

Design and Development of an automated digital assessment system for virtual examination

Dr. Raju Pal

Professor, Department of CSE

Jaypee Institute Of Information technology

Lalit Garg, Neha Agarwal, Abhinav Verma

Students, Department of CSE

Jaypee Institute Of Information Technology, Noida

Abstract—A question-answer is an efficient way of information retrieval. The objective of the research is to reduce time consumption and minimizing the human intervention in manually generating objective and subjective questions and automatically generating the responses by matching the human answer with the correct answer. Many researchers have proposed strategies for automatic question-answer generation. After analyzing various reports, journals, and optimization techniques given in the research papers, we are able to conclude a final optimized model that can result in the required examination assessment system. By this system, evaluation error will be reduced, minimizing the human intervention, and resulting in a fair platform for both students and teachers.

I. INTRODUCTION

The project's title is "Design and development of digital assessment systems for virtual examination" in which we are trying to produce a solution for the modern problem, as the world is moving towards the automation and digital age, so there is a need for automation in the Indian examination system as well. The current pandemic has raised many questions, has shown how badly Indian educational institutions' are lacking in creating a fair environment for students and teachers as well. Major problem is to get a fair platform to automatically generate questions from a given syllabus, and checking of answers requires a high concentration, time, and energy which has increased the workload of teachers immensely and is prone to mistakes as well. Hence there is a need for an automatic system that can generate questions, check the answers, and generate performance. Also, managing a huge database of question-answer systems and different unfair means like paper-leakage, and generating different sets of questions is a major concern. In the manual system, it may be possible that different marks are given for the same answer. Solving all these problems, this system can lead to more efficient, fair, and, smooth conduction of examination.

The report describes the research for information on the recent work and solutions proposed to overcome the challenges in the automatic generation of question-answer for the development of an efficient digital assessment system for virtual examination. A question-answer is an efficient way of information retrieval. The objective of the research is to reduce time consumption and manpower in manually generating objective and subjective questions and automatically generating the responses by matching the human answer with the correct answer. Many researchers have proposed strategies for an automatic question-answer generation. After analyzing all the findings and optimization

techniques in the research papers, we are able to conclude a final optimized examination assessment system. By this system, evaluation error in the marks will be reduced. We have started from the research on basic NLP pipeline tools using tokenization, lemmatization, stemming, structure tree parser, part of speech tagging and then found significant work on the improvement of part of speech tagging using optimization techniques like hidden Markov model (Viterbi Algorithm), Rule-based POS tagging, transformation-based tagging, removing lexical and syntactical ambiguities in the texts while retrieving information using text summarization techniques like TF-IDF algorithm for paraphrasing, finding an optimized set of answers synonyms, UML diagram to remove an attachment and homonymy syntactic ambiguity. We have also focused on the optimization of evaluating subjective answers using various methodologies.

II. LITERATURE REVIEW

1) Algorithm for generating questions from the text: [8]

Introduction

Today there is a need for a system of automatic evaluation where after reading certain texts, students find a need for a platform where one can analyse his/her understanding and reading skills. Manually generating question-answer is a very time consuming process for teachers. Thus this platform makes the work of the educational system much easier and flexible.

Literature Review

A framework has been proposed to evaluate the performance of students. This is done to check whether a given text has been read carefully by the student. The methodology describes the fill in the blanks, true-false generation questions using multiple tools of nlp preprocessing. It is divided in two phases, first is the training phase where the given text syllabus is trained and preprocessed to generate questions and another is the evaluation phase where student-generated answers are matched with the already stored answer.

The first training phase follows these steps:

- Load the text.
- Detect potential sentences which can generate questions.
- Mark important terms which can be marked as blank.

- Using grammar and parsing, mark your grammar rule and filter the sentences according to the given grammar.
- Store the questions and respective answers for evaluation.

Challenges faced

For Fill in the blanks and True-false questions, what type of question- What/who/whom -is to be substituted is not an issue but for subjective and wh-questions, this is an issue. For example, what is the manager of the company is not a semantically correct sentence. There should be slightly variated text i.e Paraphrasing in order to avoid straight forward question answers.

Significance of work

We are working on the problems for lexical and syntactical challenges to ensure the semantic correctness of the questions generated. This research paper has mentioned the solutions: paraphrasing, generation of distractors which we have implemented in our project. We have implemented a tf-idf algorithm to overcome the paraphrasing challenge here.

2) Generating Multiple Choice test from a Medical Text: A Pilot Study: [15]

Introduction

Multiple Choice Question is the most common way of educational assessment but it is a very time-consuming task for teachers. In order to save time and labor, Mitkov proposed a system in 2006 which extracts the important key term from the syllabus and automatically generates multiple choice questions for assessment.

Literature Review

Mitkov has proposed a four step methodology-

- Sentence Parsing
- Key-term identification
- Source Clause Selection
- Transformation to stem

After sentence-parsing using grammar, Mitkov highlighted the identification of key-term for generating questions. Multiple Choice test items should have key-term as its anchor rather than other information.

In source clause selection, following structures should not be considered while picking potential sentences.

- Subordinate clause- Although chest pain is a problem.
- Negated clause- Rohan should not cook there.
- Coordinated NP- Excessive tension causes heart diseases and blood pressure problems.
- Initial Pronoun- It associates with hypertension instead.

Significance of work

We have implemented key-term identification and source clause selection in our methodology. This has improved our quality of the questions to a great extent and our work has addressed the shortcoming listed in the Research paper that students who simply memorize the chapter without actually

understanding it if it's directly framed without paraphrasing the text.

3) Single Document Automatic Text Summarization Using Term Frequency-Inverse Document Frequency (TF-IDF): [7]

Introduction

Today students focus largely on internet text for the preparation of major exams. This emerging need has intensified the text summarization platform using Natural language processing techniques. There are many algorithms that can be used to summarize the text. One of them is Term frequency-Inverse document frequency algorithm. Text summarization results in compressed text that helps the user to extract certain information.

Literature Review

Text summarization removes less important works and reduces to its much simpler version and thus helps the reader to extract the important information out of text. This research aimed to implement TF-IDF algorithm and manages to compare it with other text-summarization algorithms. The methodology of the method used in the approach produces 67 percent of accuracy with three data samples which are higher compared to other text-summarization algorithms. The algorithm uses existing libraries NLTK and TextBlob. Tf-IDF is a numerical calculated value for each word which measures the importance of each word with respect to the given text. Summary generation is the descending order of the words with one with higher TF-IDF value listed before. TF IDF value indicates the following importance with reference to the text.

TF=total count of words in a document/total number of words in a document

IDF= $\log(\frac{\text{All document Number}}{\text{Document frequency}})$

Preprocessing of the algorithm comprises part of speech tagging, word stemming and removing stop words like a, an, the etc. Then Feature extraction is done by obtaining the sentences from a document and allotting a value between zero and one. Last, summary generation is the sentences that are put into the summary in the order of the position in the original document.

Load text document → Text preprocessing → Calculate TF-IDF value → Calculate each sentence score → Summary generation

Significance of work

We have implemented the TF-IDF algorithm to ensure paraphrasing. Students who memorize the text for evaluation without actual understanding should not be given directly framed question-answers. Thus Term frequency-Inverse Document frequency algorithm generates slightly variated text which helps in building a complex set of question-answers and thus improves the process of automated evaluation. Term frequency-Inverse frequency algorithm is used to remove stopwords and unwanted text.

4) A Study on Different Part of Speech(POS) Tagging Approaches in Assamese Language: [23]

Introduction

Part of speech tagging plays a very crucial role in Natural language processing. It has an international and worldwide approach. It is called POS tagging. POS tagging is a piece of software that pairs up each word to its speech grammatical category-pronoun, noun, adverb, verb, adjective and thus it is a crucial part in natural language processing algorithms. For example:

Word:Ram, POS tag:Noun

Word:Melodious, POS tag:Adjective

Word:come, POS tag:Verb

Literature Review

Architecture of POS tagger:

- Tokenization

The sentence in the text is divided into words(tokens). The tokens may be words,punctuation marks.

- Ambiguity look-up

Lexicon and guessor is used for unknown words to predict the part of speech. A lexical analyzer is a program which breaks a text into lexemes(tokens).

- Ambiguity Resolution

If an expression(word/phrase/sentence) has more than one POS tag associated with it is, then it is referred to as ambiguous. For example:

She saw a bear.

Your efforts will bear fruit.

The word bear in the above sentences has completely different senses, but more importantly one is a noun and other is a verb.

POS Tagging Techniques:

- Rule-based POS Tagging

It is the oldest method which lists hand written rules for tagging to identify the correct tag in case of ambiguity.

- Markov Model

A finite state machine in which each state has two probability distributions. The probability of emitting a symbol and probability of moving to a particular state. The aim of Markov Model is to find optimal sequence of tags $T=t_1, t_2, t_3, \dots$ for a proper word sequence $W=w_1, w_2, w_3, \dots$. Here probability of future state depends on the previous state.

