

Automatic Assessment of Handwritten Essays Using LCS

Reshma.V.M^[1], Linda Sara Mathew^[2]

Department of Computer Science and Engineering

Mar Athanasius College of Engineering

Kothamangalam, Ernakulam

Kerala - India

ABSTRACT

Essays are used to measure student's knowledge from previous times itself. The proposed method is to develop an automated system to evaluate handwritten student essays. The aim of proposed system is to evaluate the essays similar to manual evaluation of essays. The handwritten essays are evaluated against to the study material in order to calculate the grade and the calculated grade is displayed. In order to make a single document of study material from multiple study material, every sentence from different documents are compared and combined. The similarities between words are calculated to get the similarity between sentences. The texts of essays are extracted from the handwritten essays of students. The extracted texts are used to calculate the grade for the corresponding essay. Longest Common Subsequence (LCS) and Bigram Co-occurrence are combined to calculate the grades of the essays.

Keywords:- Longest Common Subsequence.

I. INTRODUCTION

Essays are used for measuring student's knowledge. For the human beings, essay writing and evaluation is a demanding task. Essays are evaluated according to the concepts and contents present in it. Essay evaluation can perform more faster way if it is possible to develop an automated system for evaluating the essays.

The proposed system introduces evaluation of handwritten essays. Handwriting recognition became most challenging research areas in image processing and pattern recognition. By reading different study materials, the students can increase their knowledge. Their knowledge can be measured by conducting exam. The exams are conducted for measuring the semantic similarity of the student's essay and the materials from where the student studied for that particular exam.

The proposed system includes three phases. First phase is the construction of the reference study material. In the first phase from different study materials one single study material is constructed which is used for the assessment of the handwritten essay. The second phase is text extraction from the handwritten essays. In the proposed system, the input to the assessment of the essay is a scanned copy of human written essay. From this image, the texts are extracted. The third phase in the proposed system is the grading of the essay. The extracted text from the scanned

image is compared with the constructed reference study material. Finally the grade is calculated and displayed.

Both the first phase and the second phase includes the comparison of sentences. While performing the comparison, meaning of the words present in the documents are considered, so that the assessment will make accurate results. Longest Common subsequence (LCS) is used for the comparison of essay without losing its contents and concepts present in it.

In this paper, we are describing a method to evaluate handwritten essays automatically. If it is possible to make an effective system by this method, it can be used to replace the manual evaluation for certain extend.

II. BACKGROUND

Assessment of essays is a demanding job for human beings. The methods of automatic assessment of documents are introduced in order to reduce the cost of the grading of those documents. Approaches based on the surface features and content have been introduced early before itself [1]. The system uses the method called Latent Semantic Analysis (LSA) for the essay content evaluation. Previously, the method LSA is made for information retrieval. Later it started to use for the essay evaluation in many applications. LSA will help to compare the similarity of the concepts and the contents between essays. In the

experiment the correlation between the scores given by the system and human grader varied from 0.78 to 0.82.

Another commercial software is Intelligent Essay Assessor (IEA). The IEA can be used to accurate assessment of the essay which performs similar to a well trained human being. It was using in many real life applications like Army's, Air force's, general staff colleges etc. This IEA is well suited for the distant education purposes. It gives a feedback for the essay within very few times [2].

In 1966 Ellis page developed Project Essay Grader (PEG). The purpose was to make an assessment of essays in large scale. It gives a very good quality for the essay evaluation procedure. An advantage of PEG is, the calculate grades are comparable with the human evaluated grades. It produces almost similar grades for the essay [3]. But, the drawback of the PEG is that , it does not give any importance to the contents present inside in the essay.

Another research tool is BETSY. BETSY is developed by Lawrence M Rudner. The other type of essay evaluation is implemented is AES. AES uses the concept of Bayesian theorem [4]. It has a number of applications. Some of them are identifying spam mails, sorting of resumes for job applications etc.

Another difficult and challenging job is the text extraction from the images. The process becomes more difficult because of the difference present in the images like scripts, styles, fonts, size, colour etc [5]. At first, image pre-processing is performed which include image gray scaling, noise removal, discontinuity removal, dot removal etc from the images. The next phase is the text segmentation of each character from the image. Finally, the separated character is recognized using the neural network and stored.

