Answer Evaluation Using Machine Learning

Conference Paper · March 2018

CITATIONS
6

READS
8,763

3 authors, including:

Prince Sinha
University of Mumbai
3 PUBLICATIONS **9** CITATIONS
SEE PROFILE

Ayush Kaul
University of Mumbai
3 PUBLICATIONS **9** CITATIONS
SEE PROFILE

Answer Evaluation Using Machine Learning

Prince Sinha
Thakur College of engineering and technology, Mumbai

Sharad Bharadia
Thakur College of engineering and technology, Mumba

Ayush Kaul
Thakur College of engineering and technology, Mumbai

Dr. Sheetal Rathi
Associate professor, Thakur College of engineering and technology, Mumbai

***Abstract* — In this modern age, where the world moves towards automation so, there is a need for automation in answer evaluation system. Currently, the online answer evaluation is available for mcq based question, hence evaluation of the theory answer is hectic for the checker. Teacher manually checks the answer and allot the marks. The current system takes more manpower and time to evaluate the answer. In this journal an application based on the evaluation of answers using machine learning. The objective of the journal is to specially reduce the manpower and time consumption. Since in manual answer evaluation, the manpower and the time consumption is much more. Also, in the manual system, it may be possible that the marks given to two same answers are different. This application system provides an automatic evaluation of answer based on the keyword provided to the application in form of the input by the moderator which will provide equal distribution of marks and will reduce time and manpower.**

Keywords — OCR, Backpropagation algorithm, ReLU, ANN, CNN, RNN, CRNN

INTRODUCTION

Manual answer evaluation is a very tedious task. The manual checking is very time consuming process and also requires lots of manpower. Also, the paper checker is not able to give marks equally. So, our system will evaluate answer based on some keyword and also manpower will be saved. Only one has to scan the paper then, based on the keyword in the answer the system will provide the marks to the question according to the dataset present. Also, By this system, the evaluation error of the marks to the particular question will be reduced.

So, our system will evaluate answer based on some keyword and also manpower will be saved. Only one has to scan the paper then the system will split the answer using OCR[1], based on the keyword in the answer the system will provide the marks to the question according to the dataset present [2]. There is a need for such application which will provide an easy evaluation of answer and can provide eligible marks. Also, this application will help various colleges, university, coaching institute to evaluate the answer in less time and with less manpower.

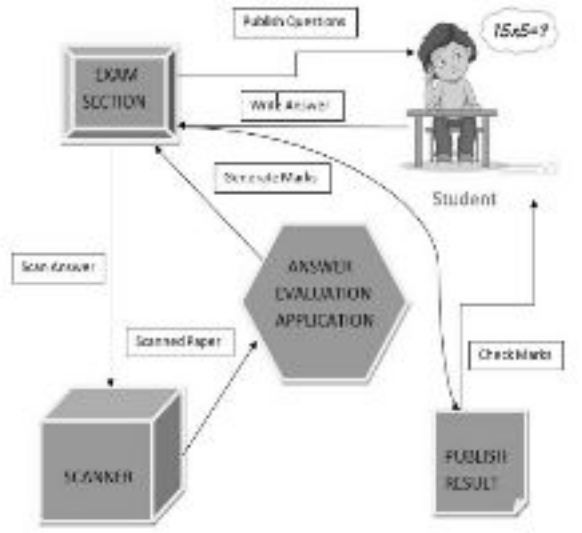

Figure 1. Overall architecture of the system

Checking answers requires high concentration for the large amount of time which often leads to mistakes. The automation of this task will increase the efficiency of answer evaluation on a large scale. After a brief discussion, it was understood that answer sheet is evaluated keeping in mind certain keywords that moderators search for the answer while evaluating an answer. Our proposed algorithm will require keywords as inputs. These keywords will be provided by the subject expert. Our proposed algorithm will match these keywords with detected words that are extracted from the answer sheet using supervised learning algorithm. Learning phase of the model will require handwritten dataset[3] for English language alphabets. These datasets are available online in various formats to be used to train the model. The machine learning model used in our proposed algorithm is

neural networks with multiple hidden layers. The model calculates the error using backpropagation[8] algorithm. The weights of the network are updated in the direction opposite to

the partial differentiation of error with respect to weighted input to the neuron in a particular layer.