- Hidden Markov Model

It is a dynamic approach of Markov Model with improvement in Emission and transition probability table. It is also known as Viterbi Algorithm. It is called Hidden Markov Models because state transitions are not observable.

Significance of work Syntactic parsing is an important part which is required for Part of Speech Tagger. For the international applications in NLP, POS tagger plays a very crucial rule and is the most basic important step

performed. Improvised POS tagging approaches Rule based tagging, Markov Model, Hidden Markov Model has successfully resolved ambiguity in POS tagging to a much greater extent.

5) Hidden Markov Models : [5]

Introduction

Hidden Markov Model has wide application in POS tagging. IT is a powerful statistical tool in the context of natural language processing tools. HMM have also been found useful in applications like signal processing and speech processing. Baum and his colleagues developed the theory of the Hidden Markov Model in the 1960s.

Literature Review

The model has three states-Bull, Bear and Even and three index observations-up,down and unchanged. The model is a finite state transition with probabilistic transition between states.

The formal definition of a HMM is as follows:

(A, B,)

S is our state alphabet set, and V is the observation alphabet set:

$S = (s_1, s_2, \dots, s_N)$

$V = (v_1, v_2, \dots, v_M)$

We define Q to be a fixed state sequence of length T, and corresponding observations O:

$Q = q_1, q_2, \dots, q_T$

$O = o_1, o_2, \dots, o_T$

A is a transition array, storing the probability of state j following state i . Note the state transition probabilities are independent of time:

$A = [a_{ij}], a_{ij} = P(q_t = s_j | q_{t-1} = s_i)$

B is the observation array, storing the probability of observation k being produced from the state j, independent of t:

$B = [b_{ik}], b_{ik} = P(x_t = v_k | q_t = s_i)$.

is the initial probability array:

$\pi = [\pi_i], \pi_i = P(q_1 = s_i)$

Two assumptions are made by the model. The first, called the Markov assumption, states that the current state is dependent only on the previous state, this represents the memory of the model:

$P(q_t | q_{t-1}) = P(q_t | q_{t-1})$

The independence assumption states that the output observation at time t is dependent only on the current state, it is independent of previous observations and states:

$P(o_t | o_{t-1}, q_{t-1}) = P(o_t | q_t)$

Significance of work

Given a Hidden Markov Model and sequence of observations, Emission probability, transition probability, we would be able to compute the probability of each hidden sequence and thus Ambiguity will be resolved in POS tagging. In case of sequence having more than one certain POS tag sequence, the one with higher probability will be

given higher preference.

6) Adaptation, Comparison, and Improvement of Metaheuristic Algorithms to the Part-of-Speech Tagging Problem.: [26]

Introduction

Metaheuristic algorithms are very common in today's emerging technology. One of the major applications is Part of Speech tagging. On the basis of statistical information and rules of transformation, Metaheuristic algorithms have been used to assign optimized tag of rules. Part of Speech tagging is a complex task in the Natural language preprocessing area. In order to tackle the complexity of POS tagging, metaheuristic algorithms are of utmost importance.

Literature Review

- Random-Restart Hill Climbing

It is a simple state metaheuristic that is an advanced version of Hill Climbing. It seeks to prevent HC from being trapped in local optimum by performing repetitive explorations in the problem space, which are generated randomly until the stop criterion occurs or a better solution is not found.

- Particle Swarm Optimization

Population metaheuristic motivated by the intelligent collective behavior of swarms in nature. Each potential solution is called a particle, the set of particles is known as a swarm, and the position of each particle changes depending on its own experience and the experience of the swarm.

- Jaya

Population metaheuristic that seeks to find the best solution in the shortest possible time, but is also always trying to get away from failure, generating an optimal balance between exploration and exploitation. Jaya is a novel, simple and efficient algorithm for optimization problem solving with and without Restrictions.

- GBHS Tagger

GBHS Tagger includes knowledge of the tagging problem using a local optimizer, adapted from the Hill Climb (HC) metaheuristic, which is applied to the best harmony in harmonic memory (HM). In addition to the GBHS parameters, three more parameters are defined: ProbOpt, which controls the percentage of times the local optimization process is carried out; MaxNeighbors, which defines the number of neighbors used in the local optimization process, and the parameter Alpha, which controls whether the components of each harmony in the population are randomly generated from their possible tags or taken from the tag with higher probability.

Significance of work

The elaborated methodology proves the good accuracy of metaheuristic algorithms with good results. Thus Metaheuristic Algorithms should be used for POS tagging

on traditional as well as non-traditional languages. This is a good optimization technique and improvement for other proposed techniques to improve POS tagging.

7) Automatic Generation of Assessment Test Items from Text: Some Quality Aspects: [16]

Introduction

This paper presents the automated generation of assessment test items from natural-language text. The main idea is to extract fragments from the main text and transform them into questions or test-items. The optimized procedure for filtering text segments and selecting key term to blank out is discussed in the generation of Assessment test-items.

Literature Review

Quality aspects of Sequential approach of text processing is listed here:

- Text preprocessing

It is the conversion from a raw text file into a well-defined sequence of linguistically correct texts. It consists of two processes: document triage and text segmentation. In Document triage, document is converted from a digital file into a well-defined text document. It is accomplished using software tools. This stage is completely technical.

- Segment filtering

Every text sentence is not certain to be a potential sentence to frame questions. For qualitative automated evaluation, proper filtering of acquired sentences should be considered so that it has the most salient segments. Various features like sentence-length cut-off (short sentences are excluded), use of cue phrases (inclusion of sentences with phrases such as "in conclusion"), sentence position in a document/paragraph, occurrence of frequent terms (based on TF-IDF term weighting), and occurrence of title words.

- Test-item generation

The main problem here is to which word to blank out to generate questions. Here blanking out special items rather than common words is given much importance.

Example:

Source: Ram has made a beautiful scenery with a perfect blend of colours using crayons.

Result: Ram has made a beautiful —— (what?) with a perfect blend of colours using crayons.

Another issue, which arises at this step, is that the processed sentences may contain anaphora. Without an implementation of automatic anaphora resolution, the user could resolve the anaphora manually (e.g. to replace pronouns with corresponding nouns) using the in-context display of the processed sentence.

Significance of work

Based on the methodology, several quality aspects - Text preprocessing, Segment filtering and test-item generation is implemented. For the anaphora issue, we have implemented

a source clause selection filtering out the sentences which have sentences starting with pronouns like He,She.Thus Anaphora issue has been resolved in the method we have implemented in our final method.

8) Generating Natural Language Questions to Support Learning ON-Line: [17]

Introduction

When teachers prepare questions from learning materials, they develop accompanying questions to test students. Natural language processing tools are most commonly used but mostly, semantic correctness of question is not ensured. The method guides about the semantic role labelling techniques to overcome the semantic correctness. It is a template based approach that has gained much accuracy and importance over years.

Literature Review

Regardless of the method used, the procedure must perform four phases:

- Content-selection: selecting spans of source text from which questions can be generated.
- target-identification: determining which specific words and phrases should be asked about.
- Question-formation: identifying the suitable question from the content selected.
- Surface form generation: producing the final surface-form realization.

The main method here used is semantic based using the techniques of semantic role labeling(SRL). Semantic role labeler identifies the semantic entities associated with each predicate. Set of SRL includes ArgM-LOC (location), ArgM-EXT (ex-tent), ArgM-DIS (discourse), ArgM-ADV (adverbial), ArgM-NEG (negation), ArgM-MOD (modal verb), ArgM-CAU (cause), ArgM-TMP (time), ArgM-PNC (purpose), ArgM-MNR (manner), and ArgM-DIR (direction).

After the process of tokenization and POS tagging, a single word is first paired with a semantic role label to ensure the semantic correctness of the question generated. For example, AM-LOC can be used to generate a where question, an AM-TMP can be framed as when question, AM-PNC can be framed as a why question, AM-MNR can be framed as How question.

Significance of work

Method describes a template-based method using predominately semantic information can be used to generate natural language questions in an online learning system. The mechanism supports rich selectional semantically correct words to generate questions thus improving the accuracy to a great extent. The method resolves the type of wh-word to be substituted in the place of question.

9) Resolving Syntactic Ambiguities in Natural Language Specification of Constraints: [3]

Introduction

The research has mentioned the shortcomings of Stanford parsers to not be able to handle particular syntactic ambiguity. There are two types discussed: attachment ambiguity and homonymy. A well-detailed approach using the novel technique of UML Diagram is proposed to resolve the problem faced by Stanford parsers.

Literature Review

Stanford POS tagger and parser is used to tag the English text and is capable of 90.7 percent accuracy. Accuracy of the Stanford parser is low for real-time applications.

