ISSN 2395-1621

Subjective Question Answering System with Automatic Result Generation using Natural Language Processing

Dr. Meenakshi A. Thalor[1], Anushi Agarwal[2], Leona Sam[2], Seona Sam[2], Vishaka Arjun Mandge[2]

[1] Associate Professor

[1]itdept_hod@aissmsioit.org

Information Technology, AISSMS Institute of Information Technology, Kennedy Road, Near RTO, Pune, Maharashtra, India

[2] Student,

Information Technology, AISSMS Institute of Information Technology, Kennedy Road, Near RTO, Pune, Maharashtra, India

ABSTRACT

Knowledge Engineering helps in understanding how a human would perform tasks in a certain domain. Knowledge Engineering is the process of solving and analyzing complex problems that require human expertise. The project aims at automating the process of paper checking and generation of results through Natural Language Processing (NLP) in order to eliminate the traditional method of giving purely subjective exams on paper. It uses several answer matching techniques like keyword matching, semantic and lexical analysis in order to generate percentage matched to deliver scorecard of the examinee to the examiner. It supports several file formats that can be used as the referral answers for matching. The worldwide use of this project will help reduce the use of papers and hence benefit society.

Index Terms: Knowledge Engineering, Natural Language Processing, Deep Learning, Question Answering System, Automatic Evaluation, Sequence Matching

ARTICLE INFO

Article History

Received: 8th March 2020

Received in revised form :

8th March 2020

Accepted: 10th March 2020

Published online :

11th March 2020

I. INTRODUCTION

In the new age of technology, the use of artificial intelligence and machine learning has gained momentum. [1] These technologies have proved to be beneficial to society in various fields, such as communication, transport, health care, commerce, etc. In spite of all these developments, the education system is still practicing the traditional method of conducting subjective exams on pen and paper.

To digitize the current practice of giving theory exams, we have proposed a subjective question answering model that will load the question paper along with its model answers into the database and use the extracted keywords as the parameter for checking the student answers. Each keyword will have its dataset consisting of synonyms. The synonyms will help in detecting and discarding the sentences that mean the same, which will eliminate redundancy. These answers will be compared with a model answer which is already loaded into the database. The system after comparing will generate a percentage depending on its closeness to the model answer.

In this proposed model, we will be using a technique known as knowledge engineering. Some of the existing models such as Paper Rater check plagiarism of the given text and even improves it grammatically using Artificial Intelligence and Machine Learning algorithms. This model only improves the quality of the given text by removing grammatical errors and checking its vocabulary. Another example is the E-rater Scoring Engine. This engine is used to grade essays written by students. The essays are graded as a whole without subdividing it. The E-rater Engine makes use of Natural Language Processing to evaluate and obtain a final score of the essay.

The existing online question answering system [2] consists of only objective type questions. The existing online question answering system, only performs keyword matching for answer evaluation without considering its synonyms. Hence this system may not be flexible for all types of answers. In this paper, we plan on implementing a Question Answer model that is able to handle subjective exams. The system will accept answers, compare it with the model answer and provide a percentage for the correctness of the answer. The model aims at considering various factors such as, repetition of keywords, repetition of sentences that mean the same, etc. A detailed explanation and flow of all the techniques and factors considered while creating this model is explained in the proposed methodology.

II. LITERATURE SURVEY

The research on NLP and ML [1], carries out extraction and mapping of data by identifying, analysing and interpreting the suitable information. According to the analysed literature, NLP Unsupervised learning is the most trending technique used but does not ensure large text collection. In the system proposed by K.T. Kodithuwakku et al [2] the system uses NLP tools and methodologies to automatically evaluate text answers. It compares semantic similarities between the model answer and the provided answer. A percentage is generated based on this similarity However, there are restrictions due to language complexity and the system can only work with selected sentence patterns and grammar conditions. [3] By using NLP and techniques such as semantic patterns, ontology and interactive conversation system, the proposed model can analyse the completeness and meaning of natural language statements. There are two types of errors; missing alphabets which change the meaning of a word and wrong position of words in the sentence.

