WYCH Debater: A Modular System For Argument Mining, Speech Formation And Debate Rebuttals Based On Artificial Neural Networks

Rishabh Yadav[1], Anujay Ghosh[2], Vidit Gholam[3], Shweta Sharma[4], Deepali Maste[5]

[1]*Student, Information Technology, Atharva College of Engineering/Mumbai University, India.*

[2,3]*Student, Computer Engineering, Atharva College of Engineering/Mumbai University, India*

[4]*Asst. Professor, Computer Engineering, Atharva College of Engineering/Mumbai University, India.*

[5]*Asst. Professor, Information Technology, Atharva College of Engineering/Mumbai University, India.*

Abstract

AI Debating system the first perceptive system which is able to debate humans on complex topics. It depends on three pioneering capabilities. The primary one is speech writing based on provided information and delivery of the speech, or the ability to automatically synthesize an entire speech, an article that is reminiscent of an opinion, and its ability to deliver it persuasively. The second is comprehension of what it listens, which is the ability to understand a long spontaneous speech made by the opponent human in order to generate a meaningful rebuttal. The third is the system's ability to represent and circumvent human dilemmas and form principled arguments made by humans in various debates to determine what constitutes an effective negation, and then follow a statistical approach to determine when an automatically generated negation can plausibly be used on a unique knowledge graph. By effectively integrating these core capabilities, it can conduct an articulate debate with human debaters.

Keywords: (*LSTM) Long Short-Term Memory, (CDC) Claim Detection Corpse, (RNN) Recurrent Neural Network, (SVM) Support Vector Machine, (DNN) Deep Neural Networks.*

1. Introduction

WYCH Debating system is an AI system that can argue with humans on a varied variety of topics and to do this efficiently, the system has to collect pertinent opinions and facts, align them into organized arguments, and then use decisive language in a coherent and convincing manner. AI Debating system relies on three pioneering capabilities: the first is data driven speech writing and delivery, listening comprehension, and the modeling of human predicament. AI Debating system analyzes large texts, forms a well structured speech on a subject provided, and delivers it with coherence and persuasion, and rebuts its opponent. Eventually, WYCH Debating system focuses on helping individuals by furnishing evaluating, evidence based arguments and reducing the prejudice that comes with emotion, bias, or ambiguity. If it disagrees, it explains its position with respect and refrains from any and all personal attacks. The goal is to help individuals build convincing arguments and make well informed decisions. The growth of one sided and tampered accounts is challenging the world. Novel enhancements in language and cognitive analysis in AI can help recognize and disprove distorted facts to provide multifaceted and perceptive viewpoints to both pro and con. The world is full with information, misinformation, and superficial thinking.

2. Problem Definition

The growth of one sided and tampered accounts is challenging the world and our platforms. Too often, we tend to speak past each other. We need a better approach. Novel enhancements in language and perceptive analysis in AI can help recognize and disprove distorted data to provide multifaceted and perceptive viewpoints to both pro and con. The world is full of information,

ISSN: 2233-7857 IJFGCN

misinformation, and superficial thinking. Because of lack of evidence, many times the innocent get punished and the existing judicial system is so slow that a decision may take years or decades many times.

Hence the WYCH Debating system can prove as a beneficial resolution for such problems, as it forms a cogent argument. Given a subject, the Debating system scours its large body of data trying to find the foremost relevant points and proof to support or contest the subject. It then picks the foremost compelling, numerous and well supported arguments and arranges them to construct a whole persuasive narrative. Debating system is aware of if a claim is for or against the subject it's given, this is one of the many things that make WYCH Debating system a unique system of its own.

3. Motivation

WYCH Debater is a computer science that focuses on increasing human thinking through impartial debate, discovering new boundaries of computer science by training systems to generate beneficial and well informed perspectives. The objective is to build a system that aids humans make confirmation based decisions when the solutions aren't just black-or-white.

We plan on teaching systems how to debate because culturally, the aim of debate lie not in disputes and competition, but in elective governance and discussion. Debate improves decision making and helps assisting people contemplate the good and bad of new ideas and philosophies. It is at the far end of educated society. We debate not only to persuade others of our own personal views, but also to comprehend and learn from each other's point of view. In the coming future, we hope that systems will be able to support humans with a number of important decisions we make every day. It is very unique and different from searching keywords because a keyword search will bring together a collection of relevant documents only.

4. Literature Survey

In this chapter we've specified and analyzed major components involved in an AI Debater system. We have initially referred a CDC model for argument mining and various architectures for word emphasis predictions. also described Natural language processing algorithms, deep neural network and weak supervision models.

Basic ideology for argument mining , Shachar Mirkin et al.[1] presented a systems listening comprehension task within the scope of reasoning and a corresponding information set in English. around two thousand spontaneous speeches arguing for or against fifty controversial subjects where being recorded by them. drawing up a question and focusing toward confirming or rejecting the occurrence of potential arguments in the speech. they assembled Labels by observing and listening to the speech and making a note of which argument were mentioned by the speaker. baseline methods where applied addressing the task, to be used as a benchmark for future work over this information set. All data utilized in this work is freely accessible for research.

Argument mining using CDC models, Yonatan Bilu et al.[2] While discussing a concrete complex subject, they found out most people will find it difficult to swiftly raise a varied variety of undoubted claims that should set the foundation of their reasoning. Hence, they defined the difficult task of automatic claim identification in a given context and discuss its associated unique changes. They further charted an introductory explanation to this task, and evaluate its performance over annotated real world data, collected specifically for that purpose over hundreds of Wikipedia articles. They reported promising results of a supervised learning approach, which is based on a cascade of classifiers designed to properly handle the skewed data which is inherent to the defined task. the introduced task's viability where demonstrated by their results.

This supervised learning approach relies on labeled data that were collected as described below. in (Aharoni et al., 2014) a detailed description of the labeling process is given . due to concise statement that directly supports or contests the given Topic. the labelers were approached to mark a book section by and by as a Claim Detection Corpus just on the off chance that it agrees to all the accompanying five criteria:

1. Quality - Strong substance that straightforwardly underpins/challenges the Topic.

2. Simplification - General substance that manages a moderately expansive thought.
3. Expressing - The parts which were marked should make a linguistically right and semantically intelligible articulation.
4. Keeping content soul - Keeps the soul of the first content.
5. Theme solidarity - Deals with one subject, or at most two related points.

These given guidelines further included concrete examples, which were taken from Wikipedia articles, to clarify these criteria. When uncertain, the labels were logically asked to form a judgment call. The labels work was judiciously monitored, and they were given detailed feedback as and when required.

