input image path (Test Case 1)

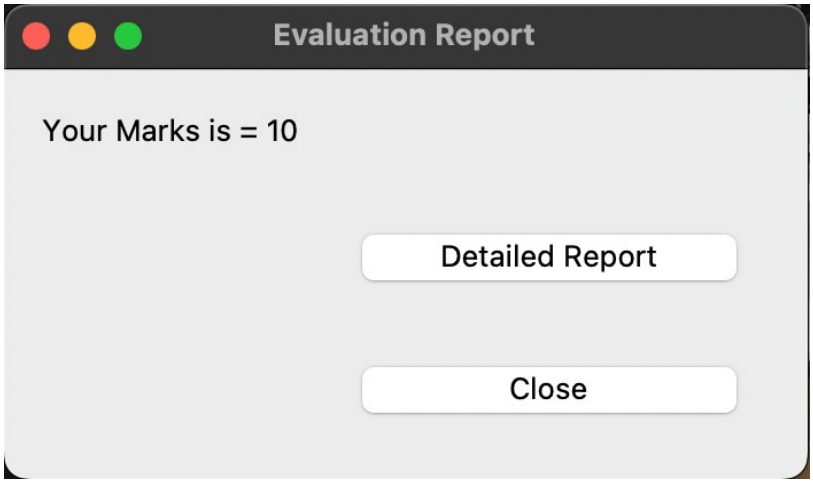

Figure 6.5 - Evaluation Report (Test Case 1)