[4]The main purpose is to experimentally evaluate association rule for mining using XML database. The XML documents are classified; it includes all the information without eliminating any information as in the existing system. Based on a study carried out by Wenpeng Yin et al [5] CNNs are Hierarchical architectures that extracts position-invariant features.RNN are sequential architectures consisting of two types- LSTM and GRU. It is used for modelling units in sequence. GRU/LSTM surpasses CNN in textual entailment while CNN dominates in answer selection. Optimization of hidden size and back size is crucial to obtain good performance of both CNN and RNN. To carry out text classification, a deep learning model was proposed [6]. The deep learning model carries out text classification using three steps- text pre-processing, text feature extraction and classification model construction. The deep learning model divides the text into long text and short text which the RNN model cannot parallelize well, CNN is best for long text. The “Matching-Aggregation” framework [7], words in the two sequences are compared to check if they are matching. Based on the number of matches a final result is obtained. The error analysis of Knowledge based question answering system suggests that complex questions are challenging and failing case of ambiguous questions.

The medical question answering model [8] consists of multiple CNN layers from which information is retrieved and the features that are obtained at the end of this model are combined to produce a solution. It makes use of multilevel CNN model. It can retrieve semantic properties and shows better performance as compared to single level neural networks. Conversion of a query into SQL languages [9] can be done by using four main steps tokenization, lexical analysis, syntactic analysis, semantic analysis. The system converts the query into structured query language which can accept text as well as speech. QAVAL [10] based on answer validation system to choose most relevant answer for the given question using learning methods. The model predicts the most relevant answer or valid answer using QAVAL system. A document summarization method [11] based on three methods-query-biased, CQA-answer-biased and expanded-query-biased and learning-to-rank-based. The quality of CQA content is accurate for optimization-based summaries and not accurate for learning-to-rank-based summaries. The Answer Extraction method [12] interprets the questions that help to generate a query. It uses NLP technology to analyse the retrieved document. Construct an answer ranking model. The proposed method improves the accuracy and recall rate. Some questions are difficult to interpret due to its complexity.

Based on the study carried out by Min-Yuh Day and Cheng-Chia Tsai [13] comprehensive analysis of imbalance and balance datasets is done using machine learning. Imbalance has more accuracy then balanced dataset but the reason is not mentioned. The Automatic Answer Validation System [14] combines the question and answer into Hypothesis (H) and the Supporting text (T) to check the entailment relation as either “VALIDATED” or “REJECTED”. The system uses Lexical information, Dependency, Chunking and Named Entities but does not use semantic features for answer validation. The QA system [15], extracts answers for a question from primary knowledge source. The system carries out tasks like Named Entity Tagging, Question Classification, Information Retrieval and Answer Extraction. However the system is not context sensitive and supports only English language. It can only extract current information.

The enhanced lexical and semantic model [16] proposed by Wen-Tau Yih et al, makes use of enhanced lexical semantic models, carries out answer selection using semantic matching with latent word alignment structure. The models main sources of errors are inaccurate entity relations, lack of robust question analysis and need of high semantic representation. Paraphrase based approach [17] for word mismatch in question retrieval. For this it makes use of two approaches, key concept identification approach and pivot language translation based approach. This model outperforms most of the existing models by eliminating problems like word verboseness and word mismatch in question retrieval. It does not support multiple linguistic resources. The metrics used for comparing different models is the MAP (Mean Average Precision) and MRR (Mean Reciprocal Rank) [18]. It lacks the use of Tree Kernel based semantic question matching and syntactic features.