Approach for sentiment composition, Noam slonim et al.[3] recommended a novel strategy for taking in opinion organization from an outsizes, unlabeled corpus, which just includes a word-level conclusion dictionary for supervision. They precipitously produce enormous opinion vocabularies of bigrams and unigrams; from which they make a lot of dictionaries for a collection of assessment organization forms. Through manual comment the adequacy of their strategy was set up, just as conclusion order explores different avenues regarding both expression level and sentence level benchmarks.

This strategy for learning supposition structure vocabularies involves the accompanying advances:

1. Train an n-gram slant classifier on a prearranged feeling dictionary for unigrams.
2. Utilize the slant classifier to consequently create huge assumption vocabularies of bigrams and unigrams.

Charles Jochim et al.[4] demonstrated that both precision and inclusion can be altogether improved through programmed extension of the underlying dictionary.

They prepared a direct SVM classifier, which incorporates the benchmark framework (with the extended vocabulary) as a component, together with a lot of logical highlights, portrayed underneath. Like the gauge framework, the classifier expects to foresee the position towards the point target "xt", and the outcome is duplicated by the offered st to acquire Stance(c, t).2

Natural language processing, Martin Gleize et al. [6] The technique for achieving excellent labeled information for natural language understanding assignments is frequently moderate, blunder inclined, unpredictable and exorbitant. This issue ends up being progressively notorious since these frameworks require a ton of stamped data to make great results with the colossal use of neural frameworks. In this manner, they proposed an approach to mix high caliber yet rare marked information with loud yet bottomless powerless labeled information during the preparation of neural systems.

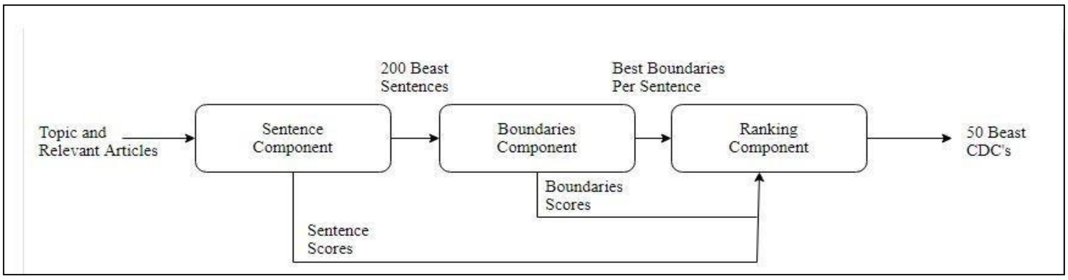
GrASP Algorithm, Eyal Shnarch el al.[7] presented the GrASP algorithm for mechanized creation of patterns that characterize subtle semantic phenomena. To the end that the GrASP augments each term of input text with multiple layers of semantic information. These diverse features of the text terms are methodically joined to expose rich patterns. They reported as expected exceedingly reliable experimental outcomes in numerous puzzling text analysis tasks within the arena of Argumentation Mining.

4. Technical Approach

The CDCD strategy we utilized is planned as a cascade, or channel, of three modules (portrayed in Figure 1), which acknowledges as input a subject alongside relevant articles and should yield the CDCs encased in that. The inspiration driving the channel is to a tiny bit at a time base on humbler and more diminutive CDC-containing content sections, while filtering through pointless substance. In this manner, the pipe parts the significant level CDCD issue into decreased and progressively unmistakable issues – given an article, distinguish sentences that involve CDCs; given a sentence, recognize the exact CDC limits; given a lot of CDC up-and-comers, rank them with the goal that genuine applicants are most noteworthy. The demonstrated numbers are the ones utilized in our examinations, and when all is said in done ought to be resolved dependent on the information and use case. To value the requirement for this channel, let us initially think about the size of this recognition issue. In named data, per Topic we have an ordinary of 10 significant Wikipedia articles that contain in any occasion 1 CDC. Each article contains a normal of 155 sentences, each sentence ranges on normal 23 words, for example 200 sub-sentences, every one of which may mean an applicant CDC.

Figure.1 Argument Mining approach via CDC

Prediction of word emphasis , Ron Hoory et al.[5] Presented a technique outperforms machine learning techniques based on hand crafted features in terms of objective metrics such as precision



and recall. By means of a listening test, we further establish that the impact of the predicted emphasized words to the expressiveness of the generated speech is subjectively perceivable. word accentuation forecast is a huge bit of expressive discourse age in current Text-To-Speech (TTS) frameworks. We present a technique for calculating emphasized words for expressive TTS, based on a Deep Neural Network (DNN)

word emphasis the index terms, synthesis of speech, and also expressive text to speech, prosody, deep learning the proposed architecture (Figure 2) receives a batch of sentences as input and processes each sentence as follows.

The Word Embedding Layer extracts a feature vector for each word using a word embedding matrix. We use the Google's pre-trained w2v [1] that represents the semantic meaning of the words. The pre-trained matrix allows us to benefit from training on a very large unlabeled data set. The Fully Connected (FC) Layer translates the original word embedding's into new representations to better fit the task at hand. It applies a linear transformation to the word embedding, followed by tanh as a non-linear activation function

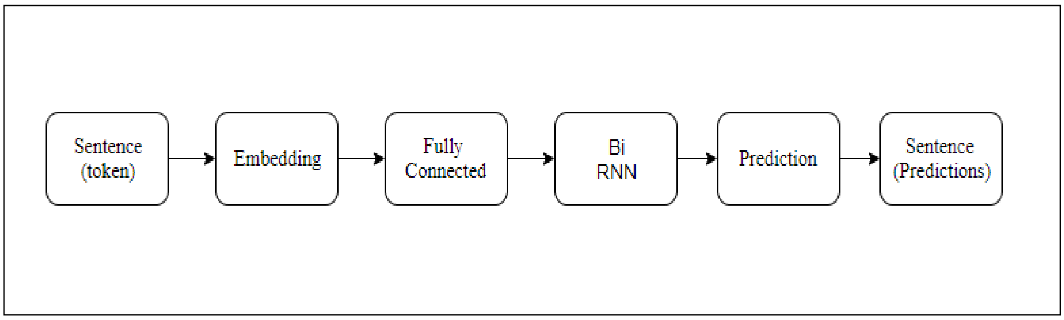


Figure.2 Model architecture of word emphasis prediction

The Bidirectional RNN Layer captures the context of each word when predicting whether it should be emphasized. Clearly, emphasizing a word depends on its context [2]. We use LSTM [3] to capture the consecutive elements in a sequence (in our case, words in a sentence). The learned demonstration of each term is dependent on the elements that come before it. To capture subsequent words, we use bidirectional LSTM (BiLSTM). As a result, the outcome of this layer captures the implication of each word together with its pertinent context. The Prediction Layer is a fully-connected layer that is used to translate the representation computed in previous layers into a probability score that represents the probability of the word being emphasized. This is done by computing the sigmoid on the inner product between a learned weight vector $\beta 1$ and the output of the previous layer x plus a bias term. namely, sigmoid ($\beta 1x + \beta 0$).

