

Intelligent Short Answer Assessment using Machine Learning



Rosy Salomi Victoria D, Viola Grace Vinitha P, Sathya R

Abstract: Education is fundamental for human progress. A student is evaluated by the mark he/she scores. The evaluation of student's work is a central aspect of the teaching profession that can affect students in significant ways. Though teachers use multiple criteria for assessing student work, it is not known if emotions are a factor in their grading decisions. Also, there are several mistakes that occur on the department's side like totaling error, marking mistakes. So, we are developing software to automate the evaluation of answers using Natural Language Processing and Machine Learning. There are two modules, in the first module, we use Optical Character Recognition to extract a handwritten font from the uploaded file and the second module evaluates the answer based on various factors and the mark is awarded. For every answer being entered, evaluation is done based on the usage of word, their importance and grammatical meaning of the sentence. With this approach we can save the cost of checking the answers manually and reduce the workload of the teachers by automating the manual checking process. The evaluation time is also reduced by using this software.

Keywords: Descriptive type answers, Grammatical Checking, Optical Character Recognition, Semantic checking.

I. INTRODUCTION

The correction of huge volumes of handwritten answers by hand is a difficult procedure that is assured to be tensed with mistakes. Programmed handwriting acknowledgment can radically cut miserable the period essential to correct huge amounts of data. Recognizing answers written by hand is a ground of investigating surrounding Artificial Intelligence computer visualization, and pattern appreciation. Online and offline are the kinds of handwriting acknowledgment scheme. Acknowledgment schemes are identified by either certain strokes or by one character or whole words.

II. EXISTING SYSTEM

Traditional paper correction process is time consuming and has many logistical activities. Onscreen Evaluation System is used to improve the traditional examination paper checking process. The student/candidate details are masked, and bar-codes are assigned for tracking. The answer scripts are scanned, digitized, encrypted, and stored in secure data centers.

Revised Manuscript Received on April 13, 2020.

* Correspondence Author

Dr. Rosy Salomi Victoria D.*, Associate Professor, Department of Computer Science and Engineering, St. Joseph's College of Engineering, Chennai, India. Email: drosysalomi@gmail.com.

Viola Grace Vinitha P, Department of Computer Science and Engineering, St. Joseph's College of Engineering, Chennai, India. Email: violagrace22@gmail.com

Sathya R, Department of Computer Science and Engineering, St. Joseph's College of Engineering, Chennai, India. Email: sathyaaraero@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

This helps to avoid tampering, misuse, fraudulent, reduces the chance of bias-based evaluation and reduces threat of errors and malpractices. These answer sheets are distributed electronically, enabling examiners (human interaction) to mark on-screen, and include comments and annotations.

III. LITERATURE SURVEY

There are several algorithms that have been proposed for handwriting recognition and conversion. Using a single technique for pre-processing, we cannot completely process the image [4]. However, after applying all these techniques it might not possible to achieve the full accuracy in a pre-processing system.

The scheme of a multiple level perceptron constructed design classifier for identifying the handwriting Bangla digits by means of a 76 component feature array is suggested [3]. Evaluation of subject answer checking isn't a new thought. It has been in the works for a decade and a half. A large number of techniques were experimented with to solve the problem efficiently: Natural Language processing, Latent Semantic Analysis, Generalized Latent Semantic Analysis, Bayes theorem, K- nearest neighbour, etc.

An analysis on sentiment classification model with preprocessing techniques to remove stop words was discussed [6]. Stop words such as "the", "and", "is" were removed. A model that takes short answers as input and constructs RDF sentences is developed [11]. Using ontology student answer is mapped with model answer they confirmed same keyword is not needed to get full marks .Even the synonym of the keyword also get marks. For One-Sentence answer, matching with the model answer is done by taking into consideration both lexical structure and synonyms. However, there is a restriction imposed on the length of the sentences of the answer.

A model to evaluate subjective papers using Semi-Automated Evaluation technique is created [5]. For that first they create Question base which contains question type, sub type, question and marks. Then Answer base is created with model answer. Evaluated answer is mapped using hash index which referred as question number. The student answer is evaluated by considering semantic meaning and length of the sentence.