III. PROPOSED METHODOLOGY

The proposed question answering model, makes use of Natural Language Processing (NLP) based Knowledge Extraction to analyse and interpret each statement of the answer. NLP, is making a computer program to understand human language. Knowledge Extraction is creating knowledge from structured text such as database, CSV [3] or XML [4] files, and unstructured text such as documents and images with the help of some already existing knowledge or some pre-determined data. In this model, we will be carrying out knowledge extraction only on structured data. Collaborative learning generally means two or more individuals learn something together. Our model makes use of Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) [5, 6] in a collaborative manner as proposed byTaihua Shao et al [19]. CNN is used for image detection and text classification. Although CNN's are most suitable for analysing images, some recent developments have shown the successful implementation of text classification using CNN. RNNs are typically used for NLP and speech recognition. While reading the text, RNNs can predict and interpret the future content of data using the already read text. RNNs can be further classified into two types Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU). In this model, we will be focusing on Bidirectional LSTM (BiLSTM) which is nothing but LSTM that can analyse the text in two ways, from past to future and from future to past a statement. This, in turn, helps in analysing text better and faster as compared to unidirectional LSTM which only analyses text from past to future. Our proposed model will carry out NLP by collaboratively using CNN and BiLSTM to analyse text.

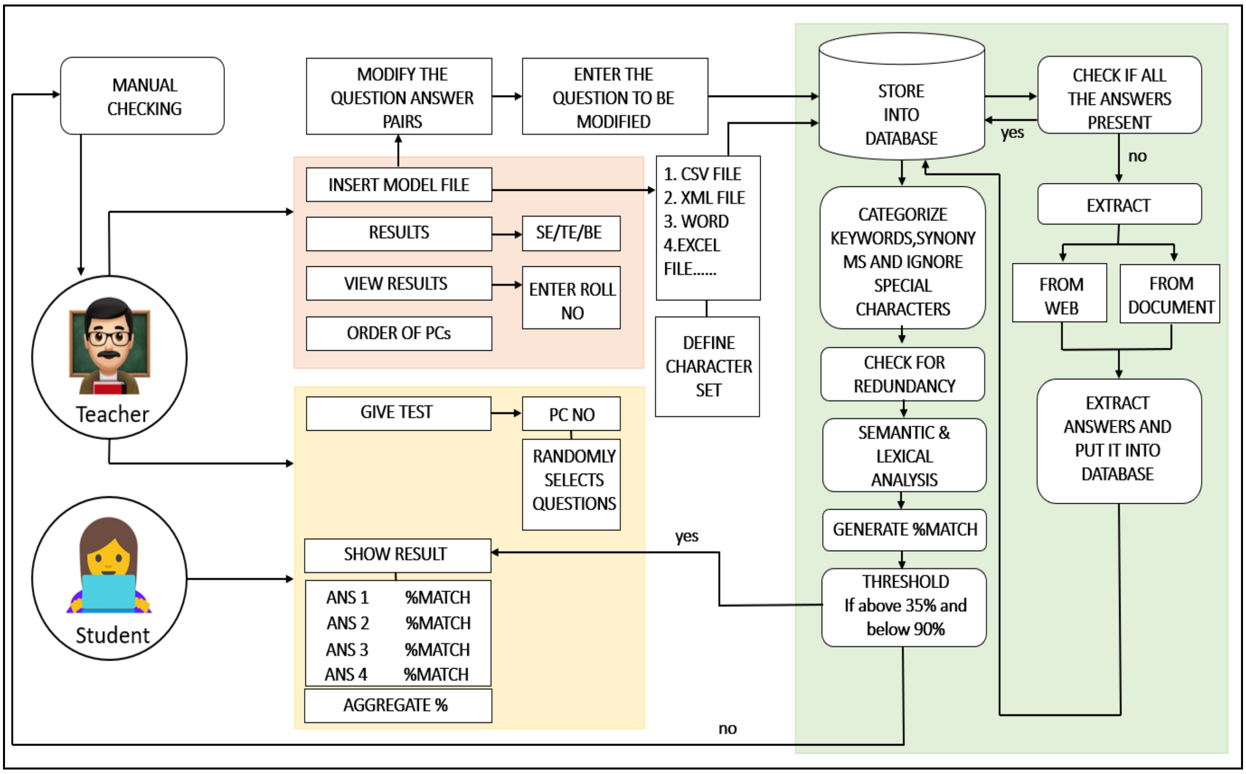

Fig. 1 Architecture of proposed model

The question answering system will require the document containing the questions and their respective model answers to be uploaded onto the database. This document must be a CSV file [3] or XML file [4]. The CSV file is then converted into the JSON format so that it can be fed into the MongoDB database. The keywords from the answer will be extracted using tokenization [9, 310]. All the words apart from the stop words will be used as keywords. It will also ignore all the punctuation marks and special characters. All these keywords will be stored as a set for each answer. There will be another dataset containing the synonyms of these keywords as show in the example below. Using this we are considering all types of words that do not match the keyword but also mean the same. The question paper for each student will contain questions generated at random from the uploaded questionnaire. This will help avoid any malpractices. Once the students have submitted their tests, it will get stored in the database for evaluation. The answers will be evaluated with the help of the model answer for that question. Each answer will undergo lexical and semantic analysis as well as keyword and synonym matching [18, 19]. The answers will be scored based on the number of keyword matches or their respective synonyms [16]. The system will also discard the repetitive sentences and even the sentences that have the same meaning. A percentage of correctness will be assigned to each question attempted by the student [2]. The final score will contain the aggregate percentage of the answers.