The network can be trained on an annotated voice corpus with binary word emphasis labels attached to each word (i.e., the emphasized words are labeled with 1, and the rest with 0). The

ISSN: 2233-7857 IJFGCN

labeled data that we used is described in Section 5 below. The loss function is the weighted cross entropy between the predictions and the actual labels.

X loss = [label x • (−log(prediction x)) • pw + x∈X (1 – label x) • (−log(1 − prediction x))] where X represents all words in all training sentences and the hyper parameter pw is used as a weight for compensating the positive (i.e., emphasized) words, due to their unbalanced ratio among all words . Another method for handling unbalanced data is to apply over/under sampling in the training set and fix the prediction bias [4] as we did in Section 6 for the Logistic Regression classifier. However, as our DNN model relies on context, it is not practical to over/under sample words within a sentence. Once the model is trained, it can be used for predicting emphasized words in a new sentence as follows. A sentence is input into the network, which outputs a prediction value for each word, as defined above. Every word with prediction value ≥ 0.5 is then defined as emphasized words.

5. Proposed Argumentation Mining Algorithm and GrASP Algorithm

5.1 Proposed Argumentation Mining Algorithm

Here we have implemented three types of Neural Networks that can be used to solve the Claim Detection. We employ our models on the IBM Datasets and for each of them we consider the pertained word embedding’s built with Glove model. For the Tree-LSTM model we follow the code of the Stanford Tree-Structured Long Short-Term Memory Networks.

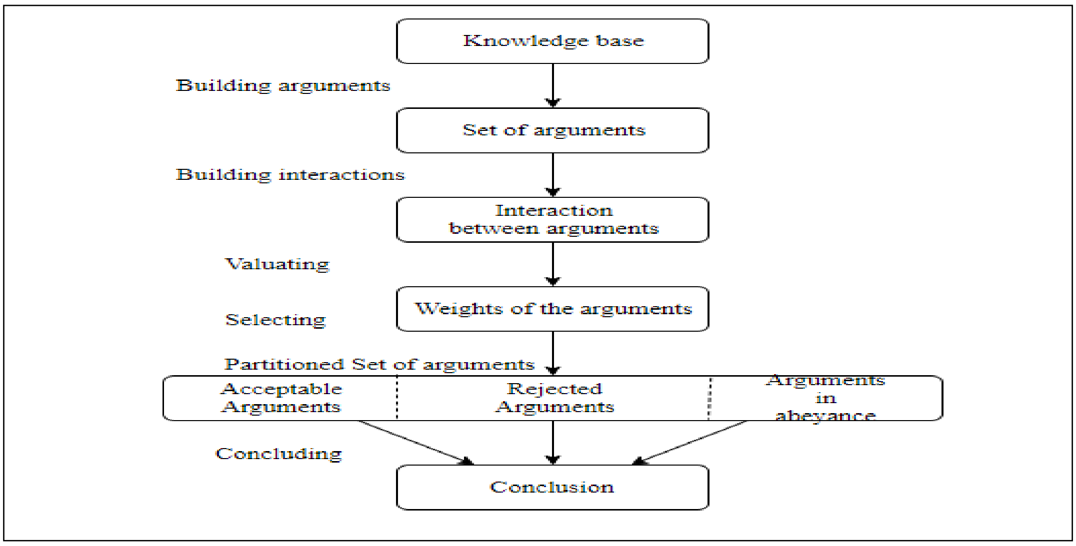


Figure.3 Mining Argument from Debating System

LSTM implementation of the LSTM: the model is defined in the lstm.py, scores.py is used to evaluate the model. The considered topics are listed in considered_topic.txt.

RNN implementation of the Recurrent Neural Network: the model is defined in the rnn.py, scores_and_charts.py is used to evaluate the model.

Tree-LSTM comprises of the modifications made to the Tree-Structured Long Short-Term Memory Networks to fit their implementation to our assignment.

5.2 GrASP Algorithm

The calculation portrayed in Algorithm 1. Its information set is a lot of positive and negative models for the objective marvel. The yield is a positioned rundown of examples, meaning to demonstrate the nearness or nonappearance of this marvel. In the accompanying, an example is viewed as coordinated in a book off the entirety of its components are found in it, in the predefined request, conceivably with holes between them, inside a window of size w.

Algorithm 1

1. Input: positive/negative content models, k1, k2, maxLen
2. Output: a positioned rundown of examples
3. (pos, neg) ← augment(positives, negatives)
4. 2 qualities ← extractAttributes(pos, neg)
5. 3 letters in order ← chooseT opK(attributes, k1)
6. 4 examples ← letters in order
7. last ← designs
8. for length ← 2 to maxLen do
9. curr ← θ
10. for p last do
11. for a letters in order do
12. curr ← curr {growRight(p, a)}
13. curr ← curr {growInside(p, a)}
14. last ← curr
15. patterns ←chooseT opK(patterns current, k2
    a. return designs

6. System design and Implementation Outputs

This section includes system design of the implemented AI Debater the above figure explains the flow of the system. the Moderator fives a motion topic to AI Debater System. This motion is then decided as to support or contest and accordingly it scans the entire corpus of dataset to come up with relevant information and then from this information it determines the effective argument and assembles them to form a persuasive narrative of the debate topic. If the human debater speaks then it listens it carefully otherwise it just delivers it's prepared speech with the help of text to speech IBM Watson API.

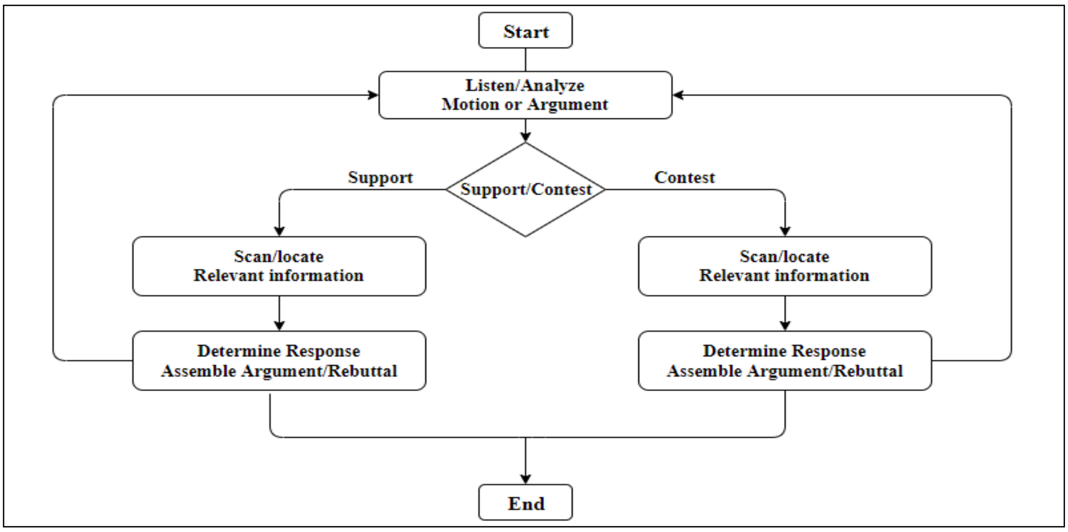


Figure.4 Activity Diagram of AI Debater System

6.1. Mining Argument

Claims and proof are the fundamental segments of a contention; recognizing and utilizing them effectively are basic to surrounding a contention in a discussion. Creating AI methods to scan enormous writings for cases and evidence and use them to deliver contentions relevant to a disputable issue.