The two types of Syntactic ambiguity are:

- Attachment Ambiguity
In the example: The pay is given to all the employees with bonus, the ambiguity is due to the prepwith phase. There can be two meanings of this same text. One meaning can be that pay is given to those employees who got bonus. Another meaning can be that pay with bonus is given to all employees. Thus prepwith phase can have two meanings depending on its attachment.
- Homonymy
It is a type of syntactic ambiguity in which a token in a sentence can exhibit one or more tag of speech and thus making the process of POS tagging ambiguous. For example: A customer books two items.
Tagging: [A/DT] [customer /NN] [books/NNS] [two/CD] [items/NNS] [./.]
Here 'books' word is misclassified as 'NNS' by Stanford Parser whereas here in this sentence, 'books' should be categorized as VBZ.

Solution for Resolving Syntactic Ambiguity:

- Addressing Attachment Ambiguity
UML diagrams with the correct mapping in between classes representing different entity of the sentence can resolve attachment ambiguity.
For generating correct dependencies of input English sentences, information is written above the arrow mapping the classes.
- Addressing Homonymy

UML class model elements	English language elements
Class names	Common Nouns
Object names	Proper Nouns
Attribute names	Generative Nouns, Adjectives
Method names	Action Verbs
Associations	Action Verbs

For example, A customer books two items, through the UML diagram listed below, one can clearly tell the books is a verb here.

Significance of work

The primary objective of the paper was to address the

challenge of resolving various cases of syntactic ambiguity such as attachment ambiguity and homonymy. By resolving the said cases of syntactic ambiguity the accuracy of machine processing can be improved. To address this challenge, the method has presented a NL based automated approach that uses a UML class model as a context of the input English (constraints) and by using the available information in the UML class model (such as classes, methods, associations, etc) we can resolve attachment ambiguity and homonymy.

10) Answer Evaluation using Machine Learning: [1]

Introduction

Manual answer evaluation is a time-consuming task and requires a lot of human intervention. Also, there is less chances of fair marks given to same answers. Thus the system proposed here will first tokenize the answer in words and then evaluate answers using machine learning tools performing automatic evaluation of answer. Only one has to scan and give answer sheet to the system, the algorithm will extract the suitable features from the user provided answer to match it with the correct answer. Such system is required to provide eligible and fair marks. This will reduce the time and cost of institutions that they pay for manual evaluations.

Literature review

The algorithm assigns marks on basis of :
The Number of keywords matched and the Length of the answers.

Significance of work

The proposed system can evaluate 5760 answers in a day whereas a human working for 8 hours can evaluate 480 answers a day. The scores are calculated for 10 students. The difference between manual evaluation and system evaluation is very close.

11) Grading Descriptive Answer Scripts using Deep Learning (Neethu George, Sijimol PJ, Surekha Mariam Varghese): [25]

Introduction

The proposed methodology in the research paper consists of LSTM-RNN layer which takes the vector representing each word in the sentence. The main objective is to extract the semantics to effectively understand the relationship in between the text. Ellis Page developed the Project Essay Grader using deep learning models. Leacock and Chodorow have developed C-rater. C-rater is an automated short answer scoring system.

Literature Review

The model is a combination of NLP and Machine learning. Using LSTM-Recurrent neural networks, the model learn the relationship in the text. LSTM was developed to overcome the shortcomings of Recurrent Neural Network.

Deep Descriptive Answer Scoring model is used to predict the answers.

Long Short Term Memory(LSTM) is an artificial recurrent neural network architecture. Unlike standard deep Learning models, LSTM has feedback connections. It has operations to remember or forget information. It can evaluate entire sequence of input data. LSTM are advanced versions of Recurrent neural networks.

Different steps in the proposed model is listed:

- Preprocessing

Preprocessing phase takes the relevant features of the text and converts it into vectors. Using tokenization and Lemmatization, features are extracted. The score corresponding to the feature is represented as a one-hot vector.

- Sentence Embedding

The embedding layer simply converts the feature corresponding to a word into glove vectors. The main aim of Using LSTM-RNN is to create a simplified and low-dimension semantic values by sequentially and recurrently. Processing and finding the corresponding values for each token in a sentence.

- Grading

Based on the embedding vector of the LSTM-RNN made in the second phase of Sentence Embedding, the output Layer, fully connected neural network layer, the dense layer will predict the one-hotted score. Keras library is used here.

Significance of work

The goal of the Deep Descriptive Answer Scoring model is to replace a system where marking and evaluation depends on factors like human mindset ,time, thinking style etc. The results predicted in the model was compared with the previous models and the method proposed here give better results thus improves accuracy.

12) Automated Essay Scoring: [9]

Introduction

Automated Essay grader has proved a boon for educational institutions. Responding bulk of students is a hectic and time-consuming process for teachers particularly in today's pandemic where a large and frequent number of assignments is manually evaluated. Many AES systems are used to overcome time, cost, and labor. The need for emerging automated evaluation research is in progress and many research have been put forward seeing the growing demand for it. Among these, majorly used is the Automated Essay scoring system. There are four types of AES system.

Literature Review

- Project Essay Grader(PEG)

On the request of the College Board, Ellis Page in 1966 developed Project Essay Grader system. It was the first

AES system to be built. Indirect measures of writing skills were considered thus it made the PES system less traceable. Along with several parsers and dictionaries, special collections and classification criteria were also used.

- Intelligent Essay Scorer(IEA)

It analyzes and scores using a semantic text analysis method called Latent Semantic Analysis. A psychologist Thomas Landauer developed the LSA approach. Latent Semantic Analysis is a method that maps semantic similarity between the words in the text. In the LSA based approach, we input the text in the form of a matrix of size (number of words in the text) * (number of features/context with which we want to analyze each word). Row in the matrix represents each word and column in the matrix represents the context./feature. Value in the cell consists of the frequency of word corresponding to the row and then the value is re-evaluated considering the frequency of the word and its importance with respect to the context mapping the particular column.

Significance of work

AES systems can be a great assistance to teachers in responding to a large number of essays and assign frequent writing assignments without worrying about scoring the first and subsequent drafts. There is a common methodology among four Automated Essay Scoring system-Project Essay Grader, Intelligent Essay Scorer, E-rater and Criterion, IntelliMetric, and MY Access. They all need to be trained on a large number of essays in order to make the learning of the system strong. Next, all systems provide strong feedback with the best possible scoring thus making the task of teachers easier and flexible.

13) Automatic Assessment of Descriptive Answers for Online Examination using Semantic Analysis: [20]

Introduction

Given a text, a question-answering system is the best way to extract information from the text. The method proposed here provides the semantic-based QA evaluation using an Artificial neural network. Two major issues-length paraphrasing are considered in evaluating subjective assessment. After pattern checking using semantic analysis, a confidence factor of positive or negative is concluded.

Literature Review

The phases of the method proposed:

- To find a set of suitable relevant answers for all questions in the question set.
- To identify the patterns of answers to match with the user answer.
- Implement stemming, stop words removal, part of speech tagging-all natural language processing pipeline tools to preprocess the text. Once unwanted words are

removed from the user answer, they are prepared to feed into a neural layer as a matrix.

- Classifying the answer using the criterion of semantic weights. To check for other relevant words, a lexical resource WordNet is used. WordNet is a lexical database dictionary NLTP toolkit majorly used to extract the relevant word considering a certain meaning. Here Artificial Neural Network is used to find the relation between the text using the sigmoid function. Long-Short Term Memory Network is used here. Feedback connections in LSTM retain the memory of the pattern in the sequence of inputs and thus gives better results. LSTM is an advanced-technology version of Neural networks.
- On the basis of length, the number of keywords, relation of the sentence formation, score for the answer of candidates can be formulated. Final results are displayed with parameters such as accuracy and false-positive ratio.

The database used here is Mysql. It is a relational database management system to store the data of the questions Corresponding to the correct answers. Eclipse an integrated development environment is here used. It contains a Base workspace for providing the interface to the code editor.

Significance of work

The system is executed on java 3-tier architecture framework with INTEL 2.7 GHz i3 processor and 4 GB RAM with a supervised learning approach. The system is tested with around 100 samples of the question-answer dataset using the proposed NLP(Natural language processing) and ANN(Artificial Neural network) and a considerable accuracy is achieved.

14) Synonyms Paraphrasing using WordNet and Internet: [6]

Introduction

When the answer is checked in reference to the correct answer, synonyms or same-meaning words that can replace each other in the text are required to look-up. Wordnet is a lexical Natural language processing tool database of Nouns, adjectives, verbs, and nouns used for the reference of words in respect to a certain meaning. WordNet is best suitable for synonyms paraphrasing which helps to generate a slightly different version of the text and produce a bucket of synonyms for each word.

Literature Review

Paraphrasing has different meanings with respect to the context.

Various Types of Paraphrasing:

- Text compression
Among the bucket of words, the token with the shortest length is taken into consideration to compress the word. This gives an advantage in space structure. The shortest synonym is taken in each synset (either independently of any statistical evaluations or selecting from the words

that passed the marginality threshold). This gives a significant gain in space only when there are abbreviations (s) among absolute synonyms.

- **Text canonization**

For this, the most frequently used synonym is taken. Of course, it may prove to be the same one as in the source text. It is also useful for persons with limited language knowledge, i.e. for foreigners or children, since this renders texts in a more intelligible way.