Consider an example,

TABLE I

Sr. No.	Question	Answer	Keywords	Synonyms
1	What is AI?	Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems.	Artificial intelligence, simulation, human, intelligence, processes, machines, computer systems	(AI), (imitation, replica), (man, individual), (intellect), (operates, acts), (tools, device, mechanism), (system)

The database will consist of a table consisting of the Sr. No., Question, Answer, Keywords and Synonyms for each question. The records of all the students will be made available to the teachers for reference. Suppose the percentage of a particular student is not within the threshold, which is between 35% and 90%, must be evaluated again by the respective teacher. We may come across a situation where a student must have entered all keywords which may lead to a score of above 90%. Such special cases may not always be detected by our system. The system may also come across a situation where it did not evaluate the answer properly and provided a low percentage which led to a student failing in that subject. In such scenarios, a revaluation must be done by the teacher. The GUI will contain a student login and a faculty login as shown in Fig. 1. Before beginning the test, the student must enter their roll number and the computer number on which they are attempting this test while logging in. This will help us tally the answers of the students who have appeared for the test next to each other. The faculty login will allow the teachers to create and edit the questionnaire for the test. They will also have access to all the checked tests for further review or revaluation as and when necessary. This question answering system will help reduce the load of checking test papers on teachers. It will also generate faster results than manual checking. Since answers are checked automatically using keywords, the evaluation will be more efficient. This system will also help the environment as there will be reduced usage of paper in educational institutes.

IV. CONCLUSION

The proposed system will help automatic process of paper checking and generation of results through natural language process (NLP) in order to eliminate the traditional method of giving purely subjective exams on paper. It uses several answer matching techniques like keyword matching semantic and lexical analysis in order to generate percentage match to deliver score card of the examinee to the examiner. It supports several files formats which can be used as the referral answer for matching. This will lead to high reduction in paper consumption in colleges, schools and any other pen-paper examination.

In future work, the system could also be used for speech. It can be used in verbal examinations such as extempore, debate, etc. It could also be used in professional purpose such as job interviews where several interviews could be recorded and compared with the desired result with the help of our software. The system can also be modified to extract answers from the web [17]for questions for which the answers are not provided.