1. Detecting claims in applicable records: We were the first to define and implement the challenging task of detecting topic-related claims within unstructured text. Our method automatically pinpoints relevant claims within a set of documents that can be used to support or contest a given controversial topic. We accomplish this using a cascade of AI algorithms exploiting various linguistic features.

2. Detecting evidence in relevant documents: We were also the first to define relevant-evidence detection as a task and to develop methods that accomplish it. Given a controversial topic and a claim, our method finds text segments in unstructured text from relevant documents that can serve as evidence supporting the claim. Our approach classifies three common evidence types: study, expert, and anecdotal.

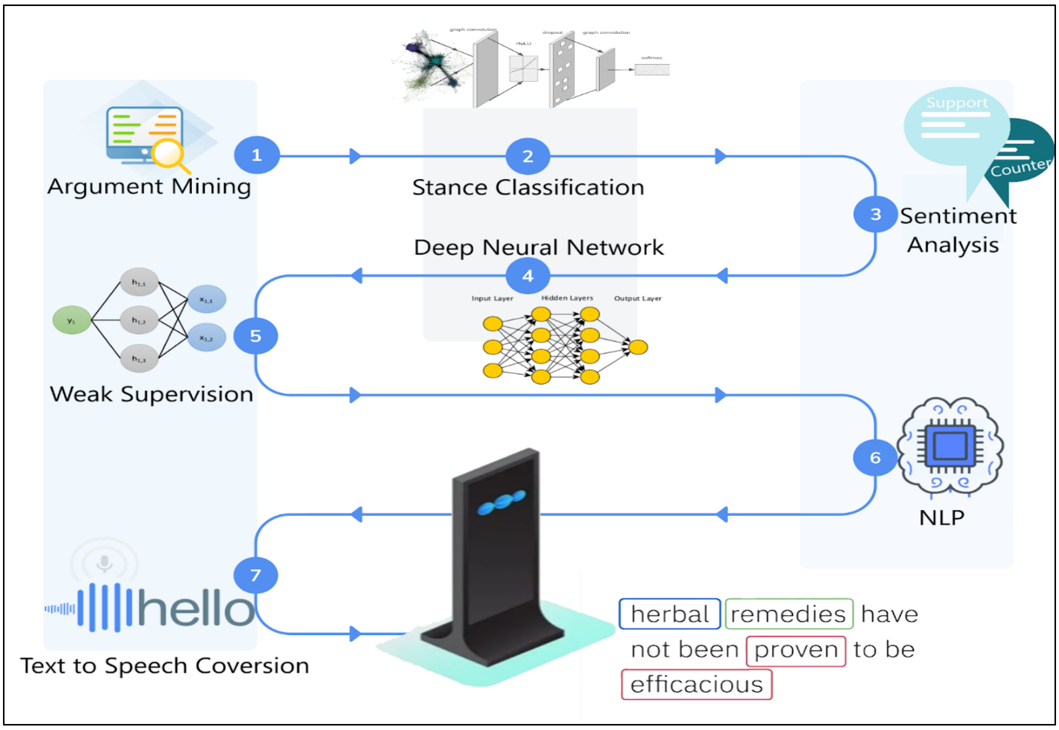


Figure.5 working flow of AI Debater System

3. Negating claims: We built up a way to deal with consequently produce a significant refutation to a given case about a dubious theme. The calculation has two parts: a standard based way to deal with figure out what comprises a powerful invalidation, at that point a measurable way to deal with decide when a consequently produced nullification can conceivably be utilized.

4. Synthesizing novel claims: It is one thing to detect claims included within relevant documents, and quite another to generate claims "de novo." We developed a method to do this by "recycling" existing arguments. Fundamental text elements extracted from a database of argumentative text are combined to construct claims that are grammatically correct, meaningful, and relevant.

5. Detecting claims throughout a corpus: We were the first to expand claim detection methods beyond preselected relevant documents by developing a framework for unsupervised, corpus-wide claim detection. Our system can pinpoint claims in a huge corpus relying solely on linguistic cues that are inherent to natural language, eliminating the need for costly and time-consuming human annotation.

6. Improving corpus-wide claim detection: We are exploring how to use corpus-wide claim detection to develop an argumentative content search engine. We have obtained high-quality results using DNNs trained via weak supervision with automatically labeled data and no human intervention.

7. Assessing argumentation quality: With academic collaborators, we are researching ways to assess the quality of machine-generated arguments. We used existing theories and

ISSN: 2233-7857 IJFGCN

approaches to derive a systematic taxonomy for computational argumentation quality assessment. We also showed that quality assessments based on theory versus practice generally agree and support one another.

```
Gambling is one of the most unproductive human behaviors
Gambling is addictive because of the sunken sum effect and so people are not rational when thay decide to gamble and we have the right to choose for the
Casinos promote other harmful behaviors like over-drinking and recklessness that can cause car accidents and other problems.
Gambling is a gateway habit to all sorts of degenerative behaviors. By banning it, governments are taking the moral high ground and making a stand.
We should ban gambling because according to most moral and ethical frameworks it is wrong to take advantage of people like casinos do.
Similar arguments
It provides a false hope that most will never get to experience; therefore, it disproportionately effects the poorest members of society, who hope to ge
Gambling can be highly addictive and can lead to devastating financial effects on an individual and a family for years
It is a metal disease that not only causes individual harm but also complete kills an entire family
Gambling is too addictive, it can spoil human's rational judgement capability.
Gambling affects asymmetrically people from a disadvantaged socioeconomic background and create less equality of opportunities for them
it is a victim-less crime. gambling is amoral rather than immoral, and poses no clear threat to public order, and since it is a victim-less crime, and a
gambling can contribute to improving and increasing tourism.
Taxing gambling is a regressive tax (this means that the poor pay a greater proportion of their income in tax than the rich).
We should ban gambling because it is a tax on those who are less educated and can least afford to pay their gambling debts.
Similar arguments
Gambling is harmful to psychological and physical health. People who live with this addiction may experience depression, migraine, distress, intestinal
Gambling is an excellent source of tax revenue. Better to control and tax it than to make it a crime.
Casinos bring low wage jobs, not a good community investment.
Similar arguments
Gambling is addictive, and casino games are in favor of the house, therefore casino gambling often costs people more than they have budgeted to spend on
Gambling serves the community at large by generating tax revenue for the economy, providing a positive outlet for mental health, and jobs. It is an impo
Because people should be able to decide for themselves if the dangers are worth it, adults can make their own decisions about how to use their money.
it can bring in a lot of tax revenue for cities.
Gambling is capable of driving people into debts (by making them lose several times straight) and so damaging their buying power, halting the economy.
A single casino, as with any other individual firm, will have limited impact on an entire area's economy.
Gambling leads to millions of dollars in loss that could have been spent on other good and stimulated the economy.
Gambling is addictive and proven to provide more financial harm than good to individuals.
Similar arguments
```