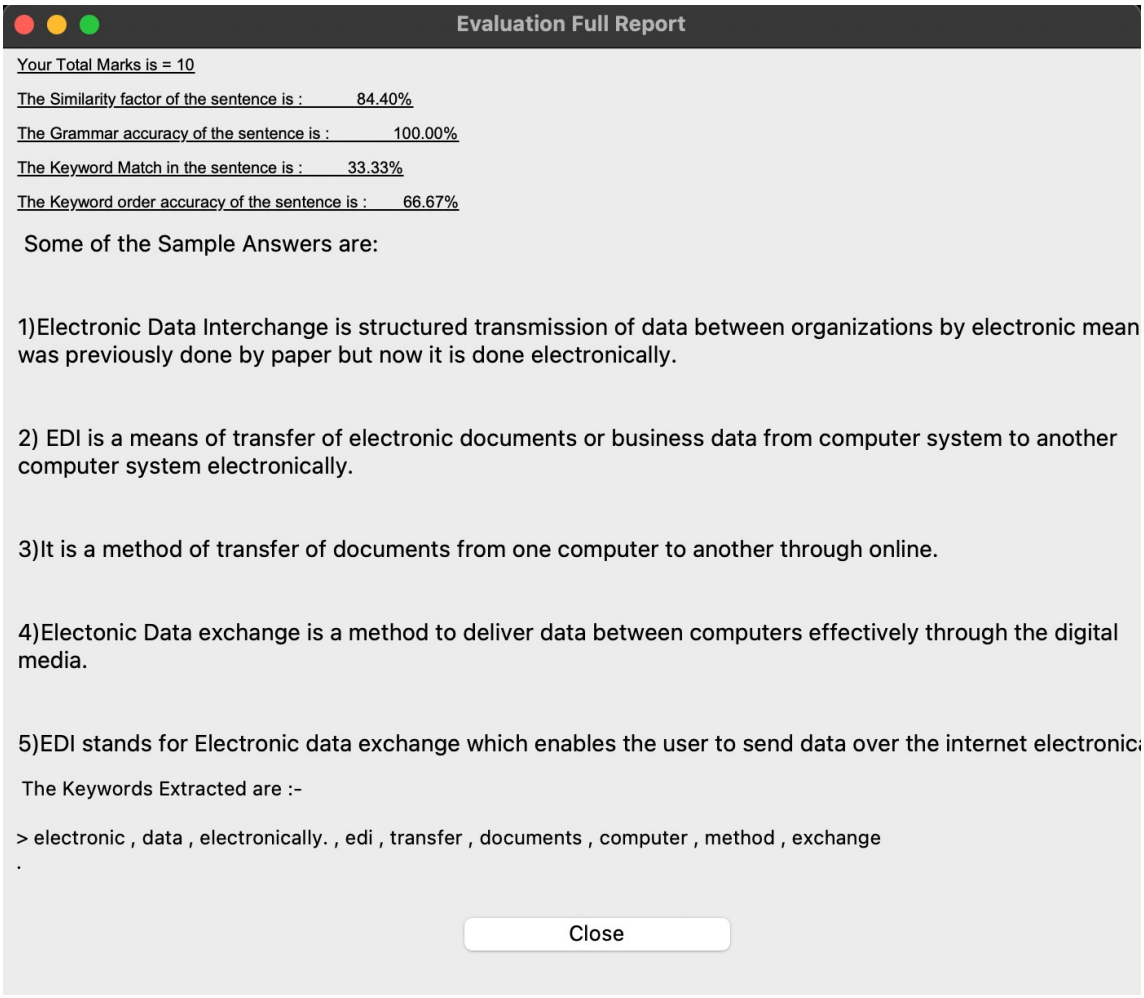

Figure 6.6 - Detailed Evaluation Report (Test Case 1)

6.2 Test case 2

This test case was to check the accuracy of the tool in recognising a handwritten paragraph, instead of just a single line.

This was for a question based on “E-Commerce”. The text from the uploaded answer key was recognised as “EDL STANDS FOR ELECTRONIC PATA EXCHANGE WHICH ENABLES THE USER TO SEND PATA OVER THE INTERNET ELECTRONICALLY” and fed to the evaluation portion of the code. The invalid word “PATA” was changed to “DATA” by the tool for further evaluation. The answer provided by the student showed a keyword match of 33.33%. The average similarity score of the answer was calculated as 100% after checking with all the sample

answers given by the evaluator. Since the sentence was grammatically correct, the grammar score was 100%. The answer was also checked for the order in which the keywords occurred, based on which a keyword accuracy score of 66.66% was given. Overall, since the answer was relevant and passed all the 4 criteria based on the weightage provided by the evaluator, a total of 10 marks was given to the student.

Table 6.2 - Results for Test Case 2

Criteria	Weightage provided by evaluator	Score assigned by Smart Assessment Tool
Keyword match	0.5	33.33%
Similarity score	1.8	100%
Grammar score	0.2	100%
Keyword accuracy score	0.1	66.67%

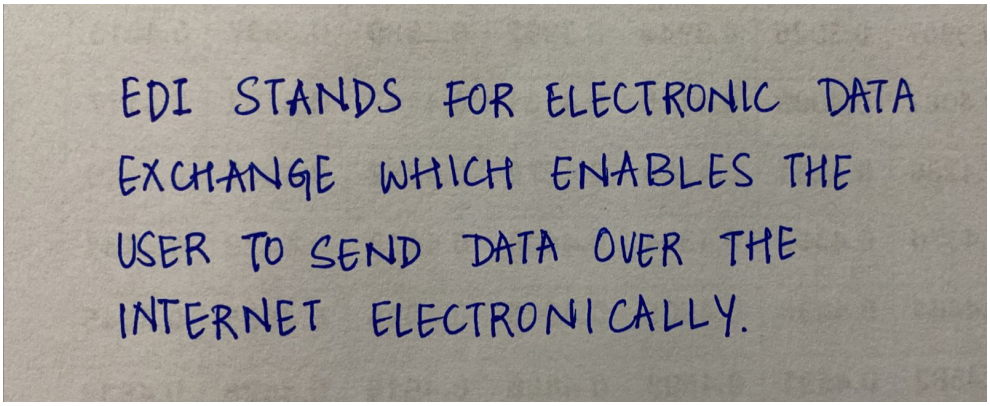

Figure 6.7 - Answer Script submitted by student (Test Case 2)