The activation function used for the model is ReLU(Rectified Linear unit) which calculates as:

$$f(x) = max(0, x)$$

Here the variable x is an input to the function. Our proposed algorithm will also consider the length of the answer as a parameter for evaluation of the answer. The ideal answer length will be taken as an input from the teacher.

Related Work

Sheeba Praveen[2] says that, in recent years it has been seen that a number of governments, semi government examinations has gone online, for example [IBPS Common Written Examination (CWE)]. According to the professor, in recent years, we are able to make online test provision for objective type question. As, online system has much more advantages over the offline examination. Online exams saves time to evaluate the answers. As, there were no provision for offline paper evaluation, he designed an algorithm that automatically evaluate the single sentence descriptive answer

B vanni, M. shyni, and R. Deepalakshmi [4] OCR refers to translate handwritten text to format which is machine readable which can be used for searching, editing and indexing. This paper is using the artificial neural network to achieve high accuracy for optically recognize the character. The proposed approach is tested and implemented on character database consisting of English characters, digits and special characters.

Another approach by Baoguang Shi, Xiang Bai and Cong Yao[5] combines the capabilities of CNN(Convolutional Neural Networks) and RNN(Recurrent Neural networks).The convolutional layers in this case is constructed by considering only the convolutional and max-pooling layers from a standard CNN model (fully-connected layers are removed). As CNN can identify the individual characters in the answer. But the variable length of answers is a problem for CNN to identify as the length of inputs is fixed. To solve this problem RNN is used with LSTM(Long Short Term Memory) to save long sequence of characters in string. This model was termed as CRNN(Convolutional Recurrent Neural Network). In CRNN, deep features are represented into sequential representations in order to be invariant to the length variation of sequence-like objects.

Yusuf Perwej and Ashish Chaturvedi [6] proposed a neural network for handwritten english alphabet recognition. In this paper they proposed and developed a scheme for handwritten character recognition by representing each english alphabet in binary value that can be used as input to the system and the output produced by the system then can be used as input to the neural network system. They have explained two phases namely pre-processing phase and neural network based recognition phase. In processing only 25 bits are used because the alphabets are divided in 25 segments. They have achieved the accuracy of 82.5% but system found less accurate for the similar alphabets.

J.Pradeep , E.Srinivasan and S. Himavathi [7] worked on off-line handwritten alphabetical character recognition system which uses multilayer feed forward neural network. A new method other than horizontal and vertical method was introduced called diagonal based feature extraction. Fifty datasets were taken to train the network. Each dataset contained 26 alphabet written by different people. The system performed quite well with higher levels of accuracy in recognizing the alphabets.

Background

Traditional answer evaluation ie. manual checking of answers takes a lot of time and energy. However, till now only multiple choice questions can be evaluated with the use of computers. When it comes to theoretical evaluation of answers, there is the need for a teacher to check the answer sheet. Hence teacher has to put his effort in answer sheet evaluation rather than providing knowledge to the students. Also, the traditional answer evaluation takes more time than that of the machine.

There is a limitation when it comes to a human. He can work for the limited amount of time. After that, he must take rest. But that is not the case with the machine. A machine can work 24*7 nonstop to give the output. The machine also eliminates the human error and hence it can be reliable.

Hence, we came up with the solution for this problem so that, it will save the time and resources. We take the answers in image format and evaluate that answers with respect to keyword, length, unique character and some other parameter. We have come up with an algorithm which will alot mark to the student according to the algorithm. We have tried this algorithm to be more precise and accurate.