Figure.6 Output of argument mining

8. Relating arguments across texts: to exploit corpus-wide argumentation mining, a framework needs to consolidate contention units from various writings. We structured a joined induction strategy for this assignment by displaying contention connection arrangement and position characterization as one. As far as anyone is concerned, this is the first-run through joint induction has been utilized right now. Here in figure 6 shows the gathering of contentions getting created by the AI Debater LSTM Model.

6.2. Position Classification and Sentiment Analysis:

A programmed discussing framework must have the option to perceive whether a contention supports or difficulties a given theme. This is genuinely simple for people however trying for machines, as it needs extraordinary affectability to the rich complexities and subtleties of regular language. We have gained significant ground right now of research.

1. Identifying expert opinion stance: Expert opinion is important evidence in constructing arguments, but its stance often hard to be determined from the text itself. We developed an innovative approach to this problem. By mining knowledge from Wikipedia with minimal human supervision, we developed a resource of over 100,000 experts and their stance toward over 100 controversial topics

2. Determining claim stance: We designed a technique to determine whether a given claim supports or challenges a new debatable topic. Our model breaks down the multifarious cognitive procedure of determining stance into a sequence of simpler sub-tasks. We recognized effective AI solutions to these sub-tasks, that can join to predict claim stance with high precision.

3. Improving claim stance classification: To improve claim stance classification, we developed a classifier that predicts the sentiment of a given word based on its context. This overcomes the limitations of manually composed sentiment lexicons. We also identified contextual features that can improve sentiment classification and enable classification of claims with no explicit sentiment.

4. Classifying sentiment of phrases: We designed a novel method for predicting the sentiment of a phrase based on its constituents. Using only the sentiment of individual words, our algorithm correctly handles complex phenomena such as sentiment reversal and mixed sentiment.

ISSN: 2233-7857 IJFGCN

5. Classifying sentiment of idioms: Claims and evidence often include idiomatic expressions, and a debating system must be able to analyses them to properly classify their stance. Because the sentiment of idiomatic expressions often cannot be deduced from their constituent words, we developed a sentiment lexicon of 5,000 common idiomatic expressions to improve sentiment analysis.

6.3. Weak Supervision along with Deep Neural Nets (DNNs):

DNNs hold enormous prospective for refining automatic understanding of language, but training them is infamously known to require a lot of high quality, manually labeled data. We developed tools and methods to train DNNs using weak supervision, alleviating that bottleneck. We also used DNNs in developing AI Debater's speaking and listening skills.

Scoring arguments: A debating system needs to score claims and evidence with respect to the topic of debate. We evaluated 19 different DNN-based methods of scoring arguments to help identify the best deep learning architecture for this task.

1. Understanding Automatic Speech Recognition (ASR) output: A debating system needs to understand arguments made by its opponent, which it receives as ASR transcripts. To do this, it must properly parse the ASR output into sentences by adding punctuation. We exploited DNNs to achieve this task.

2. Predicting phrase breaks: Phrase breaks are essential to delivering long sentences in continuous speech. We developed a novel DNN model for predicting where a phrase break or pause is needed and a new training process using phonetically aligned speech data and a weakly labeled large text corpus. This makes AI Debater's speech intelligible, natural, and expressive.

3. Improving speech patterns: We developed DNN-based models to enable controllable word-level emphasis and sentence-level emphasis in expressive TTS systems. Both models preserve quality and naturalness of the baseline TTS output while significantly improving the perceived emphasis.

4. Improving speech patterns: We built an expressive TTS system, based on DNNs, with one module that predicts which words to emphasize in a text and another that generates speech patterns based on the predictions. The prediction module outperforms methods with hand-crafted features, and the overall system is perceived as more expressive via crowd-sourced listening tests.

5. Identifying similar sentences: To train a DNN to predict thematic similarity between sentences, we automatically created a weakly labeled dataset of sentence triplets (a pivot sentence from a Wikipedia, another sentence from the same section of the article, and a third sentence from a different section of the article). Our model, trained over these data, outperformed state-of-the-art methods.

6. Improving argument mining: We developed a method to improve the performance of DNNs in argument mining by blending a small amount of high-quality, manually labeled data with a large amount of lower-quality, automatically labeled (weakly supervised) data.

7. Searching for claims throughout a corpus: searching for sentences containing claims in a large text corpus is a key component in developing an argumentative content search engine. We used DNNs trained via weak supervision (i.e., with automatically labeled data) to obtain high-quality results with no human intervention.

8. Determining concept abstractness: We used a DNN with weak supervision to determine the level of abstractness embodied within a given concept. Understanding whether the topic of the debate is abstract, as in ‘freedom of speech’, or concrete as in ‘zoo’, can guide the Debater system in developing more relevant arguments.

6.4. System For Text To Speech :

In contrast to an individual aide or guide, a discussing framework needs to talk constantly and convincingly for a couple of moments on a theme not known ahead of time, while keeping the crowd locked in. We grew new TTS calculations and systems to give WYCH Debater a solid, familiar, and persuading voice.