IV. PROPOSED SYSTEM

This paper deals with converting handwritten font to text using Neural Network based Handwritten Character Recognition system (OCR) with feature extraction and evaluating the answer scripts using wordnet.

Optical Character Recognition involves the detection of handwritten text on images and conversion of the handwritten text to encoded text that the computer can easily understand. Then the encoded text will be given as input to the evaluation Model.

The basic grammar and spellings are checked using a grammar check model. Through this model, the percentage of correctness words can be found. Stop words are removed

and stemming, lemmatization is done. The comparison of the existing system and proposed system is shown in the Table I.

Answers are checked with keywords, which will be given as a dataset. Then the total mark is calculated based on the correctness of the sentence and number of keywords matched.

Table- I: Existing System Versus Proposed System

EXISTING SYSTEM	PROPOSED SYSTEM
Involves human interaction for evaluating the answer scripts.	Completely automated.
The possibility of emotions being reflected in their correction.	No such suspicion.
More human resources are needed.	Less human resource.
Lengthy process.	Takes less time.
Logistical cost and travel management of evaluators is high.	Logistical cost and travel management of evaluators is eliminated.

This conversion is done using python Jupiter Notebook. Figure 1 illustrates the process flow.

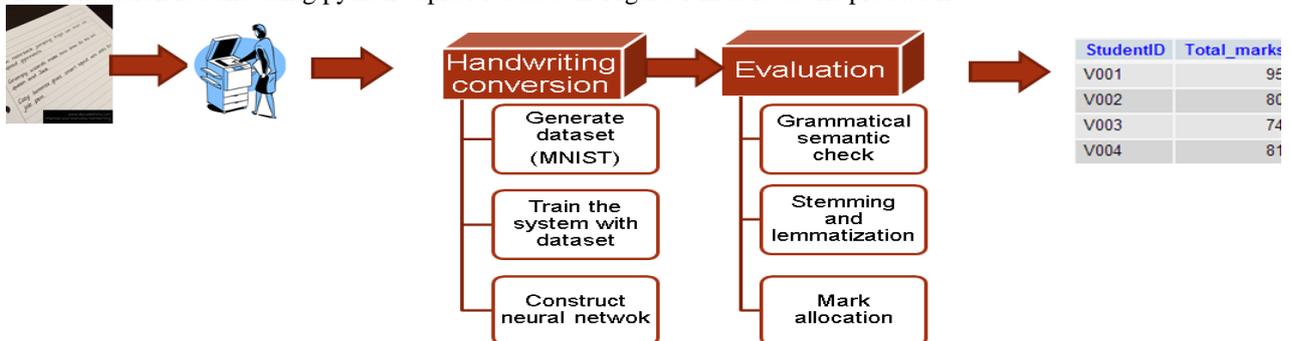


Fig .1. Architecture diagram

The answer sheet from the student is scanned and handwritten text is converted to machine encoded text, the converted text is evaluated and marks are calculated.

A. Extract text from image

The procedures used in recognizing of text can be separated into three groups: Image Pre-processing, Segmentation and Feature Extraction. The abstraction of feature is needed for handwriting to text conversion for recognizing of character. A computer compares objects within an image to letters already stored within its software. OCR converts the document into a black and white version. The scanned-in image is analyzed for light and dark areas, where the dark areas are identified as characters/text that need to be recognized/converted and light areas are identified as background.

1) Image pre-processing

It is vital in the recognizing pipe for right text estimation and alteration. Its techniques contain removing of sound, segmenting of pictures, crops, scales etc. The recognizing of text scheme agrees input as a picture scan. Arithmetical

imprisonment and alteration of a picture announces sound creates it difficult to classify the actual text. Therefore it is necessary to decrease sound by conserving the strokes of the text, because it is essential for right grouping of text.

2) Segmentation

Segmentation is breaking the whole image into subparts to further process them. A arrangement of text is divided as a subimage of a single text. Segmentation of characters is faster than any conventional method in which all the characters from the text are segmented by connected component processing only. Segmentation algorithm recognizes line by line and word by word detection. Efficient algorithm is used for segmentation.

3) Feature extraction

The structures of input information are the quantifiable possession of interpretations, which is used to examines or categorize the information. Rather than trying to recognize full letters,

feature extraction is when a computer detects certain features (lines and loops, for example) and understands that they represent a character. The letter 'F', for instance, will be picked up by the software whenever it detects two horizontal lines joined by vertical line on the left side. The output obtained after image processing, segmentation and feature extraction is shown in the result.