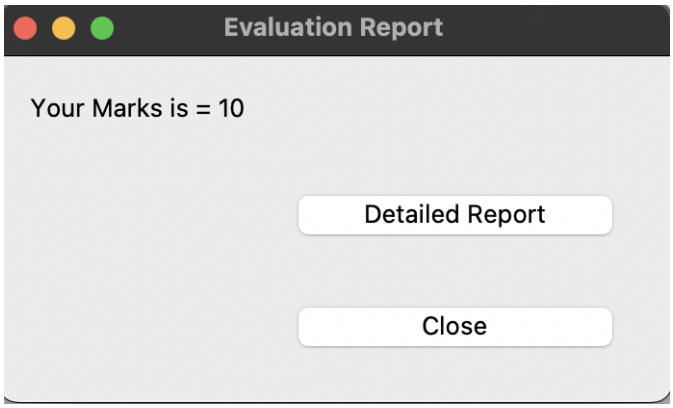

Figure 6.8 - Evaluation Report (Test Case 2)

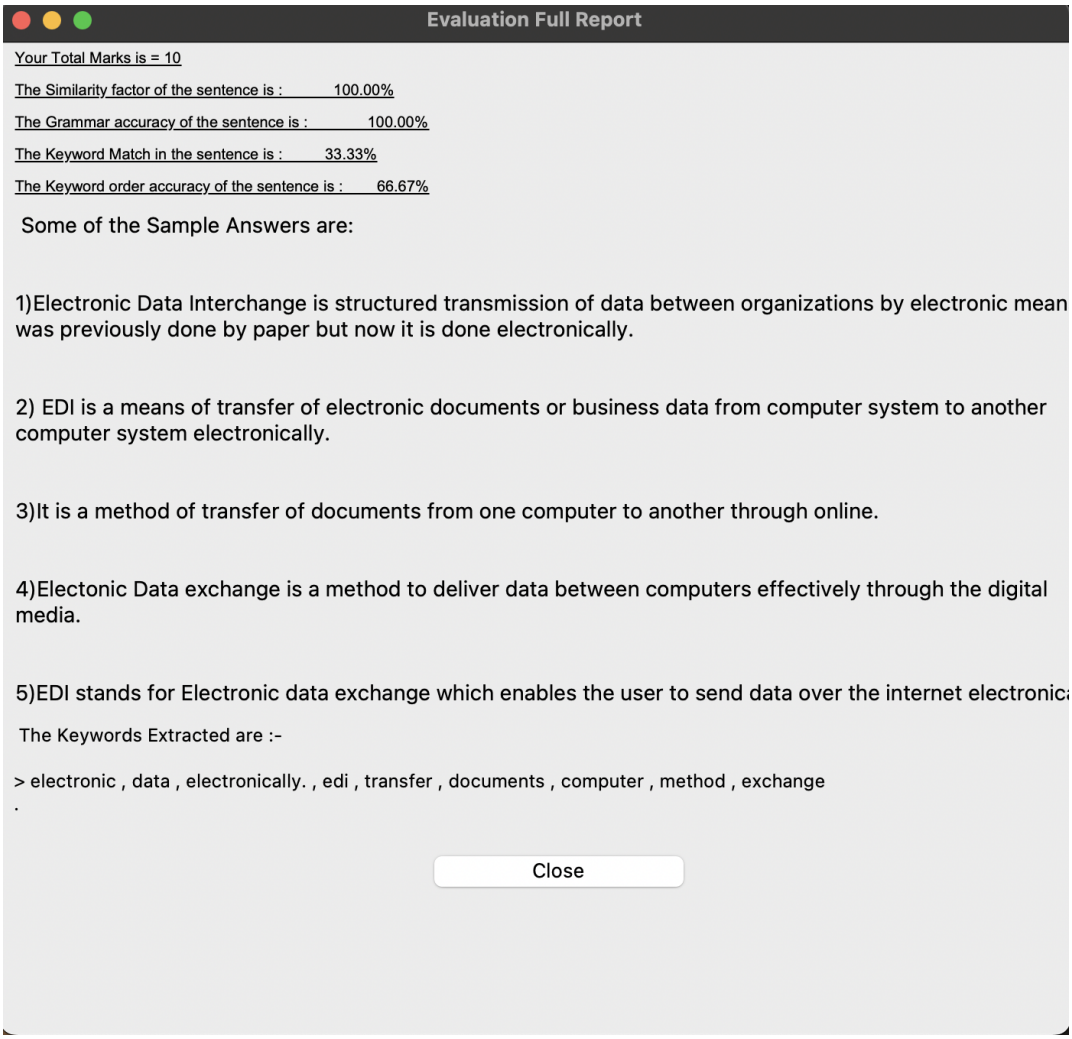

Figure 6.9 - Detailed Evaluation Report (Test Case 2)

6.3 Test case 3

This test case demonstrates the working of the system for the case where the answers consist of only keywords. This was for a question based on “E-Commerce”. The text from the uploaded answer key was correctly recognised as “ELECTRONIC EXCHANGE TRANSFER DATA COMPUTER METHOD” and fed to the evaluation portion of the code. The answer provided by the student had 6 keywords out of the 9 which were extracted from the answer key, showing a keyword match of 66.67%. The average similarity score of the answer was calculated as 93.20% after checking with all the sample