Proposed Methodology

This project is an application for automated answer evaluation using the matching keyword from dataset based on machine learning algorithm. Some applications are available but they are different than this and they use different methodology. Some available application only evaluates MCQs (multiple choice questions) not the subjective question[2].

For using this application only one has to scan the answer to that question then the system will split the answer keyword using OCR [1]. Based on keywords written in the answer and the keywords in the dataset, the application will provide marks in the range of 1 to 5.

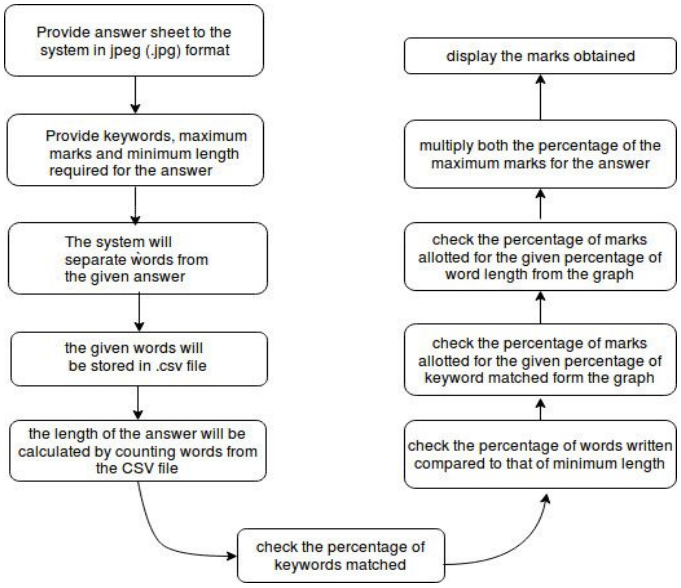

Figure 2. Flowchart depicting the algorithm

Steps to evaluate the answer

- Provide answer sheet to the system in jpeg (.jpg) format
- Provide keywords, maximum marks and minimum length required for the answer
- The system will separate words from the given answer
- the given words will be stored in .csv file
- the length of the answer will be calculated by counting words from the CSV file
- check the percentage of keywords matched
- check the percentage of words written compared to that of minimum length
- check the percentage of marks allotted for the given percentage of keyword matched form the graph
- check the percentage of marks allotted for the given percentage of word length from the graph
- multiply both the percentage of the maximum marks for the answer
- display the marks obtained