REFERENCES

[1] DiellzaNagavciMati, JauminAjdari, BujarRaufi, MentorHamiti and BesnikSelimi, A Systematic Mapping Study of Language Features Identification from Large Text Collection.In: 2019 8th Mediterranean Conference on EMBEDDED COMPUTING (MECO), 10-14 June 2019, Budva, Montenegro

[2] K.T. Kodithuwakku., W.S.J.M.C.D. Senevirathne, R.A.D.A. Randeniya, M.M.D.D.A. Wijewardane and M. P. A. W. Gamage, Self-Learning System with Automatic Feedback for Text Answers. In: 2017 11th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)

[3] ChalermpolTapsai, Information Processing and Retrieval from CSV File by Natural Language..In: 2018 IEEE 3rd International Conference on Communication and Information Systems

[4] S.Thangarasu and D.Sasikala, Extracting Knowledge from XML Document Using Tree-based Association Rules.In :2014 International Conference on Intelligent Computing Applications

[5] Wenpeng Yin, Katharina Kann, Mo Yu and HinrichSchutze Comparative Study of CNN and RNN for Natural Language Processing.In:2017

[6] JingjingCai, Jianping Li, Wei Li, and JiWang,Deeplearning model used in text classification.In:2018

[7] YunshiLan, Shuohang Wang, and Jing Jiang, Knowledge Base Question Answering with a Matching-Aggregation Model and Question-Specific Contextual Relations.In:2019 IEEE/ACM transactions on audio, speech, and language processing, vol. X, no. X, March2019

[8] Dong Ye, Sheng Zhang, Hui Wang, Jiajun Cheng, Xin Zhang, Zhaoyun Ding, Pei Li,Multi-Level Composite Neural Networks for Medical Question Answer, In: 2018 IEEE Third International Conference on Data Science in Cyberspace

[9] Abhilasha Kate ,SatishKamble, AishwaryaBodkhe, Mrunal Joshi, Conversion of Natural Language Query to SQL Query.In:2018 Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology (ICECA 2018)

[10] A. Grappy, B. Grau, M-H.Falco, A-L.Ligozat, I. Robba, A. Vilnat, Selecting answers to questions from Web documents by a robust validation process.In: 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology

[11] EviYulianti, Ruey-Cheng Chen, Falk Scholer, Bruce Croft and Mark Sanderson ,Document Summarization for Answering Non-Factoid Queries.In:2017 IEEE transactions on knowledge and data engineering, manuscript ID

[12] Huanyun ZONG Zhengtao YU Jianyi GUO Yantuan XIAN ,JianLIAn Answer Extraction

Method Based on Discourse Structure and Rank Learning.In 2011 School 0/ Information Engineering and Automation Kunming University o/Science and Technology KunmingChina

[13] Min-Yuh Day and Cheng-Chia Tsai, A Study on Machine Learning for Imbalanced Datasets with Answer Validation of Question Answering .In:2016 IEEE 17th International Conference on Information Reuse and Integration

[14] ParthaPakray, Santanu Pal, Sivajibandyopadhyay, Automatic Answer Validation System on English Language .In:20IO 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE)

[15] FaheemAbbas,MuhammadUmairRashid,Muhammad Kamran Malik,RizwanZafar, WikiQA A Question Answering System on Wikipedia using Freebase, DBpedia and Infobox.In:2016 In sixth International conference on innovating computing technology.

[16] Wen-tau Yih ,Ming-Wei Chang, Christopher Meek, AndrzejPastusiak Question Answering Using Enhanced Lexical Semantic Models.In: 2013 Association for Computational Linguistics

[17] Wei-Nan Zhang, Zhao-Yan Ming, Yu Zhang, Ting Liu and Tat-Seng Chua, Capturing the Semantics of Key

[18] Phrases Using Multiple Languages for Question Retrieval IN:2015transactions on knowledge and data engineering,vol. X, no. X, XX XXXX 1

[19] BarunPatra ,A survey of community answering.In:2017

[20] TaihuaShao ,XiaoyanKui , Pengfei Zhang , and Honghui Chen, Collaborative Learning for answer selection in Question Answering, In: IEEE 28 December 2018