answers given by the evaluator. Since the sentence was grammatically correct, the grammar score was 100%. The answer was also checked for the order in which the keywords occurred, based on which a keyword accuracy score of 66.67% was given. Overall, since the answer consisted of only keywords, a total of 3 marks was given to the student.

Table 6.3 - Results for Test Case 3

Criteria	Weightage provided by evaluator	Score assigned by Smart Assessment Tool
Keyword match	0.5	66.67%
Similarity score	1.8	93.20%
Grammar score	0.2	100%
Keyword accuracy score	0.1	66.67%

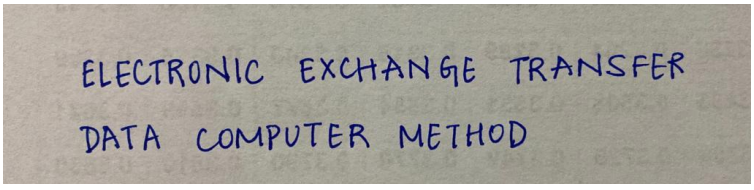

Figure 6.10 - Answer Script submitted by student (Test Case 3)

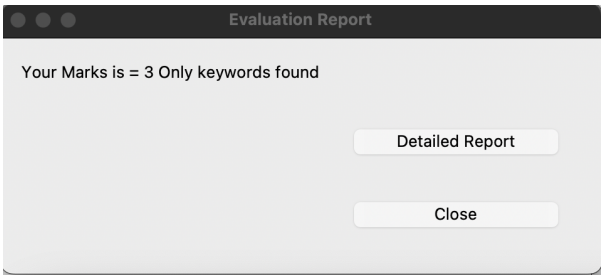

Figure 6.11 - Evaluation Report (Test Case 3)