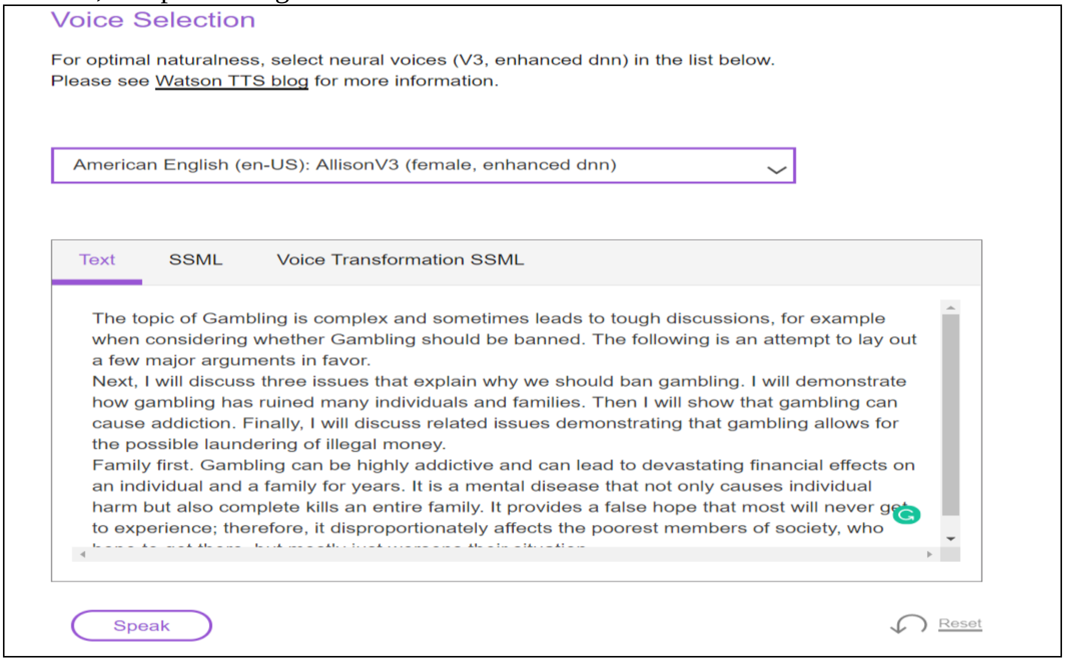


Figure.7 Output of text to speech

1. Predicting phrase breaks: Phrase breaks are essential to delivering long sentences in continuous speech. We developed a novel DNN model for predicting where a phrase break or pause is needed and a new training process using phonetically aligned speech data and a weakly labeled large text corpus. This makes Project Debater’s speech intelligible, natural, and expressive.

2. Improving speech patterns: We developed DNN-based models to enable controllable: word-level emphasis and sentence-level emphasis in expressive TTS systems. Both models preserve quality and naturalness of the baseline TTS output while significantly improving the perceived emphasis.

3. Improving speech patterns: We built an expressive TTS system, based on DNNs, with one module that predicts which words to emphasize in a text and another that generates speech patterns based on the predictions. The prediction module outperforms methods with hand-crafted features, and the overall system is perceived as more expressive via crowd-sourced listening tests.

7. Experimental Observations

We have created a web application to show the debate proceedings and formation of Pro and Con of the Debate topic provided. Figure.8 shows the home page of the web application. This contains both the pro and con of the topic in a visible representational way of which one is

International Journal of Future Generation Communication and Networking
Vol. 13, No. 1s, (2020), pp. 91- 106

dominating the other. It also consists of the no of arguments on which this classification is working on.

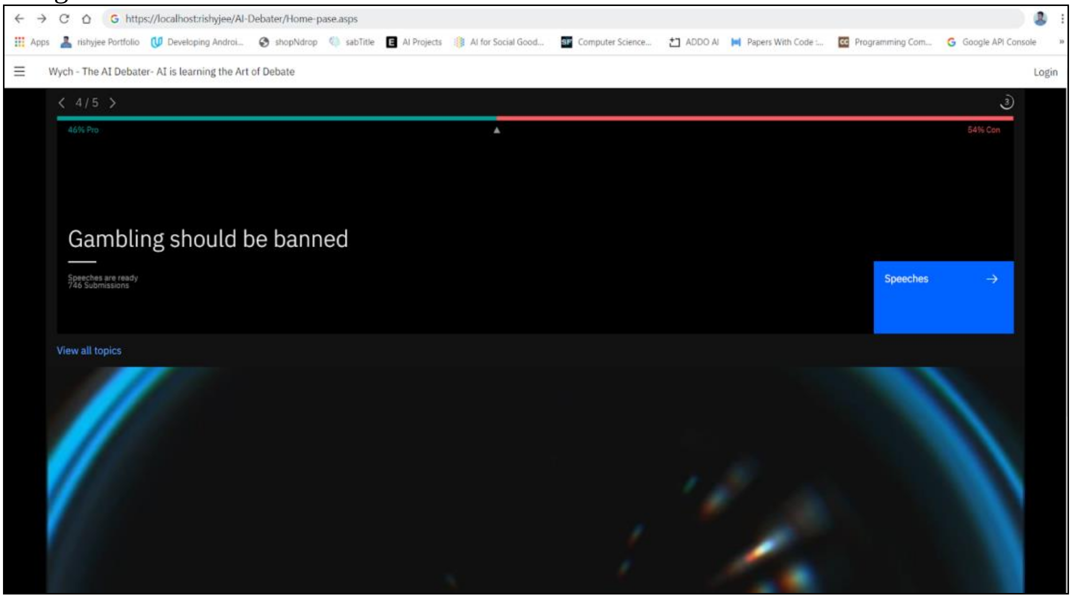


Figure.8 Home Page of AI Debater

Figure.9 shows the pro of the given topic. This explains only the arguments which support the given motion topic. It generally splits the entire dataset into 3 or 4 sections in which the first section is related to providing the chronological sequence of the debate and then in the second section is actually explains it's argument in a detailed and persuasive manner to generate the narrative and it has group of facts and figures to explain and support its' argument. This section is very important from the point of view of the opponent speaker. Now in the last section of the debate it gives concluding remark and leaves human speaker in a valid question to think.

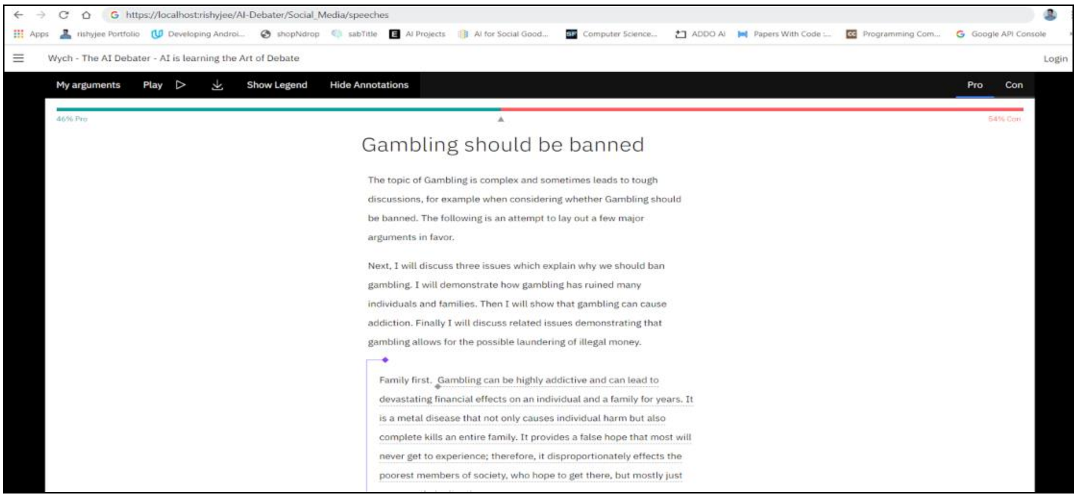


Figure.9 Gambling should be banned – Pro

Figure.10 shows the con of the given topic on whether Gambling should be banned or not. in this section it contesting the given topic. It initially collects those arguments which are very effective to contest the given motion and then it frames its' argument in 3 major section. The first section will be opening remark where it explains its' order of presentation of the content. In the second section it emphasizes on the most effective arguments and their explanation with facts and figures. in the final phase gives its' concluding remark with effective question or remark which leaves human debater to think and come up with counter arguments.