- **Text simplification**

Any text will be more intelligible for a language-impaired person if we select among synonyms a “simpler”. It is not always the most frequently used synonym.

Significance of work

Synonyms can be used as better improvements of certain words in the text that will make text compressed and canonized. Synonyms paraphrasing is an important task in the Natural language processing pipeline. For each word, we generate its relevant bucket of possible variants. The second application is keeping certain words secret. The area of cryptography where certain information needs not to be as exactly translated.

15) *Extracting Word Synonyms from Text using Neural Approaches:* [19]

Introduction

The Natural Language processing domain has been advanced by applying it to Word Synonyms technology extraction. Word2Vec tools have been used formerly to produce semantic relationships in the text. Word2vec uses a Neural network algorithm to compute the semantic relationship in between the text. Basically, it uses a two-layer neural network with a large piece of text as an input and a vector of more than a hundred dimensions containing numerical values referencing certain features as output. Word2Vec is a vector representing each piece of text with several numbers depicting a certain feature of the text. But it faced several challenges. The new technique is formed by using a two-step process. First Word embedding is built and then word embedding is feed to the neural network using the annotated dataset.

Literature Review

Synonyms refer to different words depicting the same meaning. The methodology aims to find word synonyms that can be substituted in place for each other but not every synonym is replaceable. For example, big and huge. ‘Rita has a big sister’ cannot be replaced by the sentence ‘Rita has a huge sister’. Thus Leeuwenberg discovered this problem and redefined synonyms as the word that depicts the same meaning in-reference to the text and can be used interchangeably. Thus it is evident that other factors need to be considered for synonymy. It is quite rare that a certain synonym is replaceable in context to text. Mitkov addressed

the synonymy as a classification task.

The SimLex-999 similarity score is used to compute the semantic relations in the text. A similarity score greater than 6.5 is a gold standard to extract synonyms pairs from the text. SimLex-999 is used to train the neural network model. SimLex-999 is constructed manually by paying 500 native speakers to tell semantically correct synonym pairs. For example, a similarity score of lovely and beautiful was 9.8, word-pairs like small-big, old-new received low neural network scores less than 6.5. The model succeeded to give relation in the piece of text- synonyms, antonyms, domain-similarity, Association, less antonyms, perfect antonyms, less synonyms, perfect synonyms, etc.

Significance of work

The paper proposed a neural network algorithm using backpropagation to find the semantic correctness in the piece of text. Mitkov shows that similarity does not always mean semantic relation in the text. The two-step neural architecture using Word Embeddings and SimLex-999 shows that the model achieved high accuracy. A score greater than 6.5 is considered to generate a high-synonyms similarity pair of words. Rich database WordNet fails to provide such results as it does not contain contemporary compound adjectives. The most common approach discussed here is the Continous Bag-of-Words model and Skip-Gram Model.

16) *Improving Language Understanding by Generative Pre-Training:* [22]

Introduction

Different sentences may have the same meaning as well as sentences formed by similar words have different meanings, so language understanding becomes the major step-stone to cover. In order to improve language understanding, OpenAi proposed the method of extensive use of pre-trained word embeddings. Also develop a semi-supervised approach for language understanding tasks using a combination of unsupervised pre-training and supervised fine-tuning.

Literature Review

Their framework contains two stages. The first stage is learning a high-capacity language model on a large corpus of text. This is followed by a fine-tuning stage, where they adapt the model to a discriminative task with labeled data.

Training phase steps:

- Pick a corpus of tokens and perform unsupervised pre-training.
- In unsupervised pre training ,gradient decent is used.
- Do supervised fine-tuning.

Significance of work

We are working on better understanding of the answer given by the student, so this research paper gives the way of achieving strong natural language understanding with a single task-agnostic model through generative pre-training and discriminative fine-tuning. Also the concept

of transformers are very well explained in the research paper.

17) *The Curious Case of Neural Text Degeneration:* [14]

Introduction

Efficient question generation is one of the major milestones. In order to overcome this milestone, Paul. G, HOD CS, Cape Town proposed a Nucleus Sampling to effectively draw sentences or questions from raw text. The approach avoids text degeneration by truncating the unreliable tail of the probability distribution, sampling from the dynamic nucleus of tokens containing the vast majority of the probability mass.

Literature Review

The framework proposes a new stochastic decoding method: Nucleus Sampling. The key idea is to use the shape of the probability distribution to determine the set of tokens to be sampled from. The probability mass assigned to partial human sentences. Flat distributions lead to many moderately probable tokens, while peaked distributions concentrate most probability mass into just a few tokens. The presence of flat distributions makes the use of a small k in top-k sampling problematic, while the presence of peaked distributions makes large k's problematic.

Significance of work

The paper guides us about the most common decoding methods for open-ended language generation. Also Nucleus Sampling as a solution that captures the region of confidence of language models effectively which overall helps us in generating questions from text.

18) *Encoding Word Order in Complex Embeddings:* [27]

Introduction

In order to form a good sentence, selection of words and their order are extremely important. Benyou Wang, University of Padua proposed a method using neural networks and position embeddings to maintain the sequential structure of language. He presented a novel and principled approach to model both the global absolute positions of words and their inner sequential and adjacent relationships. Easy transfer has made a major contribution by proposing a scalable and flexible platform to make it easy to develop deep TL algorithms for data science and large scale industrial applications.

Literature Review

Whole paper works at two stages:

- Modelling Word Order in Embeddings Space
 - Extending vectors to functions
 - Properties of the functions to capture word order
 - Complex-valued word embedding
- Experimental Evaluation
 - Text Classification

- Machine Translation
- Language Modeling

Significance of work

The primary objective of the paper is to make better nlp by encoding words sequentially or in order. This approach really develops a major role in how embedding also used to form a sequenced sentence. They extended word vectors to word functions with a variable i.e. position, to model the smooth shift among sequential word positions and therefore implicitly capture relative distances between words.

19) *Mimic and Conquer: Heterogeneous Tree Structure Distillation for Syntactic NLP:* [11]

Introduction

Department of Key library investigated Knowledge Distillation Method to integrate the heterogeneous structure into an LSTM. After experimenting on four different phases, the syntax-based tasks outperforms the ree encoder performance by continuous integration of rich heterogenous structure-based syntax. The experiment improved the accuracy to a great extent. Semantic analysis has gained much importance in NLP by depicting the syntactical information. By integrating the representation from tree structures, the SRL phase integrated the features and structures of two subjects. Experiments verifying the syntax-dependent tasks integrate semantic correctness. The knowledge distillation model is divided into two phases-output distillation which makes a teacher model output logits as a student model and another feature distillation which allows a student to make conclusions from the teacher's intermediate feature representations.

Litreture Review

Child sum TreeLSTM is processed to make a basic dependency structure. After execution of semantic learning and output distillation by each teacher, sequentially one teacher is distilled at a time. Turning gap g2 is used to process the dependency or constituency injection from a tree teacher. This is done to optimize the training process. The overall line is optimized after a certain number of training iterations. A 3-layer BiLSTM and a 2 layer architecture are used for the training mode. The default word embeddings are initialized randomly. It is a 300 dimension vector. The hidden size is set to 350 in the student in LSTM and 300 in the teacher models. Adam optimizer is used. Coefficients are set as 0.6, 0.2, and 0.5 for the training process. On the basis of the F1 score, correctness in the syntax was calculated. Constituent and dependency representation for a particular sentence represents POS tagging, the semantic labeling depicting the mutual dependence of each word in a text.

Significance of work

The knowledge distillation method is described for making the NLP tasks easier by distilling syntactical features of text

easily. The model outperformed the existing tree encoder models. LSTM model is used to predict the word based on its several features. And thus LSTM is trained to learn the semantic features of the text.

20) *Learning to Ask: Neural Question Generation for Reading Comprehension:* [10]

Introduction

An attention-based sequence learning model is designed in Zhiyuan College to frame an automated question generator system from reading comprehension. Human intervention in generating Question answer system is a time-consuming process. The process does not use former Natural language processing tools. It is an end-to-end trainable model. The major application can be used in chatbots like asking questions to start a conversation. Reading comprehension is a challenging task for machines. Recently many datasets have been developed to train the machine for automated question generation. For example, bAbI is a dataset featuring 29 different tasks..

Litreture Review

The well defined neural network approach uses paragraph-level information. Many researchers have focused on pure rule-based approaches for question-generation but these approaches required pre-defined rules. Heilman and Smith in the year 2010 proposed a method to improve the rule-based approach using feature set and ranking algorithm. The task is to generate a question y of a given sequence length based on the input sentence x . The associating information in between the text is considered while framing the question. Two level-sentence level and paragraph-level information are encoded to generate a question. LSTM- Long short term memory recurrent neural network architecture is stored which analyze the semantic relation in the input sequences. The softmax activation function is used. The attention bases sentence encoder is used in both sentence and paragraph based models while the paragraph encoder is used in the information that uses paragraph-level information.