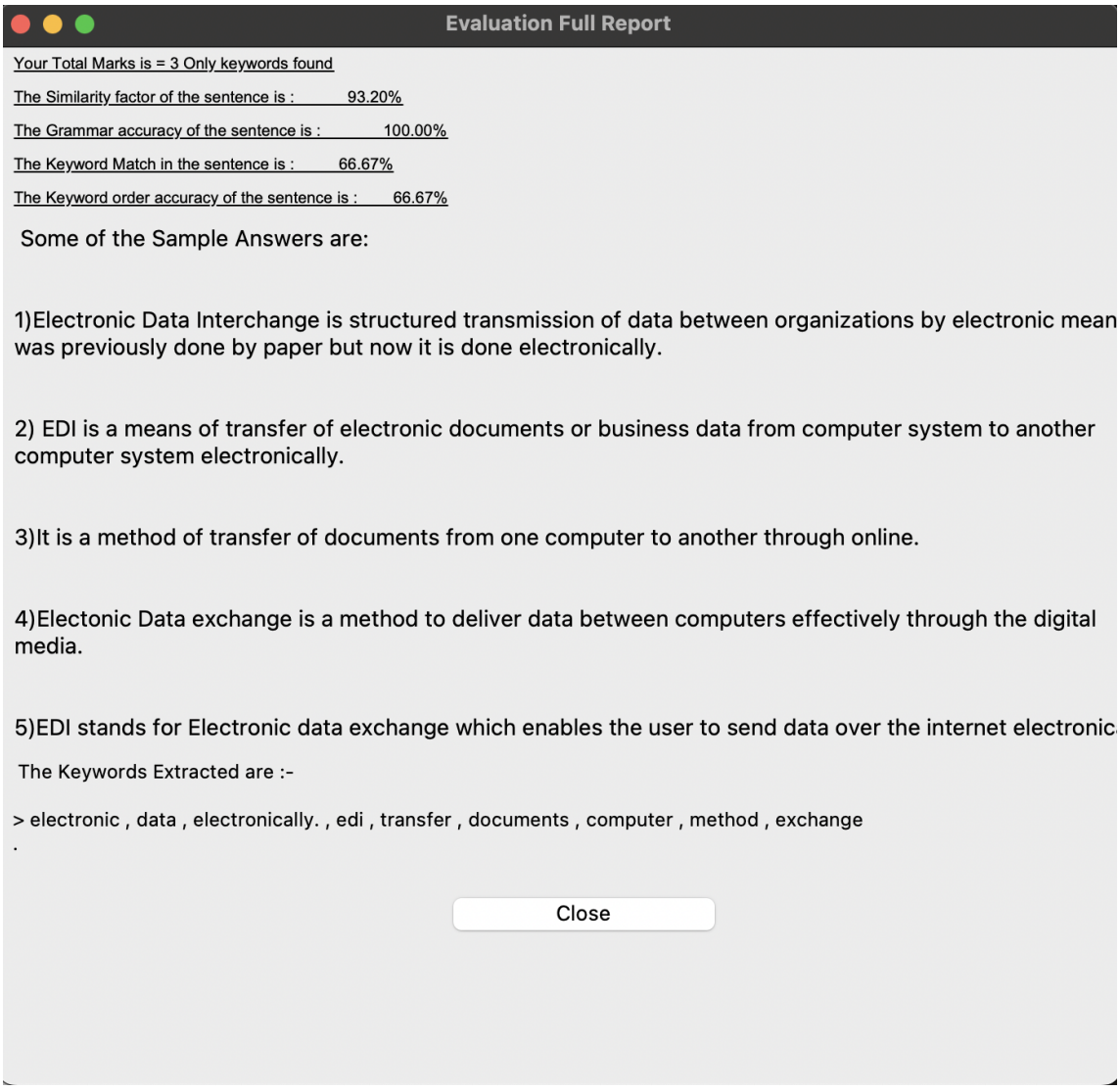

Figure 6.12 - Detailed Evaluation Report (Test Case 3)

6.4 Test case 4

This test case was for a wrong answer which contained a few keywords. Such an answer should not be awarded any marks.

This was for a question based on “E-Commerce”. The text from the uploaded answer key was correctly recognised as “MY NAME IS DATA EXCHANGE” and fed to the evaluation portion of the code. The answer provided by the student had only 2 keywords which were extracted from the answer key, showing a keyword match of 11.11%. Even though the answer has 2 keywords, the sentence written is logically wrong which cannot be considered for any marks. Therefore, the marks allotted is 0.