III. PROPOSED SYSTEM

The proposed system introduces automatic evaluation of handwritten essays. Handwriting recognition became most challenging research areas in image processing and pattern recognition. By reading different study materials, the students can increase their knowledge. Their knowledge can be measured by conducting exam. The exams are conducted for measuring the semantic similarity of the student's essay and the materials from where the student studied for that particular exam.

For an exam, there is a chance to study more than one study material for the same topics. In this scenario it requires to compare the students essay with all these study materials under their study. But this will take more time and the evaluation will not finish quickly. In order to make

the comparison of essays easier, it is better to make a single study material from all these study material. Sentences from different study materials are compared sentence by sentence by checking its semantic similarity. Longest common subsequence occurrence is used for this purpose. Sentences from different study materials are compared sentence by sentence.

Each and every sentence from both the study materials are compared semantically. For that each word are compared semantically. In order to compare each word of two different sentences synonyms of those words are compared. The synonyms of words are obtained from the synonym set of Microsoft word application. With this document the handwritten essays are compared and grade is evaluated. Before that, the texts that are appearing in the handwritten essays are extracted using image processing techniques.

The grades are calculated using the method Longest Common Subsequence (LCS). The Longest Common Subsequence is one method to use which can use with a set of reference translation and candidate translation. The longest occurring similar words from the two different sentences are identified by longest common subsequence method. It is done by considering the takes into account sentence level structure similarity of the considered two sentences.

The proposed system includes three phases. The first phase is summarization of the study materials. The second phase is the extraction of the texts from the handwritten student essays. The third phase is the essay grading phase.

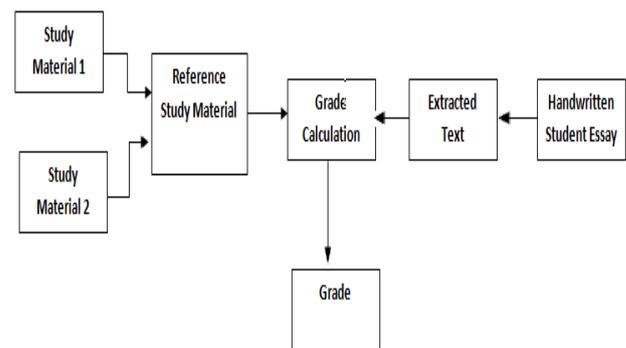


Fig. 1 Proposed system

A. Summarization of The Study Material

In the proposed system two study materials are using. From those two study materials, one single study material is constructed. It is referred as the reference study material.

This material is used in the third phase of the proposed system, to evaluate the grade for the student written essay. Each sentence from the two documents is compared for constructing the new study material. For a single topic under study, there may be more than one study materials. In those study materials the same data may be containing but the way of writings may be different.

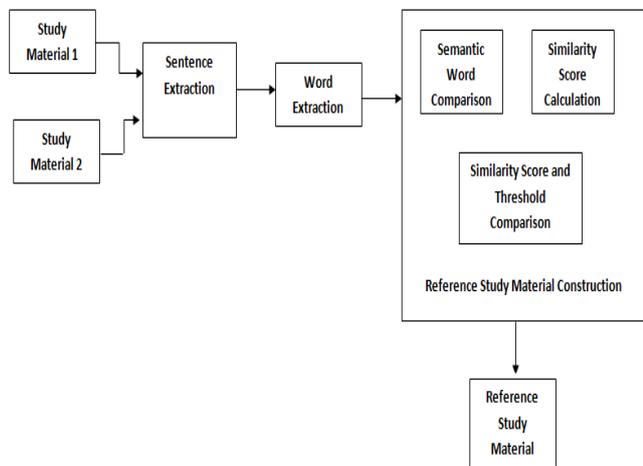


Fig. 2 Reference study material construction

The reference study material is made by the checking the semantic similarity of the two study materials. So if the semantic meanings of the sentences are not considered, the document summarization will be meaningless. For avoiding this problem, the semantic comparison of the sentences is made. If the two comparing sentences are semantically similar, then any one of them is considered to the newly created reference study material. Similarly if the comparing strings are not semantically similar, then both sentences will add to the reference study material.