1. Percentage of Grammatical Correctness = (Number of correct words / Total number of Words in answer)*100.

2. Percentage of keyword match = (Number of Keywords matched / Total Number of Keywords)*100.

3. Percentage of correctness of the sentence = (percentage of grammatical correctness + percentage of keyword match)/2

4. Mark calculation = (Percentage of correctness of the sentence/100)*Total Marks. **B. Answer Evaluation**

1) *Grammar check*

A grammar checker function will find each word in a sentence, look up each word in the dictionary, and then

Table- II: Mark Allocation

Question No	Correct mark	Manual Evaluation Mark	System Evaluation mark
1	4.8	4.7	4.8
2	4	4	4
3	2.4	2	2.5
4	3.1	4	3
5	4	4.5	4
6	5	3	5
7	4	3.5	4
8	3	3.2	3

2) *Remove stop words*

The correct sentence from the Grammar check will be sent to stop words removing function. The function will remove stop words like and, but, then, is, etc., from the correct sentence. It will return the array containing keywords only.

C. Stemming and Lemmatization

1) *Stemming*

Stemming is the process of reducing inflection (prefix,suffix) in words to their root forms such as mapping a

group of words to the same stem even if the stem itself is not a valid word in the Language. It reduces a word to its word stem that affixes to suffixes and prefixes or to the roots of words

known as a lemma. Stemming will remove suffix and prefix from the word.

attempt to parse the word into a form that matches a grammar.

Using various rules, the program can then detect various errors, such as tense, spelling mistakes, etc. Number of mistakes in the text will be found. Using this percentage of correctness of the sentence will be calculated and then word with mistake will be changed to correct word.

Table II contains marks by manual evaluation and machine evaluation. It used to compare the marks between human and machine evaluation. Figure 2 blue lines shows marks given by human and the Red line shows marks given by machine. For question 1, the correct mark will be 4.8, by human evaluation it gets 4.7 marks assigned, whereas by using machine we can get accurate answer marks as 4.8. For question 4, the correct mark is 3.1, human evaluates this question and gives 4 marks, whereas machine evaluation gives 3 marks, which is closer to correct mark.

2) *Lemmatization*

To extract the proper lemma of word, it is necessary to look at the morphological analysis of each word. Rather than stemming, lemmatization provides lemma which has a dictionary meaning. The key to Lemmatization is linguistics.

D. WordNet

WordNet is a lexical database .It contains semantic relations between words in more than 200 languages. It links words into semantic relations including synonyms, meronyms. The synonyms are grouped into sets called synsets with short definitions and usage examples. Using Synsets, the meaning of the words in the answer and keywords are found. Check how many keywords present in the answer by calculating distance.

E. Mark Calculation

Using the percentage of correctness of the sentence, number of keywords presents in the answer and the length of the answer mark are calculated. Formula:

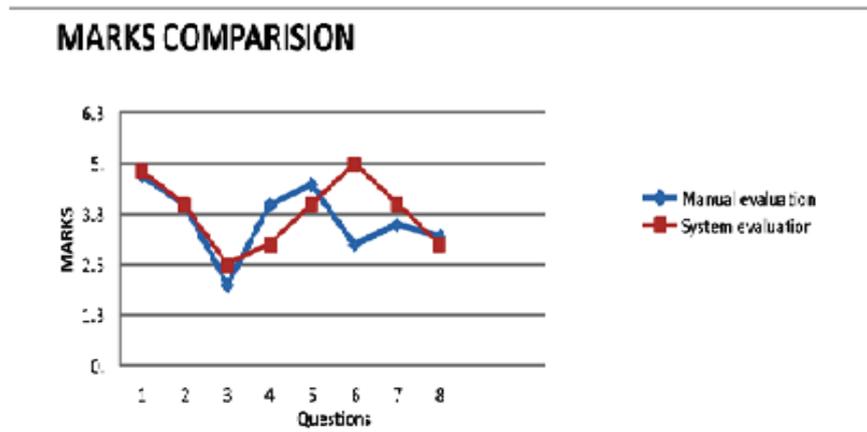


Fig. 2: Relationship of system evaluation and manual evaluation

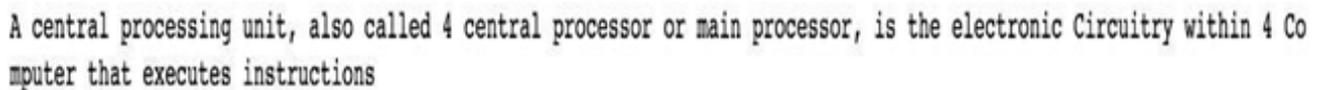
V. RESULT AND DISCUSSION

The results, using the above proposed algorithm are brought up in this section. The handwritten text, given as input to the handwriting conversion module is extracted and converted to machine encoded text using Optical Character Recognition Algorithm. The machine encoded text is given as input to the evaluation module. This module evaluates the answer based on grammatical meaning of the sentences; number of keywords matched and gives marks as output. Figure 3 is the output obtained after image processing, segmentation and feature extraction. Figure 4 show the count of mistaken words in the sentence and change the

sentence to correct one. The machine encoded text is checked and the number of mistaken words is counted by comparing with the English dictionary and is corrected to grammatically meaningful sentence. The stop words like 'and', 'but', 'is', 'the' are removed from a sentence as shown in Figure 5. The output after lemmatization is depicted through Figure 6. Lemmatization process removes the prefix and suffix meaningfully. This is done to get the root word of each word. Figure 7 depicts the output after using WordNet. The number of keywords that matches with the answer is counted. The mark is calculated based on the number of keywords matched. The calculated mark is displayed as shown in Figure 8.



(a)



(b)

Fig. 3. (a) Image of the handwritten text given to the handwriting conversion module. (b)Image obtained as machine encoded text.

A central processing unit, also called a central processor or main processor, is the electronic circuitry within 4 Co mputer that executes instructions

Number of wrong words in the sentence: 0

Number of wrong words in the keywords: 4

Correct sentence is:

A central processing unit, also called a central processor or main processor, is the electronic circuitry within 4 C omputer that executes instructions

Fig. 4. Grammar Check

['A', 'central', 'processing', 'unit', 'also', 'called', 'a', 'central', 'processor', 'or', 'main', 'processor', 'is', 'the', 'electronic', 'circuitry', 'within', '4', 'Computer', 'that', 'executes', 'instructions']
['A', 'central', 'processing', 'unit', 'also', 'called', 'central', 'processor', 'main', 'processor', 'electronic', 'circuitry', 'within', '4', 'Computer', 'executes', 'instructions']

Fig. 5. Stop Words Removal

['A', 'central', 'process', 'unit', 'also', 'call', 'central', 'processor', 'main', 'processor', 'electron', 'circuit', 'ri', 'within', '4', 'comput', 'execut', 'instruct']

Fig. 6. Lemmatization

Number of keywords matched: 7

Fig. 7. Keywords Match

The mark is: 4.9

Fig. 8. Display marks

An analysis between manual evaluation and system evaluation is done and shown in the previous section.

VI. CONCLUSION AND FUTURE WORK

We have developed a software using an Optical Character Recognition and Evaluation algorithm which evaluates theoretical answers and gives marks according to the keyword match. Optical Character recognition is used to extract text from the uploaded file. Evaluation algorithm is used to assess the answer paper and award marks based on grammatical meaning of the sentence, length of answer, number of important keywords match. This software reduces manual work and saves time with fast evaluation of answers and our software has produced good intermediate results.

The proposed system can be extended to work on degraded text or broken characters. Recognition of digits in the text, half characters and compound characters can be done to improve the word recognition accuracy. Font Independent OCR can be used to identify different font styles used by students/candidates.

The model can be trained in different languages. From this, we can create a dataset of different handwritten languages.

Hence an answer with a non-English language can also be evaluated. And the System will also evaluate the overwritten alphabets and other words with absolute accuracy.

The model can also be trained in such a way that it can check the complete paper instead of a single answer. Hence the model will evaluate the answers according to the answer number provided in the answer sheet. The model can be trained to evaluate diagrams and provide the marks accordingly. Thus the model with the user interface would need the user to upload the answer script which will be evaluated by the system and the marks will be stored in the database.