Table 6.4 - Results for Test Case 4

Criteria	Weightage provided by evaluator	Score assigned by Smart Assessment Tool
Keyword match	0.5	11.11%
Similarity score	1.8	7.20%
Grammar score	0.2	100%
Keyword accuracy score	0.1	0%

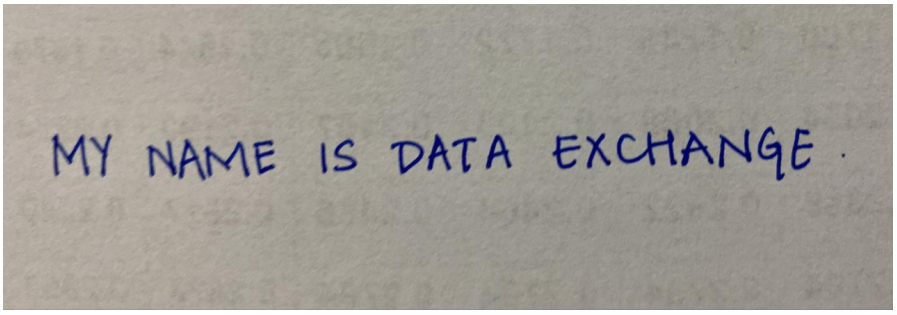

Figure 6.13 - Answer Script submitted by student (Test Case 4)

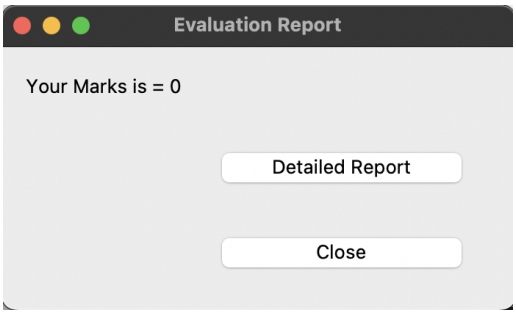

Figure 6.14 - Evaluation Report (Test Case 4)