B. Text Extraction from Handwritten Student Essay

In the field of image processing and pattern recognition, handwritten recognition became the most fascinating and challenging research area.

The first stage is image pre-processing. In the pre-processing step, a clear binary image is constructed from the input image of the handwritten essay by removing the noises. This step will help to increase the performance of the text recognition task. Gray scaling is the first phase in the pre processing step. It will help to remove the noises and disturbances present in the scanned image of the student’s essay. After that line removal and discontinuity removal are performed. Then dot removal is performed. The next phase is the text localization and the text segmentation. Both of these stages will help to separate

each character from the entire word present in the image of the essay. By applying this method, the whole sentences comes in the essay are converted into single character image. Text localization is done for every single character. This will help to separate each character from the entire word.

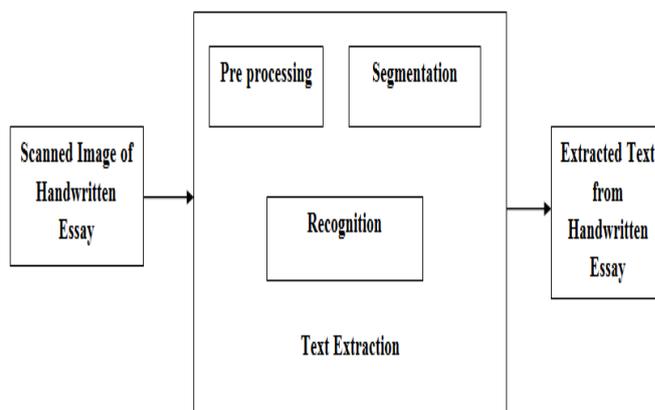


Fig. 3 Text Extraction from handwritten essay

The final stage is text recognition. In the text recognition stage, the segmented characters are thinned and scaled. After this, each character is compared with the stored neural network. From this the more nearest match is identified and stored.

C. Grading of Handwritten Student Essay

The system takes the summarized study material as input. The same document is used for calculating the score and this value is stored. In the analysis phase the student essay obtained from the second phase will take as the input. The study material and the essay are compared line by line and score is calculated. This score is compared with the previously calculated score and the grade is assigned.

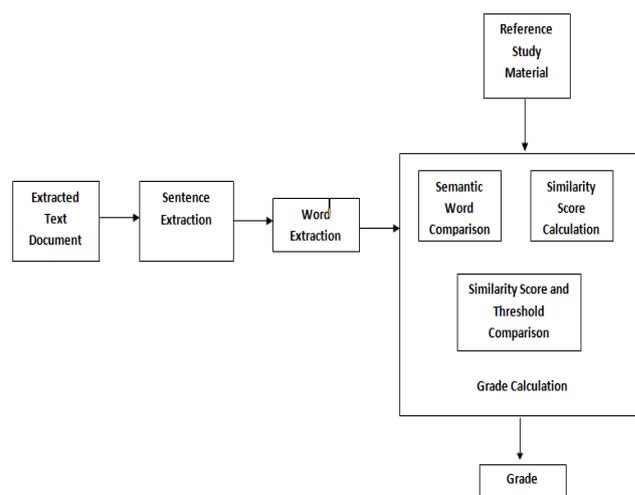


Fig. 4 Grade Calculation

The grades are calculated using by the approach called Longest Common Subsequence. The method is based on longest common subsequence (LCS) between a candidate translation and a set of reference translations. The candidate translation is a sentence of words from where it is trying find out the similar words by comparing with the reference translation. The term reference translation is a sentence of words from where comparison is made. Longest common subsequence identifies the longest occurring common words of given two sentences by using the semantic similarity. The proposed system is trying to find out the similar words by checking the meaning of the words present in the sentence.