REFERENCES

1. Marina Nakano, Emi Masuda, Masaru Kamada, "A Structure Editor for the English Language", in *Proc. Int. Conf. on Electrical,*

- Electronics, and Optimization Techniques (ICEEOT)*, 2016, pp. 155-161.
2. N. VenkateswaraRao, A. Srikrishna, B. RaveendraBabu, G.R.M. Babu, "An efficient feature extraction and classification of handwritten digits using neural networks", *Int. Journal of Computer Science Engineering and Applications*, vol.1, 2011, pp. 47-56.
3. Subhadip Basu, Nibar Das, Ram Sarkar, Mita Nasipuri, "An MLP based Approach for Recognition of Handwritten 'Bangla' Numerals", in *Proc. Indian International Conference on Artificial Intelligence*, 2005.. pp. 407-417.
4. K. Gaurav and Pradeep Kumar Bhatia, "Analytical Review of Preprocessing Techniques for Offline Handwritten Character Recognition", *Int. Journal of Advances in Engineering Sciences*, vol.3, no. 3, 2013, pp. 14 -22.
5. Chhanda Roy, Chitrita Chaudhuri, "Case Based Modeling of Answer Points to Expedite Semi-Automated Evaluation of Subjective Papers", in *Proc. Int. Conf. IEEE 8th International Advance Computing Conference (IACC)*, 2018, pp. 85-90.
6. Ms. Kranti Viithal Ghag, Dr. Ketan Shah, "Comparative Analysis of Effect of Stopwords Removal on Sentiment Classification", in *Proc. Int. Conf. IEEE International Conference on Computer, Communication and Control (IC4-2015)*, 2015.
7. Jentrisi Priyatno, Moch Arif Bijaksana, "Clustering Synonym Sets in English WordNet", in *Proc. 7th Int. Conf. on Information and Communication Technology (ICOICT)*, 2019, pp.1-4.
8. G. Abdul Robby, Antonia Tandra, Imelda Susanto, Jeklin Harefa, Andry Chowanda, "Implementation of Optical Character Recognition using Tesseract with the Javanese Script Target in Android Application", *Procedia Computer Science*, vol. 157, 2019, pp. 499-505.
9. S.B. Patil, "Neural Network based bilingual OCR system: Experiment with English and Kannada bilingual document", *Int. Journal of Computer Applications*, vol. 13, no. 8, 2011, pp. 6-14.
10. Rahul Pramanik, Soumen Bag, "Shape decomposition-based handwritten compound character recognition for Bangla OCR", *Journal of Visual Communication and Image Representation*, vol. 50, 2018, pp. 123-134.
11. Jie Mei, Aminul Islam, Abidrahman Moh'd, Yajing Wu, Evangelos Milios, "Statistical learning for OCR error correction", *Information Processing and Management*, vol. 54, Nov. 2018, pp. 874-887.
12. V.Senthil Kumaran, A.Sankar, "Towards Automated System For short-answer assessment using ontology mapping", *Int. Arab Journal of e-Technology*, vol. 4, no. 1, 2015, pp. 17-24.
13. Mohammed Alshahrani, Spyridon Samothrakis, Maria Fasli, "Word mover's distance for affect detection", in *Proc. Int. Conf. on the Frontiers and Advances in Data Science (FADS)*, 2017, pp. 18-23.

AUTHORS PROFILE



D. Rosy Salomi Victoria received her Master of Engineering degree in Computer Science and Engineering from Sathyabama University, Chennai in 2003. She obtained Ph.D. from Anna University, Chennai in 2015. She is an Associate Professor at St. Joseph's College of Engineering. Her area of interest are Computer Networks, Object Oriented Analysis and Design, Distributed systems and machine learning.



P. Viola Grace Vinitha is a final year student of Bachelor of Engineering at St. Joseph's College of Engineering, Chennai. She is the 3rd topper in the Department of Computer Science and Engineering. Her area of interest are Machine Learning, Object Oriented Programming, Programming and Data Structures.



R. Sathya is a final year student of Bachelor of Engineering at St. Joseph's College of Engineering, Chennai. She is the 2nd topper in the Department of Computer Science and Engineering. Her area of interest are Machine Learning, AutoML, Programming and Data Structures.