Significance of work

The model is trained with the dataset containing 100,000 questions about 536 articles. The dataset is created by employing Amazon Mechanical Turks crowd workers to frame questions. They were asked to make slightly variated questions for the better training of the model. In deep learning recurrent neural network LSTM, hidden unit size is set to 600 and number of layers of LSTMs to two. Stochastic gradient descent algorithm is used with learning rate at epoch 8, batch size as 64. The professional English speakers were asked to rate the model on a range of 1-5 and then human raters were asked to give a ranking of questions. The model achieved good ranking and high accuracy. The two main criterion used in the model were naturalness and semantic relation in the question.

21) *Contextual LSTM (CLSTM) models for Large scale NLP tasks:* [13]

Introduction

The main contribution of the paper was an Contextual Long Short Term Model (CLSTM). The model was supposed to improve the word prediction, scoring of the next sentence, and, prediction of topic sentence. Experiments done in CLSTM model shows the improvement compared to LSTM model and have significance on wide variety of NLP applications like sentence completion, paraphrase generation, dialog systems. In the experiment, we trained the model using two large document (English corpora) from Wikipedia, and a recent screenshots of Google News document. Having total vocabulary of 140K in Wikipedia and 100K words in Google News. And the experiment in the paper demonstrate, that the power of topics with world level features shows decent improvement in the LSTM. CLSTM shows improvement of 21 and 18 percent respectively on Wikipedia corp and Google news.

Literature Review

In the approach, As we have already introduced that we are working on the improvement of LSTM model, LSTM is an rnn model that is useful for capturing long short term dependency in the data given. The model has multiple layers of neuron, where every neuron is like digital memory to store the information over the time training and testing. Initially, the LSTM was implemented over Google news corpus. The model takes one hot encoding from the input, converts it into the vector, and consumes vector one at a time. The model is trained to predict the next word in the sentence, on the already seen sequence of words. The core algorithm of LSTM uses BPTT, using softmax layer that gives the next word ground truth.

To adapt the LSTM cell that is taken from CLSTM cell, taking two inputs words and the topic and given to the modified equation representing the CLSTM cell. Note that we have added the topic input to the every LSTM cell as each LSTM cell can represents the different topic. For example : when the topic is based on the sentence segment seen so far, the topic is based on the current sentence prefix.

Significance of work

The experiment in the paper clearly demonstrate, that the power of topics with world level features shows decent improvement in the LSTM. CLSTM shows improvement of 21 and 18 percent respectively on Wikipedia corp and Google news. The CLSTM improvement on the state of art of LSTM by 2-3 percent and the improved accuracy of about 20 percent over the old LSTM. these gains are quite significant and we get similar gains on the Google news as well. which shows the span of the algorithms. The improvement obtained by using the CLSTM model has major increase in the performance improvements in many NLP applications, ranging from sentence completion, question and answering, and paraphrase generation to

different applications in dialog systems.

22) *Improved Neural Relation Detection for Knowledge Base Question Answering:* [30]

Introduction

The paper proposed to build an both relation and world level relation level representation and used the bidirectional LSTM (BiLSTM) to learn the different level of questions in order to compare the level of relation information. Finally, the paper propose residual network for learning the sequence matching, that makes the training easier and result in more depth question representation. The proposed KBQA composed of two step relation detection. Given an input question and a set of candidates linked with the question. The proposed relation detection plays an important role in the KBQA algorithm. Re-ranking the entity candidates according to the high confident relations detected from the raw question helping in finding the core relation for every entity candidate. Secondly, the selection from a much smaller candidate entity set after re-ranking.

Literature Review

Relation detection is important for any NLP applications mainly knowledge base question Answering (KBQA). We propose a hierarchical rnn improved using the residual learning which detects KB relations given an input question. Our method uses deep residual bidirectional LSTMs use to compare questions and relation at a different levels of abstraction. And, the paper propose a simple KBQA system that connects entity linking and the proposed relation to make the two components improves each other, And the experimental results shows that our approach not only achieves outstanding relation detection performance, but importantly, it helps in KBQA system to achieve state-of-the-art accuracy.

Significance of work

KB relation detection is an important step in KBQA and it is very different from general relation extraction methods. Paper has proposed a KB relation detection model,that performs matching between questions and KB relations. Model proposed in the model out performed the previous methods on KB relation detection and allows our KBQA system to achieve state-of-the-arts. Also it will help in investigating the integration of our model into end-to-end systems. For example, our model could be integrated with the decoder to provide better sequence prediction. We will also investigate datasets like graph question and complex questions to handle more characteristics of general question answering.

23) *MATINF : A Jointly Labeled Large-Scale Dataset for Classification, Question Answering and Summarization:* [29]

Introduction

In this paper, Maternal and Infant dataset (MATINF), the first that converts three major NLP tasks, text classification, question answering, and, summarization. MATINF works really well on question answering. MatInf supports three NLP application as follow, In Text classification it chooses fine grained category list, different from previous news classification, MatInf distance between different categories is similar, which provides more challenging stage to test the state-of-art of the model. For Question Answering Given a question, the task is to produce an answer in natural language, this task is slightly different from machine reading comprehension since documentcontains the correct answer are not provided. Therefore, how to collect the domain knowledge from massive QA data becomes extremely important.

Literature Review

On large datasets it has made development in nearly all domains of NLP. However, there is currently no cross task dataset in NLP, which affects the development of multitask learning. This research proposed MATINF, the first jointly labeled large scale dataset for question answering and summarization. It contains 1.07 million question answer pairs with categorized by user generated question descriptions. On the basis of these knowledge, it works in three major NLP tasks, including classification, question answering, and summarization. It benchmarked previous techniques and methods to inspire further research. Our comprehensive comparison and experiments over it and other datasets demonstrate the merits held by MATINF .

Significance of work

The significance of this work in our project, the MATINF helped jointly in labeling large scale dataset for classification, question answering and summarization. The benchmark improved by this method from previous existing methods and a straightforward method with a multitask paradigm on MATINF and analyze their performance on these three tasks. Our extensive experiments proves that the potential of the proposed dataset accelerates the results in the three tasks and question answer learning.

24) *What BERT Sees: Cross-Modal Transfer for Visual Question Generation:* [24]

Introduction

This version of BERT is very popular in NLP. The text based question generation has been very common application of NLP. However, the text generation that is creating questions with image as an input has been a studious of at task. The paper presents BERT to generate visual based questions. It was first introduced the VQG task in the neural self talk model. The methodology investigated if BERT can give good results on the input is given beyond textual data. VQG task has several application, chatbots can ask to start a conversation with the user on the basis of modeled data. Since BERT was not previously developed for visual

content, it is challenging for BERT to deal with the data present in the image format.

Literature Review

Datasets were used to train the model. The dataset consists of 5000 images with 500 questions per image. A variety of methods have been used, Most of use a standard rnn and a CNN encoder. Then input is feed to GRU to output the result. The result are type of questions starting with what, why, when based on a certain image. Cnn process the image and LSTM process the caption to convert visual information to textual data to generate questions. Faster-RCNN first process the images on the pre-trained on Visual Genome. In the result, each image represented as a sequence of data. Thus, BERT attention can be calculated at image-level. Three types of inputs can be considered like given-text, image only and image caption.

Significance of work

This research becomes useful whenever we deal with the data present in the form of image, BERT model shows the excellent performance, ensuring the benefit of pre-trained dataset models. Even training on the data and freezing the weights of the model, it performed well. This shows the scope of BERT working as a model beyond normal text. Due to the flexibility of the model, with image only as input, the model is found to be improved further resulting in the better output. Data in the form of image caption as an input, produces the highest accuracy was obtained. This shows that BERT performs really well in textual data as well as visual data. It is a strong model to generate questions from data present in the form of image.

25) *EasyTransfer – A Simple and Scalable Deep Transfer Learning Platform for NLP Applications:* [21]

Introduction

The paper presents a EasyTransfer platform designed to develop Transfer learning applications for NLP applications. Transfer learning is used to deal with complicated dataset in place of Machine learning applications. Today major applications use transfer learning due to complexity of the dataset and method used. Large immense research is carried to make easy to apply transfer learning.

Literature review

A few Transfer learning toolkits- NVIDIA Transfer learning kit, Amazon Xfer, Tsinghua transfer learning toolkit, The Huggingface toolkit have been designed to make it easy to apply TL algorithms. However, the challenge is that these toolkits fail to give good results when it comes to industrial applications. It is due to the complexity and huge size of deep learning models. The Easy Transfer architecture is a scalable source and platform to make it easy to apply transfer learning algorithms. The architecture consists of both low level and high level layer API for users to build their own models. The Easy transfer toolkit supports Model Fine-Tuning, Feature based transfer learning, instance based

transfer learning and Model based transfer learning.