D. Longest Common Subsequence (LCS)

Let A and B are two sentences consisting of words in which A and B can be represented as $A = (a_1, a_2 \dots a_m)$ and $B = (b_1, b_2 \dots b_n)$. For each $i = 1$ to m and $j = 1$ to n , a_i and b_j indicates the words present in the sentences. Then the longest common subsequence (LCS) of the two sentences A and B is the longest sequence C which is a subsequence of both sentences A and B.

Each and every sentence from both the study materials are compared semantically. For that each word are compared semantically. In order to compare each word of two different sentences synonyms of those words are compared. The synonyms of words are obtained from the synonym set of Microsoft word application. With this document the handwritten essays are compared and grade is evaluated.

By applying the LCS, a score is calculated for comparing each sentence in the extracted text of the handwritten essay. The score value obtained for each sentences in the essay are added to get the final score of the essay. The final score obtained for the essay is compared with the threshold score value. According to that, a grade is assigned for the essay. If the comparing two documents of sentences are exactly same, then the calculate final score value for the essay will be maximum. This maximum value of the final score is taken as the threshold score value.

IV. RESULTS

The following table gives the relation between the number of lines and the time required for combining the texts in the first and third phase.

TABLE I
RELATION BETWEEN NO: OF LINES AND THE TIME REQUIRED TO COMBINE SENTENCES

No: of Lines	Time Required to Combine Sentences (s)
1	184.54
2	342.87
3	735.46
4	1394.55

Fig. 5 shows the graphical representation of the details described in the TABLE 1. From the Graph it is clear that the time required to combine sentences is increasing according to the number of the lines increases.

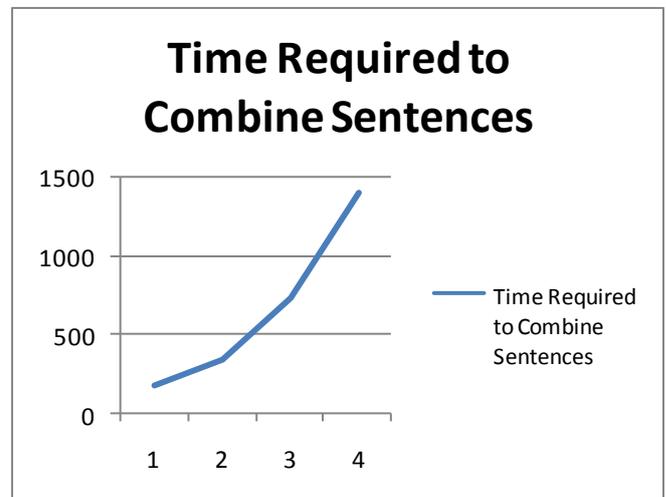


Fig. 5 Relation between No: of Lines and Time Required to Combine Sentences

Since, the texts are semantically compared it requires to consider each and every sentence present in both documents. Thus the construction of the reference study material takes little more time than expected.

The following diagrams indicate different performance measurements of the neural network data and its training. The

Fig. 6 and Fig. 7 indicate the validation performance for the neural network and the training state of the neural network. Next gives the detailed description is of the neural network confusion