Function to calculate % marks for key words matched

$f(x) = 25 \quad for \;\; 0 =< x < 20$

$f(x) = (1.5 * x) - 5 \quad for \;\; 20 =< x < 40$

$f(x) = x + 15 \quad for \;\; 40 =< x < 80$

$f(x) = (0.25 * x) + 75 \quad for \;\; 80 =< x < 100$

Function to calculate % marks for key words matched

$y = x \quad for\ 0 >= (y, x) <= 100$

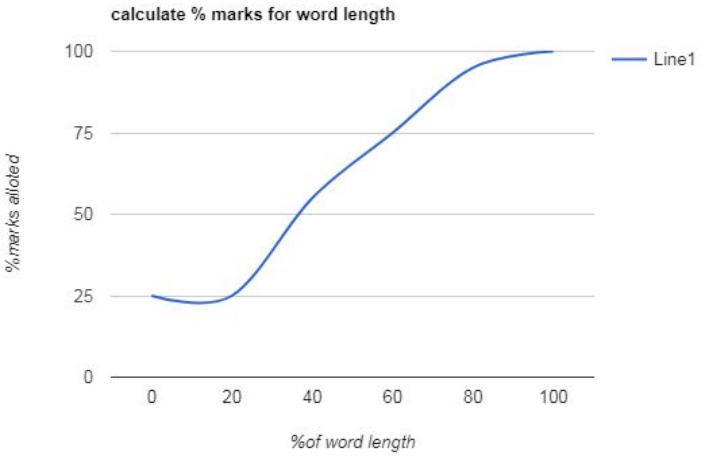

Figure 3. percentage of marks allotted based on answer length

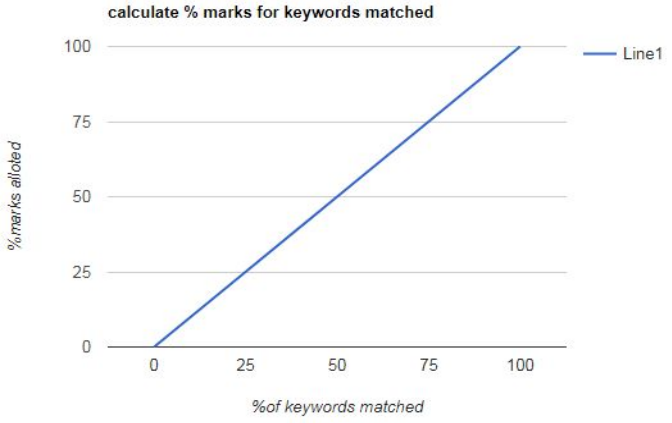

Figure 4. percentage of marks allotted based on number of keywords matched.

$$maximum\ marks \times f(n) \times \left[\frac{No.of\ keyword\ matched}{Total\ Keyword}\right]$$

Results and Discussion

The current manual evaluation takes about 60 seconds to evaluate an answer whereas the proposed system takes about 15 seconds to evaluate an answer. The proposed system is 300% more time efficient as compared to manual answer evaluation system. The proposed system is about 75-87.5% accurate with comparison to manual system. The proposed system completely eliminates the human effort and time to evaluate an answer.

Proposed system can evaluate 5760 answers in a day whereas a human working for 8 hours can evaluate 480 answers a day.

Hence, proposed system can evaluate 1100% more answers compared to that of manual evaluation system.

The amount of money spent on manual answer evaluation is much more if compared to our proposed system. The installation of our proposed system is one time investment with negligible maintenance cost.

Sample. No.	No. of words	No. of Keywords Matched	No of sentences written
1	179	12	10
2	137	12	7
3	36	7	1
4	200	12	7
5	200	12	13
6	173	10	6
7	200	12	7
8	29	4	2
9	131	12	4
10	167	11	4

Table 1. Evaluation Parameters

Table 1 illustrates that the parameters like number of words, no of keywords matched, no of sentences, the machine calculate the marks to be allotted to a particular answer. Where the weights of the parameters are taken as per the above figure 3 and figure 4.

Sample No.	Marks alloted in Manual Checking	Marks alloted b the System
1	5	5
2	4.5	4
3	2	1.5
4	4	4
5	4.5	5
6	4	4
7	5	5
8	1.5	1
9	2.5	3
10	3.5	3.5

Table 2. Comparison between Manual checking and by Machine

According to Table 2, the results are compared.The scores are calculated for 10 student. In this table, Second column evaluate the scores based on manual checking and the third column evaluate the score according to program. The difference between the manual evaluation and system evaluation are very close.

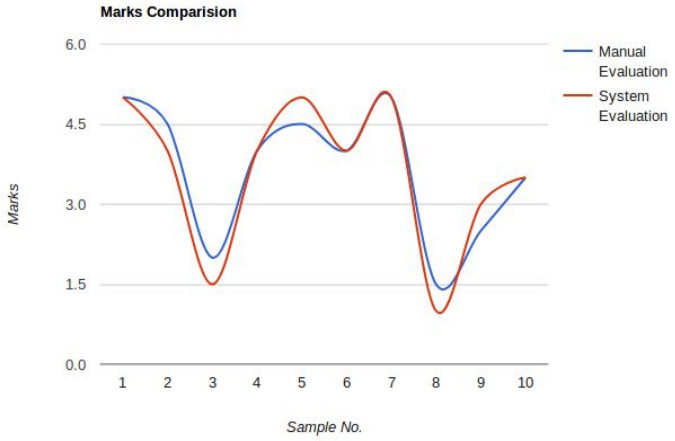

Figure 5. Line chart for comparison of normal evaluation and system evaluation

The graph in figure 5 shows that for samples ranging from sample no. 1 to 3 the marks given by our system is less in comparison to the marks given by the moderator. The samples ranging from 3 to 8 are almost coinciding with each other which shows high accuracy. For samples ranging from 8 to 10 the lines do not coincide, So there is a notable difference between manual evaluation and system evaluation.