ISSN: 2233-7857 IJFGCN

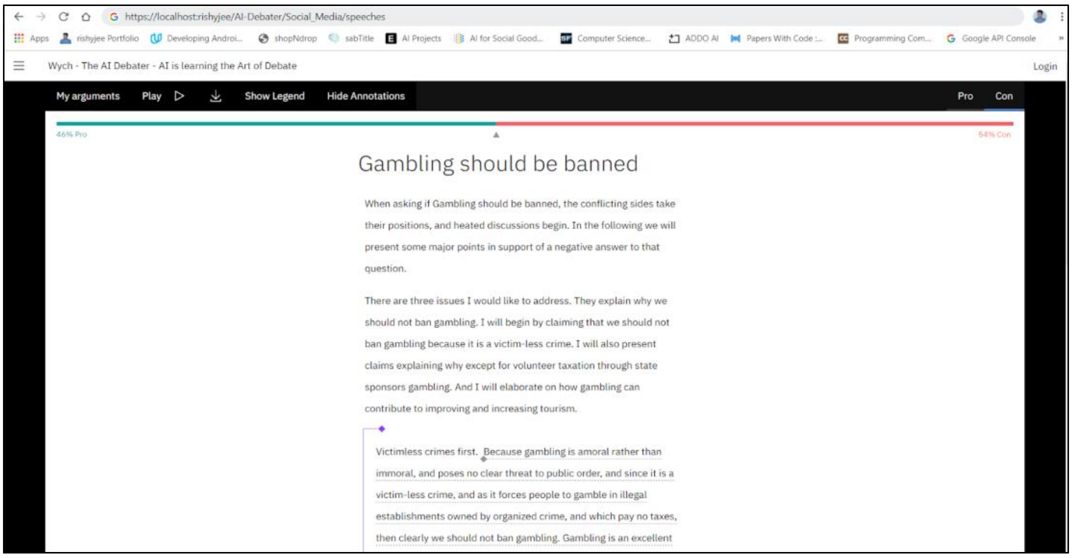


Figure.10 Gambling should be banned – Con

7.1. Tables

This section will describe the methodology and tools used for implementing the generative model. It will also describe the experiment conducted to determine which Luong score function to use for generating responses.

Table 1. Comparison between text generations based on different Luong score functions

Sentence	dot	general	concat
Global warming will make winters a thing of the past	the global economy is not bad	A global warming is no longer a small priority	global warming is a very short term economic and development measure to have produced human rights
Yes, but private investment will always want to improve negative features of gambling	for example the private sector has been investing for decades	as the best way to progress is that it is impossible to say whether she is being spent or because of the risk that is being spent on the condition	but what a lot of money is going to be able to drive a service based market

So therefore there exist no further al- ternative path to ex- plore beyond educa- tional programs?	this is a problem that is simply not about what people want to learn about how they play a nat- ural process	this is a clear ap- proach that can eas- ily be used malev- olently and is not consistent	all this is true will not offer that all women are expected to abide by
I argue that saving anyone goes against natural selection	i am sure it is a good thing that you have to be scientifi- cally scientifically	i think that is ex- actly why id is an in- correct theory	in fact behes claim is that the know not only because they are unable to abuse their intel- ligence if i know there is no god
There are situations where you would have to choose which persons to save as all cant be saved, would it not be ok in this situation to let some die to save more people?	however if a person has a right to educa- tion those who are legally prepared to use guns for crimi- nal purposes to be able to save their life and then put their lives in life	the court s respon- sibility is to the state and to iso- late people from the feeling of language and religion and is thus unacceptable in most cases when the state is being punished for society	there is no reason why a uk ban on cer- tain areas where in a year there would be no need for the vast majority of life

7.3. Results

LSTM stance classifier and the orange color represent the results for the Chabot version that used the LSTM SAM as a stance classifier. Ratings for both classifiers had a distribution towards the lower ratings, with the majority of the votes for rating 3 and the average of 4.56 ± 2.03 for the LTSM stance classifier and 3.38 ± 1.94 for the LSTM SAM stance classifier. The t examination is a statistical hypothesis examines that was used to identify if there was any significant uniqueness between the means of two different classifiers for the conversation flow's naturalness. The significance level was set to 0.05 (the most commonly used significance level that was used

ISSN: 2233-7857 IJFGCN

to compare the t-test value with). According to the t-test, that showed 0.08, there is a slightly significant difference between the results for the naturalness of the conversation flow for the different classifiers used.

Users felt that the Chabot did not understand their inputs. For the LSTM classifier, the users said that it was hard to understand the Chabot's stance, as it always replied with either "I agree" or "I disagree" to every user argument. For the LSTM SAM stance classifier, the users felt that the Chabot's responses were unrelated to what the users said and having "I agree"/"I disagree" statements felt rehearsed and broke the continuity of the conversation. One user suggested to add statements of the type "I agree but... "to improve the conversation flow.

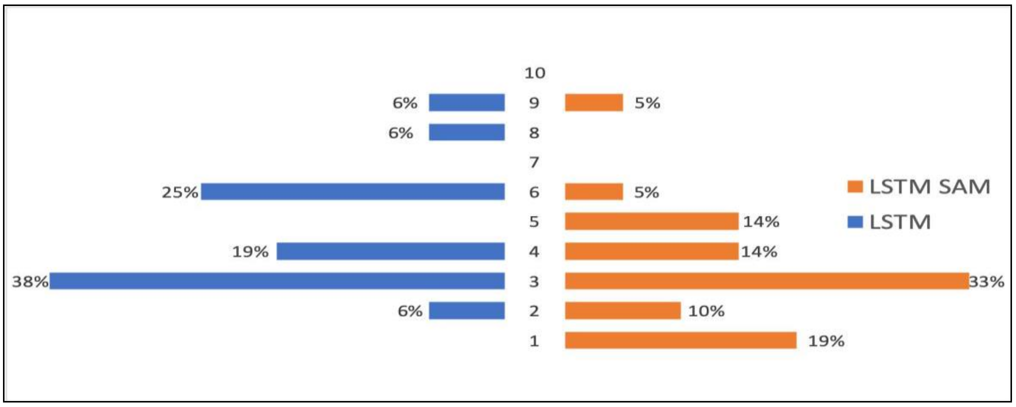


Figure.11 User ratings for how natural (human-like) the conversation flow with the Chabot felt, where 1 is unnatural and 10 is natural. The percentage score shows the distribution for the rating amongst the users for LSTM (blue color) and LSTM SAM (orange color) as a stance classifier

8. Future Scope

Following are the future scopes of the proposed system: The success of Debater opens up a whole bunch of opportunities ranging from intelligent speech assistants to solving customer queries for enterprises.

