MAULANA AZAD
NATIONAL INSTITUTE OF TECHNOLOGY BHOPAL
INDIA, 462003

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Stance Detection with Sentiment Analysis and Sarcasm Detection

Minor Project Report

Semester VI

Submitted by:

Dishant Pandauria	181112226
Ragini Kalvade	181112256
Raj Mehroliya	181112244
Vanshika Agrawal	181112246

Under the Guidance of

Dr. Saritha S. K.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Session: 2018 -2022

MAULANA AZAD

NATIONAL INSTITUTE OF TECHNOLOGY BHOPAL

INDIA, 462003

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that the project report carried out on "**Stance Detection with Sentiment Analysis and Sarcasm Detection**" by the 3rd year students:

Dishant Pandauria	181112226
Ragini Kalvade	181112256
Raj Mehroliya	181112244
Vanshika Agrawal	181112246

Have successfully completed their project in partial fulfilment of their Degree in Bachelor of Technology in Computer Science and Engineering.

Dr. Saritha S. K.
(Minor Project Mentor)

DECLARATION

We, hereby declare that the following report which is being presented in the Minor Project Documentation Entitled as "**Stance Detection with Sentiment Analysis and Sarcasm Detection**" is an authentic documentation of our own original work and to best of our knowledge. The following project and its report, in part or whole, has not been presented or submitted by us for any purpose in any other institute or organization. Any contribution made to the research by others, with whom we have worked at Maulana Azad National Institute of Technology, Bhopal or elsewhere, is explicitly acknowledged in the report.

Dishant Pandauria	181112226
Ragini Kalvade	181112256
Raj Mehroliya	181112244
Vanshika Agrawal	181112246

ACKNOWLEDGEMENT

With due respect, we express our deep sense of gratitude to our respected guide and coordinator Dr. Saritha S.K., for her valuable help and guidance. We are thankful for the encouragement that she has given us in completing this project successfully.

It is imperative for us to mention the fact that the report of minor project could not have been accomplished without the periodic suggestions and advice of our project guide Dr. Saritha S.K. and project coordinators Dr. Dhirendra Pratap Singh and Dr. Jaytrilok Choudhary.

We are also grateful to our respected director Dr. N. S. Raghuwanshi for permitting us to utilize all the necessary facilities of the college.

We are also thankful to all the other faculty, staff members and laboratory attendants of our department for their kind cooperation and help. Last but certainly not the least; we would like to express our deep appreciation towards our family members and batch mates for providing the much-needed support and encouragement.

ABSTRACT

Stance detection on social media is an emerging opinion mining paradigm for social and political applications. Social Media channels allow for people to express their views and opinions about any public topics. Public sentiment related to future events, such as demonstrations or parades, indicate public attitude and therefore may be applied while trying to estimate the level of disruption and disorder during such events.

One can often detect from a person's utterances whether he/she is in favor of or against a given target entity— their stance towards the target. However, a person may express the same stance towards a target by using negative or positive language.

The project created factors the sentiment of the tweet along with detecting the stance of a target-entity. Stance detection is related to, but different from, sentiment analysis. Sentiment analysis determines whether a piece of text is positive, negative, or neutral based on the text presented. It is common that generally sentiment analysis doesn't take into account the sarcasm of the text so the model takes that as a factor also. In stance detection, systems are to determine favor-ability towards a given selected target of interest. The target of interest may not be explicitly mentioned in the text and it may not be the target of opinion in the text.

Therefore, a multilevel framework is proposed which takes into consideration the sentiment and the presence of sarcasm in the text and provides us with the overall stance towards the target.

TABLE OF CONTENTS

Certificate ii
Declaration iii
Acknowledgement iv
Abstract v

1. Introduction……………………………………………………………………………………9

2. Literature review and survey……………………………………………………………10

3. Gaps Identified……………………………………………………………………………11

4. Problem Definition………………………………………………………………………12

5. Proposed work and methodology………………………………………………………13

5.1 Proposed Work……………………………………………………………………………13

5.2 Methodology………………………………………………………………………………15

6. Results and Discusssion.…………………………………………………………………18

6.1 Linear Support Vector Classifier …………………………………………………………18

6.2 Logistic Regression ………………………………………………………………………21

7. Working Model……………………………………………………………………………28

8. Conclusion and Future Scope.……………………………………………………………30

9. References…………………………………………………………………………………31

LIST OF FIGURES

Fig. 1. Architecture of the project……………………………………………………………………………13

Fig. 2. Flow of model creation…………………………………………………………………………………15

Fig. 3. Accuracy of the sarcasm detection model………………………………………………………18

Fig. 4. News Headlines Dataset before pre-processing…………………………………………………19

Fig. 5. News Headlines Dataset after pre-processing……………………………………………………19

Fig. 6. Precision, Recall and F-1 Score for Sarcasm Detection……………………………………19

Fig. 7. Confusion Matrix for Sarcasm Detection……………………………………………………………19

Fig. 8. User Input tested for Sarcasm Detection Model…………………………………………………20

Fig. 9. Proportion Target Class for Sarcasm Detection Dataset………………………………………20

Fig. 10. Most common words in sarcastic headlines……………………………………………………20

Fig. 11. Most common words in non-sarcastic headlines………………………………………………20

Fig. 12. Accuracy of the sentiment analysis model………………………………………………………21

Fig. 13. Accuracy of the stance detection model…………………………………………………………21

Fig. 14. Comparison between Linear and Logistic Regression…………………………………………21

Fig. 15. Stance Detection Dataset before pre-processing…………………………………………………22

Fig. 16. Stance Detection Dataset after pre-processing…………………………………………………22

Fig. 17. Precision, Recall and F-1 Score for Stance Detection ………………………………………22

Fig. 18. Confusion Matrix for Stance Detection……………………………………………………………22

Fig. 19. User Input tested for Stance Detection Model…………………………………………………22

Fig. 20. Proportion Target Class for Stance Detection Dataset………………………………………23

Fig. 21. Most common words in Against tweets……………………………………………………………23

Fig. 22. Most common words in Favor tweets………………………………………………………………23

Fig. 23. Most common words in None tweets………………………………………………………………24

Fig. 24. Sentiment140 Dataset before pre-processing……………………………………………………25

Fig. 25. Sentiment140 Dataset after pre-processing………………………………………………………25

Fig. 26. Precision, Recall and F-1 Score for Sentiment for Analysis…………………………………26