Significance of work

There are a variety type of TL toolkit algorithms available yet no kit is available for users to examine different types of state of the art tL algorithms. To overcome thus, Easy toolkit has been designed and is available to the open source platform. EasyTransfer has to integrate into a number of deep learning products in Alibaba. Easy transfer has made a major contribution by proposing a scalable and flexible platform to make it easy to develop deep TL algorithms for data science and large scale industrial applications.

26) *The Question Answering System Using NLP and AI International Journal of Scientific Engineering Research Volume 7, Issue 12, December-2016:* [28]

Introduction

Today there is a need for a system of automatic evaluation where after reading certain texts, students find a need for a platform where one can analyze his/her understanding and real time-consuming. Manually generating question-answer is a very time consuming process for teachers. Thus this platform makes the work of the educational system much easier and flexible.

Literature Review

The questions can be divided into two categories-Factual and Expert. Questions starting with wh-words like what, when, where, how are factual whereas One starting with how, why are expert questions.

Approaches in QA:

- Linguistic Approach

The phase here comprises of Part of Speech, tagging, parsing, and natural language texts. The questions extracted from this approach have semantic correctness. Part of Speech tagging is used to pair the tokenized words to grammatical categories like noun, adjective, verb, adverb.

- Statistical Approach

Techniques like Support MACHine Classifier, Bayesian Classifier is used statistically to formulate the question. It is due to the huge amount of data on the World Wide Web.

- Pattern Matching Approach

A well-formulated pattern approach helps in finding a suitable answer with good results. The phase makes use of the pattern extracted from the text and compares the features of user answers from the correct answer.

Significance of work

Using the three approaches, the methodology deals with online platform AI and NLP approach to provide a flexible Question answer approach. This has improved our quality of the questions to a great extent and our work has addressed the shortcoming listed in the Research paper that students who simply memorize the chapter without actually

understanding it if it's directly framed without paraphrasing the text.

27) Comparative study on Python web frameworks: Flask and Django : [12]

Introduction

Process of integrating software applications that can run on websites. Even though web application development follows the software development process, the technology and the architecture used for it are quite different. The software application that runs on a personal computer (PC) might not depend on the internet in contrast to the web application that depends upon the remote servers. .

Literature Review

Front End Technologies: HyperText Markup Language is the basic toolkit to start for the web development field. The skeleton of a web application is framed using the complete attributes and tag list in HTML. The tag starts with `<tag>` and closes with angular close brackets `</tag>`. There are plenty of tags and attributes supporting the web design for different purpose-color, input boxes, submit buttons, etc. Cascading Style Sheets make web design flexible and easy to present. With the advanced features added in CSS, templates in CSS can be applied to a single web page very easily.

Python Web Frameworks:

- Flask

Flask is a web-development flexible micro-framework used extensively for web-development. The framework is majorly used for the pure back-end. It is discovered by an international bunch of python programmers. It is an easy to deploy, extensible framework. Its importance is proved by its frequent and growing use of Django over other frameworks.

- Django

Adrian Holovaty and Simon Willson 2003 developed Django. In order to schedule the applications, Django was used. In 2005, it was released an open-source in 2005. Django is compatible with Python versions 2.6.5 and 2.7 but Djangodjango also offers mental support from version 3.2.3 to 3.3. Its main goals are simplicity, flexibility, reliability, and scalability. Django has its own naming system for all functions and components (e.g., HTTP responses are called "views").

Significance of work

Every front end technology and framework is built with a motive to strengthen the technology. The article supports the web application growing need with elaboration on front-end technologies-HTML, CSS and web frameworks-flask, and Django. Based on the comparative study, it was found that Django is the best fit for large-scale projects whereas Flask is more suitable for small projects. Flask can be learned and set up quickly, but when it comes to managing

and maintaining, it requires more work than the former frameworks.

28) Making It Work for Everyone: HTML5 and CSS Level 3 for Responsive, Accessible Design on your Library's Website

Stewart Baker, Western Oregon University: [4]

Introduction

The article talks about the web application feature seeing the immense need for web applications and the potential of web developers. Building a simplified web application using tools-HTML, CSS, and javascript method have been so popular today. The excellent frameworks in the emerging technology make website universally accessible and even user-friendly.

Literature Review

- HTML5

HyperText Markup Language is the basic toolkit to start for the web development field. The skeleton of a web application is framed using the complete attributes and tag list in HTML. The tag starts with `<tag>` and closes with angular close brackets `</tag>`. There are plenty of tags and attributes supporting the web design for different purpose-color, input boxes, submit buttons, etc. Hyperlinks are also used here. Adding spaces, forming a new paragraph, modifying a particular word to bold, italic, underline. There are H1 Attribute specifying a range of sizes for heading. HTML5 supports multimedia, audio and video links. HTML is the most flexible, easy to learn and implement markup language for beginners. It is a powerful tool for web development. Once one starts to learn the basic HTML, with strong massive features of HTML, web design can be made very strong and powerful.

- CSS

Cascading Style Sheets make web design flexible and easy to present. With the advanced features added in CSS, templates in CSS can be applied to a single web page very easily. CSS is for the content presentation of the web page. This is done using an external style sheet. CSS provides portability, easy layout with multiple numbers of features. Since the content is completely separated from the design, changes across the website can be implemented all at once. This reduces delivery times and the costs of future edits.

Significance of work

The paper provides the tools and techniques commonly used in web development. In our project, we have integrated our project as a web application using HTML and CSS listed in the Research paper. Brophy and Craven have highlighted the overview of tech stack in web application and its emerging need in the growing digital world.

29) RepoAI: A Novel Approach Towards Automated Project Reporting

Department of Computer Science Engineering, BNM Institute of Technology, Bangalore: [18]

Introduction

In order to evaluate students' knowledge about the content, it is important to take some questions out of the text but about the topic. The Department of CSE from BNM Institute proposed a method to web scrape data and develop some questions using nlp which helps to evaluate students better.

Literature Review

There are some steps of implementation-

- Backend

The development and deployment of the backend part is done in a number of steps which is described in great detail.

- Login and Signup

The data entered by the user is filtered to check for SQL Injections and then passed on to the FireBase API. The response from the FireBase API is sent back to the user through our API.

- Google Scrapping

The backend development starts with the usage of Python. Some background checks are done at this stage. If an article on the input query is present on Wikipedia, then it jumps to step 3. Else, a predefined set of topic headings are employed namely [“Introduction”, “Cause”, “Effect”, “Precautions”, “Applications”, “Conclusion”]. Subsequent Google searches will be performed on the predefined headings. Taking an example of ‘Marine Pollution’ as the input string, Google searches on “Marine pollution introduction”, “Marine pollution causes” etc., will be performed.

- NLP Wikipedia Scrapping

The input query which is received by the API backend script is cleaned by using NLP for filtering out unwanted words and leaving behind only essential keywords. These keywords are used to find related Wikipedia pages, if any, and the Table of Contents of that particular Wikipedia page is scrapped and a list of necessary headings again after cleaning and filtering is returned back to the user.

Significance of work

A report is fundamentally a document that gives out information on a given topic in a methodical fashion. Writing the report is not as overwhelming as some may initially think, more so, for school students. A student's interaction with prevailing literature often requires perusing a substantial amount of text. In this paper, a novel method using web scraping and NLP have been proposed to summarize project reports automatically so that students can perform their tasks by consulting shorter summaries instead

of entire literature.

30) Efficient Way Of Web Development Using Python And Flask

Fankar Armash Aslam, Department Of Computer Engineering Kalsekar Technical Campus Panvel, India: [2]

Introduction

The web application is the most emerging technology. The good UI web portal attracts many customers and thus there are many options and frameworks to build the web application. Python is an open source language that makes code simpler and readable. Python community provides a large number of libraries to support machine learning and its integration with flask. Python was integrated by web servers mainly to deal with incoming traffic on the server. Flask is a micro framework of python to build web applications. It allows more plugins and functionalities which makes its easy to integrate.

Literature Review

There are advantages of building web applications with Flask. It is extensible, Robust, and open-source. Flask has an extensive easy to read documentation, integrated support, support several cookies, etc. Flask is a lightweight framework and provides a scalable and simple template for web development and the popular thing about Python and flask that it is portable and interactive. Python is one of the most widely used languages. The enormous standard libraries make it compatible to use with any database, web applications, other software, etc. Jinja2 provides a modern and scalable user-friendly template for python..

Significance of work

Every front end technology and framework is built with a motive to strengthen the technology. The article supports the web application growing need with elaboration on the most widely used open-source language-python and scalable and simple web frameworks-flask. Efficient deployment using python and flask has been easy with extensive documentation. Python can be used for making our web application simple, scalable, impressive with the help of an extensively used framework flask template engine.