$$\mathrm{MAE} = \frac{1}{n}\sum_{j=1}^{n} |y_j - \hat{y}_j|$$

Calculating Mean Absolute Error for above is sample: 0.25.
Mean absolute error is calculated between the marks calculated by our proposed algorithm and the evaluated samples that were evaluated by the moderator. It is calculated by taking difference between marks calculated by our proposed algorithm and the marks calculated by moderator evalu ated answer samples. There is a possibility that difference can also be negative so absolute values are only considered while considering the sum of all differences as shown in above formulae for MAE. The absolute error for the system is 25% compared to manual evaluation. Hence, it can be concluded that accuracy of 75% is achieved in comparison to manual evaluation calculated for the given sample of 10 answers.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2.$$

Calculating Mean Squared Error for above sample: 0.125.
The squared error doesn't consider the absolute values. Rather than that is takes square of difference between individual samples. Hence the difference terms are positive and error calculated for given 10 samples is about 12.5%. Therefore, the accuracy comes out to be about 87.5% as compared to the manual marks obtained from evaluation by moderator.

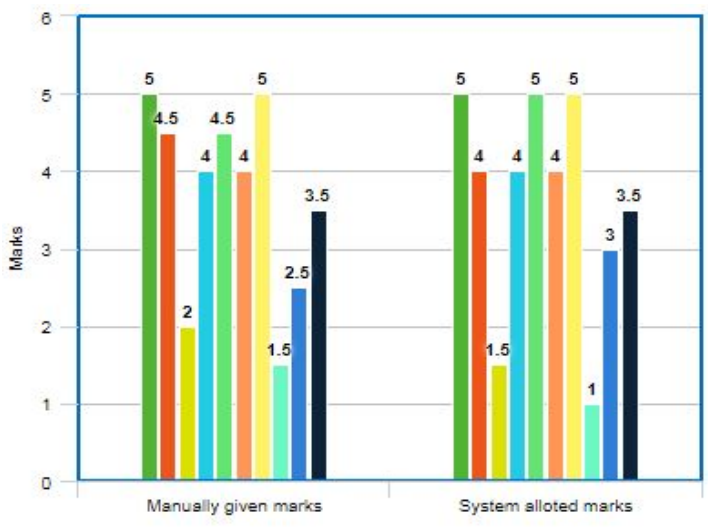

Figure 6. Bar graph depicting comparison of manual and system allotted marks

The graph in the figure is a bar graph showing the comparison of manual and system allotted marks. In this graph there are 2 groups with 10 bars in each group. The bar shows the answer number which is from 1 to 10. In the graph. the first group is for manually given marks by the teacher which is in left side while the second group is for marks alloted by our system.

Conclusion

In this paper, we have developed an algorithm which will evaluate theoretical answers and give marks according to the keyword matching which will reduce manual work and saves time with faster result evaluation. A person should collect the answer copy from the student and scan it. The machine will take the image as input and will evaluate the answer based on the length of the answer and important keywords covered which are specified by the teacher with each answer which is to be evaluated. The algorithm assigns marks on basis of :

- The Number of keywords matched.
- Length of the answer.

Future Scope

In future, We can extract, handwritten text from the image instead of printed text from image. This will be more realistic and more useful.We can use recursive neural network (RNN) to train our model with different handwriting. This will make our model more accurate.