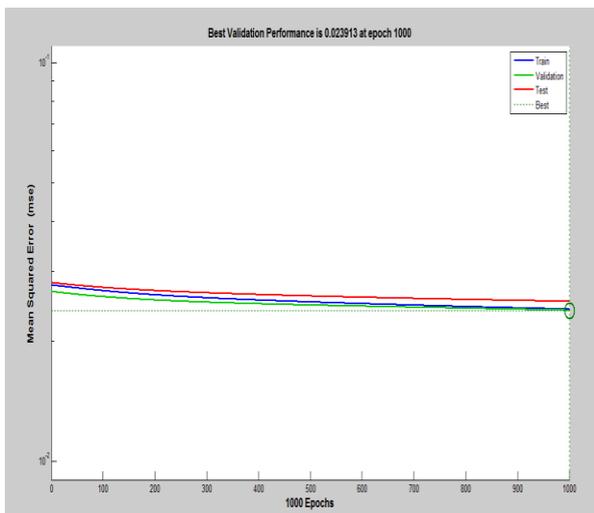


Fig. 6 Validation Performance of Neural Network

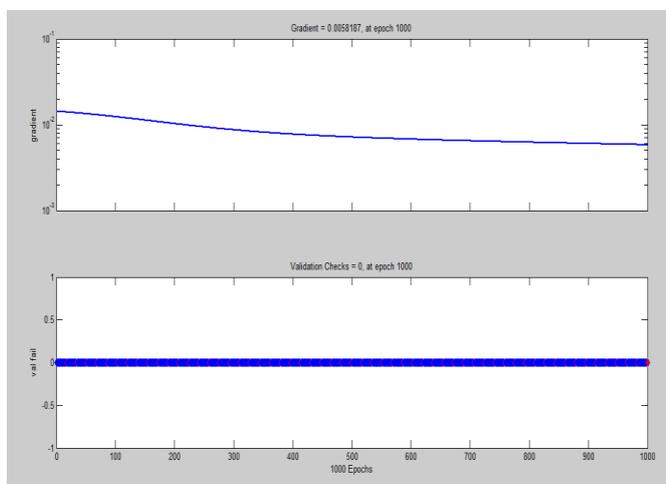


Fig. 7 Neural Network Training State

V. CONCLUSIONS

Essays are used for measuring student knowledge. Essays, is a demanding task even for a human. From previous times itself essays are evaluated according to conceptual content present in that. If there is an automated system for evaluating the essays, the evaluation can do faster, especially for evaluating handwritten essays which will perform similar to the manual human evaluation. The proposed system introduces evaluation of handwritten essays. Handwriting recognition is one of the most challenging research areas in field of image processing and pattern recognition. Student's knowledge is acquired by reading the study materials and the student's knowledge can be measured as the degree of semantic similarity between the essay and the parts of the textbook covering

the topic under study. Students can refer more than one study material to study. The proposed system includes three different modules. The first module is text extraction from handwritten student essays. The next phase is and the final phase is grading of the students essays. The proposed system evaluates the essays successfully.

If there is an effective automated system for evaluating essays, it will be more helpful in the education field. Essays can evaluate more easily without requiring manual effort. The proposed system extracts texts from the image, but the texts extracted are not accurate always. The system results wrong words instead of giving the actual word present in the handwritten essay. In this case, the actual grade calculation is not possible. The grade calculated will not be effective. If it is possible to add a feature like automatic spell correction of the words along with the text extraction of the system, the more accurate texts can be extracted and thus more accurate grade also can calculate.

REFERENCES

- [1] Tuomo Kakkonen and Erik krisutinen, "Automatic Assessment of the Content of Essays Based on Course materials", Dept of Computer Science and Engineering,
- [2] Lynn Streeter and Others, "The Credible Grading Machine: Automated Essay Scoring in the DOD", Dept of Computer Science and Engineering,
- [3] Semire Dikli, "An Overview of Autoamted Scoring of Essays", Turkish Online Journal of Distance Education – TOJDE, January 2006, Vol: 7, No:1, Article:5
- [4] Tuomo Kokken and Others, "Comparison of Dimension Reduction Methods for Automated Essay Grading", Educational Technology & Society. 11(3), Pageno: 275-288, 2008
- [5] Paraag Agarwaal and Rohit Varmma, " Text Extraction from Images", IJCSET, April 2012, Vol 2, Issue 4, 1084-1087
- [6] Neeta Nain Subhash Panwar, " Handwritten Text Recognition System Based on Neural Network", Malavya National Institute of Technology, Jaipur
- [7] Chin-Yew Lin and Franz Josef Och, " Automatic Evaluation of Machine Translation Quality Using Largest Common Subsequence and Skip Bigram Satictics", Information Science Institute
- [8] Md Monjurul Islam and A Sm Latiful Hoque, " Automated Essay Scoring Using Genaralized Latent

- Semantic Analysis”, Journal of Computers, Vol 7, No 3, March 2012
- [9] Chin Yen Lin and Eduard Hony, “ Automatic Evaluation of Summaries Using N-Gram Co-occurrence Statistics”, Proceedings of HLT-NAACL 2003, PP 41-78
- [10] Jesse Hansen, “A Matlab Project in Optical Character Recognition (OCR)”