III. SUMMARY

S.No	Papers	Significance of work-Findings
1	Algorithm for generating questions from the Text	Natural text processing pipelines method Tokenization, Stemming, POS tagging, Structure tree Parser to generate a question-answer system from the text. Lexical and Syntactical Challenges were discussed like Paraphrasing, Pos tagging Ambiguity process, inefficiency to learn connections in the answer sentences.
2	Generating Multiple Choice tests from a Medical Text.	The method here gave us two key points to strengthen our NLP pipeline methods. Importance of answer in the fill in the blanks as a key term. Filtration of inappropriate structures for the main text chosen-Subordinate clause, Negated clause, Coordinated NP, initial pronouns.
3	Single Document Automatic Text Summarization Using Term Frequency-inverse Document Frequency (Tf-IDF)	Text summarization technique TF-IDF algorithm to overcome the challenge of paraphrasing(questions should have a slight variation from the original text) and dealing with complex text by first simplifying the text.
4	A Study on Different Part of Speech(POS) Tagging Approaches in Assamese Language	The technique of assigning an appropriate part of speech tag for each word in an input sentence of a language is called Part of Speech Tagging. It is commonly referred to as POS tagging. It is the main method in our naive process.
5	Hidden Markov Models	The detail on the optimization Model to remove ambiguity in POS tagging. Given an HMM, and a sequence of observations, we'd like to be able to compute P(O—), the probability of the observation sequence given a model
6	Adaptation, Comparison, and Improvement of Metaheuristic Algorithms to the Part-of-Speech Tagging Problem	Using metaheuristic algorithms like Particle Swarm Optimization, Hill Climbing in place of POS tagging for better results
7	Automatic Generation of Assessment Test Items from Text: Some Quality Aspects	Based on the methodology, several quality aspects-Text preprocessing, Segment filtering and test-item generation is implemented. For the anaphora issue, we have implemented a source clause selection filtering out the sentences which have sentences starting with pronouns like He,She.Thus Anaphora issue has been resolved in the method we have implemented in our final method.
8	Generating Natural Language Questions to Support Learning ON-Line	Method describes a template-based method using predominately semantic information can be used to generate natural language questions in an online learning system.The mechanism supports rich selectional semantically correct words to generate questions thus improving the accuracy to a great extent.
9	Resolving Syntactic Ambiguities in Natural Language Specification of Constraints	The primary objective of the paper was to address the challenge of resolving various cases of syntactic ambiguity such as attachment ambiguity and homonymy. By resolving the said cases of syntactic ambiguity the accuracy of machine processing can be improved. To address this challenge, the method has presented a NL based automated approach that uses a UML class model as a context of the input English (constraints) and by using the available information in the UML class model (such as classes, methods, associations, etc) we can resolve attachment ambiguity and homonymy.
10	Answer Evaluation using Machine Learning	The system proposed here will first tokenize the answer in words and then evaluate answers using machine learning tools performing automatic evaluation of answer.Only one has to scan and give answer sheet to the system, the algorithm will extract the suitable features from the user provided answer to match it with the correct answer.Such system is required to provide eligible and fair marks.This will reduce the time and cost of institutions that they pay for manual evaluation.
11	Grading Descriptive Answer scripts using Deep Learning	It is a combination of NLP and machine learning. The learning is done by using the Recurrent Neural Network and LSTM cells. Deep Neural networks are able to capture the semantics of text in order to and the similarity between texts.The goal of the Deep Descriptive Answer Scoring model is to replace a system where marking and Evaluation depends on factors like human mindset ,time, thinking style etc.

12	Automated Essay Scoring	Automated Essay Scoring Systems- AES systems can be a great assistance to teachers in responding to a large number of essays and assign frequent writing assignments without worrying about scoring the first and subsequent drafts. There is a common methodology among four Automated Essay Scoring system- Project Essay Grader, Intelligent Essay Scorer, Erater and Criterion, IntelliMetric, and MY Access. They all need to be trained on a large number of essays in order to make the learning of the system strong. Next, all systems provide strong feedback with the best possible scoring thus making the task of teachers easier and flexible.
13	Automatic Assessment of Descriptive Answers for Online Examination using Semantic Analysis	The system is executed on java 3-tier architecture framework with INTEL 2.7 GHz i3 processor and 4 GB RAM with a supervised learning approach. The system is tested with around 100 samples of the question-answer dataset using the proposed NLP(Natural language processing) and ANN(Artificial Neural network) and a considerable accuracy is achieved.
14	Synonyms Paraphrasing using WordNet and Internet	Using a wordnet dictionary to produce synonyms of the word. Various types of paraphrasing-Text compression, Text canonization, Text simplification to fulfill the need to produce slightly deviated questions from the text. Another application is in producing synonyms of the correct answer so that different words with the same meaning as the correct answer should be evaluated correctly. Synonyms can be used as better improvements of certain words in the text that will make text compressed and canonized. For each word, we generate its relevant bucket of possible variants. The second application is keeping certain words secret. The area of cryptography where certain information needs not to be as exactly translated
15	Extracting Word Synonyms from Text using Neural Approaches	The paper proposed a neural network algorithm using backpropagation to find the semantic correctness in the piece of text. Mitokov shows that similarity does not always mean semantic relation in the text. The two-step neural architecture using Word Embeddings and SimLex-999 shows that the model achieved high accuracy. A score greater than 6.5 is considered to generate a high-synonyms similarity pair of words. Rich database WordNet fails to provide such results as it does not contain contemporary compound adjectives. The most common approach discussed here is the Continous Bag-of-Words model and Skip-Gram Model.
16	Improving Language Understanding by Generative Pre-Training	We are working on better understanding of the answer given by the student, so this research paper gives the way of achieving strong natural language understanding with a single task-agnostic model through generative pre-training and discriminative fine-tuning. Also the concept of transformers are very well explained in the research paper.
17	The Curious Case of Neural Text Degeneration	The paper guides us about the most common decoding methods for open-ended language generation. Also Nucleus Sampling as a solution that captures the region of confidence of language models effectively.
18	Encoding Word Order in Complex Embeddings	The primary objective of the paper is to make better nlp by encoding words sequentially or in order. This approach really develops a major role in how embedding also used to form a sequenced sentence. They extended word vectors to word functions with a variable i.e. position, to model the smooth shift among sequential word positions and therefore implicitly capture relative distances between words.
19	Mimic and Conquer: Heterogeneous Tree Structure Distillation for Syntactic NLP	The knowledge distillation method is described for making the NLP tasks easier by distilling syntactical features of text easily. The model outperformed the existing tree encoder models. LSTM model is used to predict the word based on its several features. And thus LSTM is trained to learn the semantic features of the text.

20	Learning to Ask: Neural Question Generation for Reading Comprehension	The model is trained with the dataset containing 100,000 questions about 536 articles. The dataset is created by employing Amazon Mechanical Turks crowd workers to frame questions. They were asked to make slightly variated questions for the better training of the model. In deep learning recurrent neural network LSTM, hidden unit size is set to 600 and number of layers of LSTMs to two. Stochastic gradient descent algorithm is used with learning rate at epoch 8, batch size as 64. The professional English speakers were asked to rate the model on a range of 1-5 and then human raters were asked to give a ranking of questions. The model achieved good ranking and high accuracy. The two main criterion used in the model were naturalness and semantic relation in the question.
21	Contextual LSTM (CLSTM) models for Large scale NLP tasks	The experiment in the paper clearly demonstrate, that the power of topics with world level features shows decent improvement in the LSTM. CLSTM shows improvement of 21 and 18 percent respectively on Wikipedia corp and Google news. The CLSTM improvement on the state of art of LSTM by 2-3 percent and the improved accuracy of about 20 percent over the old LSTM. these gains are quite significant and we get similar gains on the Google news as well. which shows the span of the algorithms. The improvement obtained by using the CLSTM model has major increase in the performance improvements in many NLP applications, ranging from sentence completion, question and answering, and paraphrase generation to different applications in dialog systems.
22	Improved Neural Relation Detection for Knowledge Base Question Answering	KB relation detection is an important step in KBQA and it is very different from general relation extraction methods. Paper has proposed a KB relation detection model,that performs matching between questions and KB relations. Model proposed in the model out performed the previous methods on KB relation detection and allows our KBQA system to achieve state-of-the-arts. Also it will help in investigating the integration of our model into end-to-end systems. For example, our model could be integrated with the decoder to provide better sequence prediction. We will also investigate datasets like graph question and complex questions to handle more characteristics of general question answering.
23	MATINF : A Jointly Labeled Large-Scale Dataset for Classification,Question Answering and Summarization	Since MATINF is a web-crawled dataset, it would be inevitable to be noisier than a dataset annotated by hired annotators though we have made every effort to clean the data. On the bright side, it can encourage more robust models and facilitate real-world applications. For future work, we would like to see more interesting work exploring new multi-task learning approaches.
24	What BERT Sees: Cross-Modal Transfer for Visual Question Generation	This research becomes useful whenever we deal with the data present in the form of image, BERT model shows the excellent performance, ensuring the benefit of pre-trained dataset models. Even training on the data and freezing the eightts of the model, it performed well. This shows the scope of BERT working as a model beyond normal text. Due to the flexibility of the model, with image only as input, the model is found to be improved further resulting in the better output. Data in the form of image caption as an input, produces the highest accuracy was obtained. This shows that BERT performs really well in textual data as well as visual data. It is a strong model to generate questions from data present in the form of image.