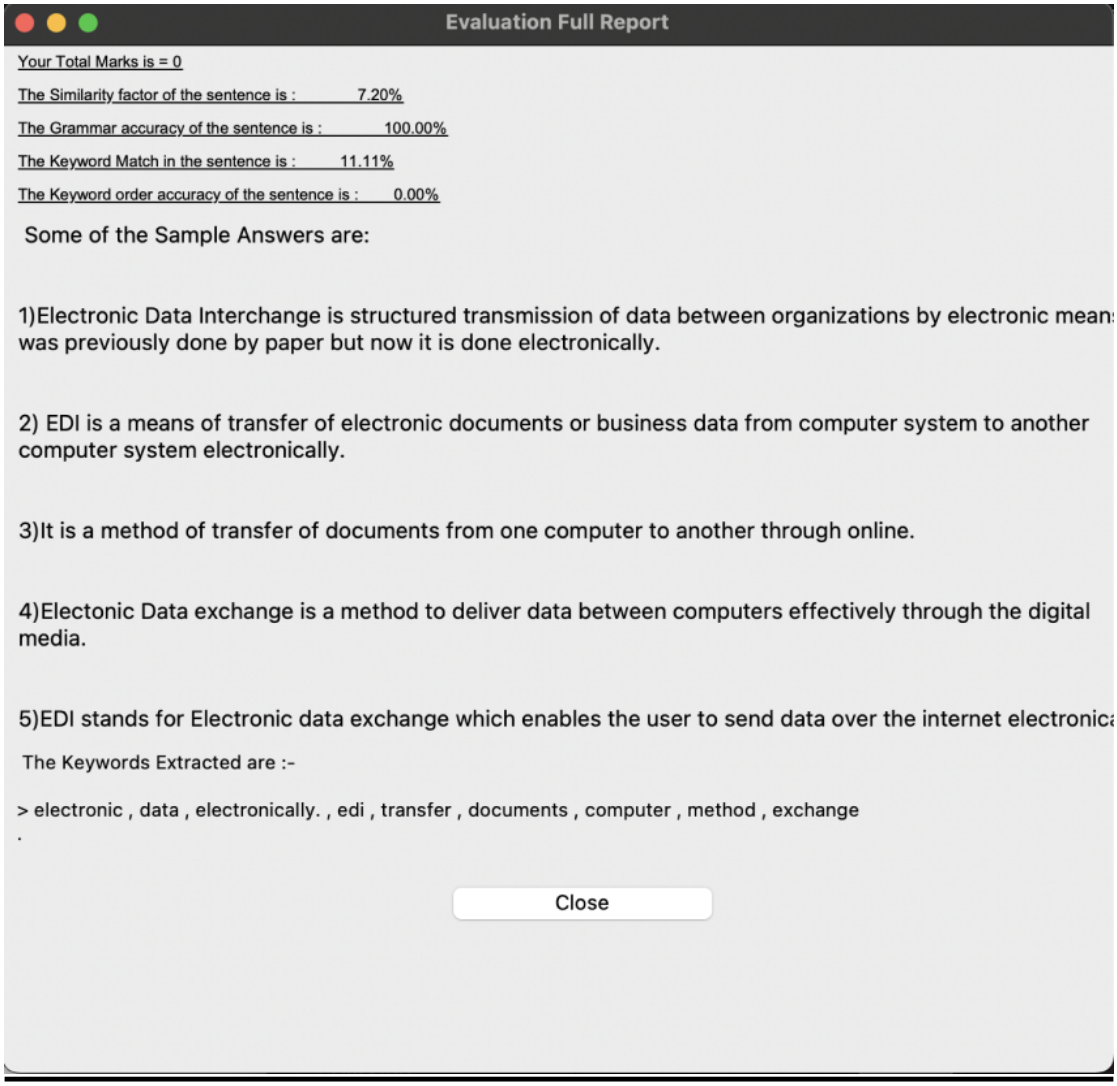

Figure 6.15 - Detailed Evaluation Report (Test Case 4)

Chapter 7

CONCLUSION AND FUTURE SCOPE

The development of an automated system for evaluating handwritten subjective exam papers using OCR technology and string matching algorithms has the potential to revolutionize the grading process. The system can reduce the workload of examiners and provide faster and more accurate evaluation. With the use of WordNet to handle variations in student responses, the system can ensure uniformity in grading and eliminate the possibility of biased evaluation.

In terms of future work, the accuracy of the system can be further improved by incorporating deep learning models such as Recurrent Neural Networks (RNN) which can handle more complex and nuanced responses. Additionally, the system could be extended to evaluate flow charts and diagrams in the provided answers, which would require the development of new algorithms to handle image recognition and evaluation. Furthermore, the system could be integrated with Learning Management Systems (LMS) to provide learners with immediate feedback on their performance and identify areas where they need to improve. Overall, this project has significant potential to streamline the grading process and improve the quality of education.

BIBLIOGRAPHY

[1] A. Nikitha, J. Geetha and D. S. JayaLakshmi, "Handwritten Text Recognition using Deep Learning," 2020 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), Bangalore, India, 2020, pp. 388-392, doi: 10.1109/RTEICT49044.2020.9315679.

[2] Bagaria, Vinal et al. "An Intelligent System for Evaluation of Descriptive Answers." 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) (2020): 19-24.

[3] Obaid, Ahmed & El-Bakry, Hazem & Eldosuky, M.A. & Shehab, Abdulaziz. (2019). Handwritten Text Recognition System based on Neural Network. International Journal of Advanced Research in Computer Science & Technology. 4. 72-77.

[4] Rahman, Md & Siddiqui, Fazlul. (2018). NLP-based Automatic Answer Script Evaluation. 4. 35-42.

[5] Rajeshwari S B and Jagadish S Kallimani. (2017). "Smart Online Assessment System for Descriptive Responses using Machine Learning Techniques". International Conference on Communication and Signal Processing, April 6-8, 2017, India

[6] Piyush Patil, Sachin Patil, VaibhavMiniyar, Amol Bandal, "Subjective Answer Evaluation Using Machine Learning",International Journal of Pure and Applied Mathematics | 2018 | Volume 118 No. 24.

[7] R.Parthiban, R.Ezhilarasi, D.Saravanan, ''Optical Character Recognition for English Handwritten Text Using Recurrent Neural Network,'' IEEE, 2020.

[8] Das, B. Sharma, S. S. Rautaray and M. Pandey, "An Examination System Automation Using Natural Language Processing," 2019 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2019.