The model can be trained for different languages across India. In this, we can collect dataset of different handwritten languages. Hence a answer with language other than English can be evaluated. System will also evaluate the overwritten alphabets and other words with absolute accuracy.

The software can be trained in such a way so that, it can check the complete paper instead of a single answer. Hence the software will evaluate the answers according to the answer no provided in the answer sheet. The model can also be trained to evaluate diagrams and hence give the marks accordingly.

Acknowledge

This research was supported by our faculty Ms. Harshala Yadav and Ms. Venna Kulkarni (Thakur College of Engineering and Technology). We thank our colleague Mr. Monark Unadkat (Thakur College of Engineering and Technology) who helped and provided insight that greatly assisted the research.

References

[1] S. Impedovo, L. Ottaviano and S.Occhinegro, "*Optical character recognition*", International Journal Pattern Recognition and Artificial Intelligence, Vol. 5(1-2), pp. 1- 24, 1991

[2] Sheeba Praveen, *"An Approach to Evaluate Subjective Questions for Online Examination System"*, Assistant Professor, Dept. CSE,Integral University, Lucknow ,U.P, India.

[3] J.Pradeep , E.Srinivasan and S.Himavathi, "*Diagonal Based Feature Extraction For Handwritten Alphabets Recognition System Using Neural Networks*", International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 1, Feb 2011.

[4] B vanni, M. shyni, and R. Deepalakshmi, *"High accuracy optical character recognition algorithms using learning array of ANN"* in Proc. 2014 IEEE International Conference on Circuit, Power and Computing Technologies (ICCPCT), 2014 International Conference.

[5] Baoguang Shi, Xiang Bai and Cong Yao, "*An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition*", arXiv:1507.05717v1 [cs.CV] 21 Jul 2015.

[6] Yusuf Perwej, Ashish Chaturvedi, "*Neural Networks for Handwritten English Alphabet Recognition*", International Journal of Computer Applications (0975 – 8887) Volume 20– No.7, April 2011.

[7] J.Pradeep , E.Srinivasan and S.Himavathi, "*Neural network based handwritten character recognition system without feature extraction*", International Conference on Computer, Communication and Electrical Technology – ICCCET 2011, 18th & 19th March, 2011.

[8] W.A.M. Ahmed, E.S.M. Saad, E.S.A. Aziz, "*Modified back propagation algorithm for learning artificial neural*

networks", Radio Science Conference, 2001. NRSC 2001. Proceedings of the Eighteenth National, 27-29 March 2001

[9] K Arlitsch, J Herbert *"Microfilm, paper, and OCR: Issues in newspaper digitization. the Utah digital newspapers program*" Microform & Imaging Review, 2004 - degruyter.com

[10] Shai Shalev-Shwartz and Shai Ben-David "*Understanding Machine Learning: From Theory to Algorithms*", ©2014

[11] George Nagy, Stephen V. Rice, and Thomas A. Nartker, "*Optical Character Recognition: An Illustrated Guide to the Frontier (The Springer International Series in Engineering and Computer Science)*"

[12] Batuhan Balci, Dan Saadati, Dan Shiferaw, *"Handwritten Text Recognition using Deep Learning*", CS231n: Convolutional Neural Networks for Visual Recognition Spring 2017 project report.

[13] Convolutional Recurrent Neural Network, https://github.com/bgshih/crnn.

[14] Tensorflow implementation of handwritten sequence of small letter recognition, https://github.com/johnsmithm/handwritten-sequence-tensorflow.

[15]Handwritten names Recognition, https://github.com/ikergarcia1996/Handwritten-Names-Recognition.

[16] Ivor Uhliaqp Ñrik, *"Handwritten Character Recognition Using Machine Learning Methods*", Comenius University In Bratislava Faculty Of Mathematics, Physics And Informatics Thesis.

[17] S. Mori, C.Y. Suen and K. Kamamoto, "*Historical review of OCR research and development*," Proc. of IEEE, vol. 80, pp. 1029-1058, July 1992.