25	EasyTransfer – A Simple and Scalable Deep Transfer Learning Platform for NLP Applications	There are a variety type of TL toolkit algorithms available yet no kit is available for users to examine different types of state of the art tL algorithms. To overcome thus, Easy toolkit has been designed and is available to the open source platform. EasyTransfer has to integrate into a number of deep learning products in Alibaba. Easy transfer has made a major contribution by proposing a scalable and flexible platform to make it easy to develop deep TL algorithms for data science and large scale industrial applications.
26	The Question Answering System Using NLP and AI	Using the three approaches, the methodology deals with online platform AI and NLP approach to provide a flexible Question answer approach. This has improved our quality of the questions to a great extent and our work has addressed the shortcoming listed in the Research paper that students who simply memorize the chapter without actually understanding it if it's directly framed without paraphrasing the text.
27	Comparative study on Python web frameworks: Flask and Django	Every front end technology and framework is built with a motive to strengthen the technology. The article supports the web application growing need with elaboration on front-end technologies-HTML, CSS and web frameworks-flask, and Django. Based on the comparative study, it was found that Django is the best fit for large-scale projects whereas Flask is more suitable for small projects. Flask can be learned and set up quickly, but when it comes to managing and maintaining, it requires more work than the former framework.
28	Making It Work for Everyone: HTML5 and CSS Level 3 for Responsive, Accessible Design on your Library's Website	The paper provides the tools and techniques commonly used in web development. In our project, we have integrated our project as a web application using HTML and CSS listed in the Research paper. Brophy and Craven have highlighted the overview of tech stack in web application and its emerging need in the growing digital world.
29	RepoAI: A Novel Approach Towards Automated Project Reporting	A report is fundamentally a document that gives out information on a given topic in a methodical fashion. Writing the report is not as overwhelming as some may initially think, more so, for school students. A student's interaction with prevailing literature often requires perusing a substantial amount of text. In this paper, a novel method using web scraping and NLP have been proposed to summarize project reports automatically so that students can perform their tasks by consulting shorter summaries instead of entire literature.
30	Efficient Way Of Web Development Using Python And Flask	Every front end technology and framework is built with a motive to strengthen the technology. The article supports the web application growing need with elaboration on the most widely used open-source language-python and scalable and simple web frameworks-flask. Efficient deployment using python and flask has been easy with extensive documentation. Python can be used for making our web application simple, scalable, impressive with the help of an extensively used framework flask template engine.

REFERENCES

- [1] Tariq Ahmad, Lars H Lund, Pooja Rao, Rohit Ghosh, Prashant Warier, Benjamin Vaccaro, Ulf Dahlström, Christopher M O’connor, G Michael Felker, and Nihar R Desai. Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients. *Journal of the American Heart Association*, 7(8):e008081, 2018.
- [2] Fankar Armash Aslam, Hawa Nabeel Mohammed, Jummal Musab Mohd, Murade Aaraf Gulamgaus, and PS Lok. Efficient way of web development using python and flask. *International Journal of Advanced Research in Computer Science*, 6(2), 2015.
- [3] Imran Sarwar Bajwa, Mark Lee, and Behzad Bordbar. Resolving syntactic ambiguities in natural language specification of constraints. In *International Conference on Intelligent Text Processing and Computational Linguistics*, pages 178–187. Springer, 2012.
- [4] Stewart C Baker. Making it work for everyone: Html5 and css level 3 for responsive, accessible design on your library’s web site. *Journal of Library & Information Services in Distance Learning*, 8(3-4):118–136, 2014.
- [5] Phil Blunsom. Hidden markov models. *Lecture notes, August*, 15(18-19):48, 2004.
- [6] Igor A Bolshakov and Alexander Gelbukh. Synonymous paraphrasing using wordnet and internet. In *International Conference on Application of Natural Language to Information Systems*, pages 312–323. Springer, 2004.
- [7] Hans Christian, Mikhael Pramodana Agus, and Derwin Suhartono. Single document automatic text summarization using term frequency-inverse document frequency (tf-idf). *ComTech: Computer, Mathematics and Engineering Applications*, 7(4):285–294, 2016.
- [8] Shweta Devare Phadke, Sabina Kazi Hakim, Meera Deshmukh, and Pranati Tilak. Comparison of physiotherapy and management students’ perception towards online examination introduced during covid-19 pandemic. 07:5403–5409, 08 2020.
- [9] Semire Dikli. Automated essay scoring. *Online Submission*, 7(1):49–62, 2006.
- [10] Xinya Du, Junru Shao, and Claire Cardie. Learning to ask: Neural question generation for reading comprehension. *arXiv preprint arXiv:1705.00106*, 2017.
- [11] Hao Fei, Yafeng Ren, and Donghong Ji. Mimic and conquer: Heterogeneous tree structure distillation for syntactic nlp. *arXiv preprint arXiv:2009.07411*, 2020.
- [12] Devendra Ghimire. Comparative study on python web frameworks: Flask and django. 2020.
- [13] Shalini Ghosh, Oriol Vinyals, Brian Strope, Scott Roy, Tom Dean, and Larry Heck. Contextual lstm (clstm) models for large scale nlp tasks. *arXiv preprint arXiv:1602.06291*, 2016.
- [14] Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. The curious case of neural text degeneration. *arXiv preprint arXiv:1904.09751*, 2019.
- [15] Nikiforos Karamanis, Ruslan Mitkov, et al. Generating multiple-choice test items from medical text: A pilot study. In *Proceedings of the fourth international natural language generation conference*, pages 111–113, 2006.
- [16] Andrey Kurtasov. Automated generation of assessment test items from text: Some quality aspects. In *AIST (Supplement)*, pages 91–95, 2014.
- [17] David Lindberg, Fred Popowich, John Nesbit, and Phil Winne. Generating natural language questions to support learning on-line. In *Proceedings of the 14th European Workshop on Natural Language Generation*, pages 105–114, 2013.
- [18] P Manikiran, Abhay Subramanian, and Kshithij R Kikkeri. Repoai: A novel approach towards automated project reporting.
- [19] Nora Mohammed. Extracting word synonyms from text using neural approaches. *Int. Arab J. Inf. Technol.*, 17(1):45–51, 2020.
- [20] Toshavi Patil. Automatic assessment of descriptive answers for online examination using semantic analysis. *Journal of the Gujarat Research Society*, 21(5):413–419, 2019.
- [21] Minghui Qiu, Peng Li, Hanjie Pan, Chengyu Wang, Cen Chen, Yaliang Li, Dehong Gao, Jun Huang, Yong Li, Jun Yang, et al. Easytransfer—a simple and scalable deep transfer learning platform for nlp applications. *arXiv preprint arXiv:2011.09463*, 2020.
- [22] Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. Improving language understanding by generative pre-training, 2018.
- [23] Bipul Roy and Bipul Syam Purkayastha. A study on different part of speech (pos) tagging approaches in assamese language. *International Journal of Advanced Research in Computer and Communication Engineering*, 5(3), 2016.
- [24] Thomas Scialom, Patrick Bordes, Paul-Alexis Dray, Jacopo Staiano, and Patrick Gallinari. What bert sees: Cross-modal transfer for visual question generation. In *Proceedings of the 13th International Conference on Natural Language Generation*, pages 327–337, 2020.
- [25] PJ Sijimol and Surekha Mariam Varghese. Handwritten short answer evaluation system (hsaes). *IJSRST*, 4, 2018.
- [26] Miguel Alexis Solano-Jiménez, Jose Julio Tobar-Cifuentes, Luz Marina Sierra-Martínez, and Carlos Alberto Cobos-Lozada. Adaptation, comparison, and improvement of metaheuristic algorithms to the part-of-speech tagging problem. *Revista Facultad de Ingeniería*, 29(54):e11762–e11762, 2020.
- [27] Benyou Wang, Donghao Zhao, Christina Lioma, Qiuchi Li, Peng Zhang, and Jakob Grue Simonsen. Encoding word order in complex embeddings. *arXiv preprint arXiv:1912.12333*, 2019.
- [28] Joy T Wu, Franck Dernoncourt, Sebastian Gehrmann, Patrick D Tyler, Edward T Moseley, Eric T Carlson, David W Grant, Yeran Li, Jonathan Welt, and Leo Anthony Celi. Behind the scenes: A medical natural language processing project. *International journal of medical informatics*, 112:68–73, 2018.
- [29] Canwen Xu, Jiaxin Pei, Hongtao Wu, Yiyu Liu, and Chenliang Li. Matinf: A jointly labeled large-scale dataset for classification, question answering and summarization. *arXiv preprint arXiv:2004.12302*, 2020.
- [30] Mo Yu, Wenpeng Yin, Kazi Saidul Hasan, Cicero dos Santos, Bing Xiang, and Bowen Zhou. Improved neural relation detection for knowledge base question answering. *arXiv preprint arXiv:1704.06194*, 2017.