1) It can improve basic reasoning and basic composing aptitudes of youths, which will help them in their scholastics.
2) Machines that understand language are being used for Chabot's, speech assistants where the user can get their grievances settled or approved loan just by answering a few yes or no questions. It can have implications for these enterprises. Chabot's, if run on Debater's algorithm can have long, smooth conversations with the customers and help acquire real-time feedback.
3) In the future, Debater can be used to establish a platform to promote more elegant and professional debates in online comment forums.
4) It very well may be utilized by a lawyer planning for a preliminary where it could survey legitimate points of reference and test the qualities and shortcomings of a case utilizing a fake lawful discussion.
5) Whereas, it is very well may be utilized by a lawyer planning for a preliminary where it could survey legitimate points of reference and test the qualities and shortcomings of a case utilizing a fake lawful discussion.
6) In what can be a possible future scenario, we might see the Lok Sabha Election debates hosted by AI instead of some biased news anchors who come to the dais with their own agenda. AI Debater can be fed with all the grievances of the public and it can form a list of queries to ask the candidates, without any reluctance.
7) These applications can pave the way for a future where people can have healthy debates without the danger of running into self-made echo chambers.

10. Conclusion

Thus, we have identified implicitness as a major remaining problem in argument mining. Our proposed AI Debating system is effective in providing unbiased viewpoints for any debate topic. The system proves to be cost effective and requires less maintenance due to automation. Hence, a common man can afford to purchase such debater system to keep his legal services affordable and will actively participate in legal cases without any compromise. The merit of this system is that the user has knowledge of what the opponent is going to come up and anticipation of such thing is going to be crucial in critical cases. This system is useful in commercial places such as enterprises, courts, General assembly.

Acknowledgments

We have great pleasure in presenting the report on WYCH Debater. We take this opportunity to express our sincere thanks towards our guide Prof. Shweta Sharma for providing the technical guidelines and the suggestions regarding line of this work. We would like to express our gratitude towards her constant encouragement, support and guidance throughout the development of the project.

We are grateful to Dr.S.P.Kallurkar (Principal), Prof. Suvarna Pansambal (HOD, Department of Computer Engineering), without their support and advice our project would not have shaped up as it has.

We also thank all the staff of Atharva College of Engineering, Mumbai for their valuable help rendered during the course of this work. We wish to express our deep gratitude towards all our colleagues at Atharva College of Engineering, Mumbai for their encouragement.

References

11.1. Conference Proceedings

[1] Ran Levy Yonatan Bilu Daniel Hershcovich Ehud Aharoni Noam Slonim. COLING “Context Dependent Claim, Detection Proceedings” of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, pages 1489–1500, Dublin, Ireland, August 23-29 2014.

[2] Ruty Rinott, Lena Dankin, Carlos Alzate Perez, Mitesh M. Khapra, Ehud Aharoni, Noam Slonim. EMNLP Show Me Your Evidence – “An Automatic Method for Context Dependent Evidence Detection”, Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 440–450, Lisbon, Portugal, 17-21 September 2015. c 2015 Association for Computational Linguistics.

[3] “Automatic Claim Negation: Why, How and When” Yonatan Bilu and Noam Slonim. 2nd Argument Mining Workshop, NAACL, Proceedings of the 2nd Workshop on Argumentation Mining, pages 84–93, Denver, Colorado, June 4, 2015. c 2015 Association for Computational Linguistics

[4] Yonatan Bilu and Noam Slonim ACL “Claim Synthesis via Predicate Recycling”, Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pages 525–530, Berlin, Germany, August 7-12, 2016. c 2016 Association for Computational Linguistics

[5] Unsupervised Corpus-“Wide Claim Detection Recycling”, Proceedings of the 4th Workshop on Argument Mining, pages 79–84 Copenhagen, Denmark, September 8, 2017. c 2017 Association for Computational Linguistics

[6] Ran Levy, Ben Bogin, Shai Gretz, Ranit Aharonov, Noam Slonim. COLING Towards an “Argumentative Content Search Engine Using Weak Supervision Detection”, Proceedings of the 27th International Conference on Computational Linguistics, pages 2066–2081 Santa Fe, New Mexico, USA, August 20-26, 2018

[7] Henning Wachsmuth, Nona Naderi, Yufang Hou, Yonatan Bilu, Vinodkumar Prabhakaran, Tim Alberdingk Thijm, Graeme Hirs, tBenno Stein. EACL “Computational Argumentation Quality Assessment in Natural Language”, Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers, pages 176–187, Valencia, Spain, April 3-7, 2017. c 2017 Association for Computational Linguistics

[8] T. Mikolov, K. Chen, G. Corrado, and J.-f. Dean, “Efficient estimation of word representations in vector space,” ArXiv e-prints, Mar. 2013 Journal of Machine Learning Research, 3:1137-1155, 2013.
[9] K. Lee and Oenoehag, “Sentence stress in information structure”, vol. 27, 2013. Disfluencies signal thee, um, new information. J journal of Psycholinguistic Research , 32(1), 25-36
[10] S. Hochreiter and J. Schmidhuber, “Long short-term memory,Neural computation”, vol. 9, pp. 1997, t. In IEEE 1st International Conference on Neural Networks, San Diego, volume 2, pages 609-618
[11] Slava Shechtman, Moran Mordechay, “EMPHATIC SPEECH PROSODY PREDICTION WITH DEEP LSTM NETWORKS”, 978-1-5386-4658-8/18/$31.00 2018 IEEE
[12] Yosi Mass, Slava Shechtman, Moran Mordechay. “Word Emphasis Prediction for Expressive Text to Speech”, Interspeech 2018 2-6 September 2018, Hyderabad
[11] G. King and L. Zeng, “Logistic regression in rare events data, Political analysis”, vol. 9, no. 2, pp. 137–163, P1: FJJ/Shraban WV006-01 February 16, 2001.
[12] https://dl.acm.org/citation.cfm?id=2850417
[13] https://arxiv.org/abs/1503.0007512. Copyright Forms

ISSN: 2233-7857 IJFGCN