Fig. 27. Confusion Matrix for Sentiment Analysis…………………………………………………………26

Fig. 28. User Input tested for Sentiment Analysis Model………………………………………………26

Fig. 29. Proportion Target Class for Sentiment 140 Dataset……………………………………………26

Fig. 30. Most common words in Negative tweets…………………………………………………………27

Fig. 31. Most common words in Positive tweets……………………………………………………………27

Fig. 32. Snapshot of Home Page…………………………………………………………………………………28

Fig. 33. Snapshot of Post Form…………………………………………………………………………………28

Fig. 35. Snapshot of Blog…………………………………………………………………………………………29

Fig. 34. Snapshot of Target Stance……………………………………………………………………………29

LIST OF TABLES

Table I. Testing scores of different Sarcasm Detection models……………………………….19

Table II. Accuracy scores of different Sentiment Analysis models…………………………..25

INTRODUCTION

Stance detection is the task of automatically determining from text whether the author of the text is in favor of, against, or neutral towards a proposition or target. The target may be a person, an organization, a government policy, a movement, a product, etc.

One can infer from Barack Obama's speeches that he is in favor of stricter gun laws in the US. Similarly, people often express stance towards various target entities through posts on online forums, blogs, Twitter, YouTube, Instagram, etc. Automatically detecting stance has widespread applications in information retrieval, text summarization, and textual entailment. Over the last decade, there has been active research in modeling stance.

The task undertaken can be formulated as follows - given a target-entity, the framework must gather all tweets posted by users about it and provide a sentiment and its stance towards the target. The model will also provide an overall stance based on the tweets gathered.

For example - *Target: Donald Trump*

Tweet: Joe Biden is the only sane presidential candidate.

In the given tweet, the target entity is not mentioned directly yet it is extracted for stance detection. To utilize social media data to its maximum extent we have to consider all factors to gain an opinion. Factors include considering tags relevant to target entity, sentiment analysis of tweet, and detection of sarcasm.

LITERATURE REVIEW AND SURVEY

Literature on stance detection includes [Somasundaran and Wiebe 2010] where a stance detection approach was presented, based on sentiment and arguing features, along with an arguing lexicon automatically compiled. This approach was reported to perform better than baseline systems which were distribution-based and unigram-based systems [Somasundaran and Wiebe 2010]. In studies such as [Walker et al. 2012a,b], it was concluded that considering the dialog structure of online debate posts improved stance classification performance on these posts.

In [Ebrahimi et al. 2016], a log-linear model for stance classification in tweets was proposed where the interactions between the stance target, stance, and sentiment were modeled. In [Gadek et al. 2017], the authors showed that extracting and using contextonyms ("contextually related words") helped improve stance detection in tweets. In [Sobhani et al. 2017], a data set for multi-target stance detection was presented together with experiments on this data set. Another recent topic closely related to stance detection is argumentation (or, argument) mining. The aim of the argumentation mining task is to identify the particular arguments, related components, and relations in natural language texts [Nguyen and Litman 2015]. These texts are usually in the form of on-line debates, legal documents, and student essays [Nguyen and Litman 2015]. There are also studies that performed joint argument mining and stance detection [Sobhani et al. 2015].

The paper, Multi-Task Stance Detection with Sentiment and Stance Lexicons,by Yingjie Li and Cornelia Caragea, on which we based most of our research on proposed an attention-based multitask learning framework and integrate lexicon information to achieve better performance. Their experimental results show that that model outperforms state-of-the-art deep learning methods for this task.

GAPS IDENTIFIED

In the literature review, it is identified that while stance detection is a highly researched topic, one of the major drawbacks is that while analyzing sentiment of a tweet, sarcasm present in the tweet is not given consideration. This can give misleading results, when a level of the framework is solely dependent on sentiment analysis.

Further, it is also observed that no feature is provided to present the overall stance of a target entity. With the high influx of tweets every second, it is hardly useful to know the stance of every tweet. Rather, one is more inclined to be certain on the overall public stance. Adding on to that, the lack of datasets present on the topic of research also cause a stagnancy in the research and we cannot further our study without the provision of more datasets.

PROBLEM DEFINITION

Social media is a vital part of interaction with public. Twitter has now become an important medium to stay up to date with the worldly on-goings. From every politician, celebrity to smallest brands, have their presence on Twitter.

Twitter 2019 Usage Stats showed -

- **500 million tweets are sent per day.**
- **Twitter users spent an average of 3.39 minutes on the social networking platform per session.**
- 83% of 193 UN member countries have a Twitter presence.

On the basis of these stats, it is inferred that a fairly accurate public stance of trending topics be it political actions, debates, controversial topics from analyzing social media posts on Twitter can be gathered. With access to social media data, it can interpret the sentiment towards the entity.

The created system analyses the posts that user post on the given web platform and analyse it according to the model previously trained and displays the stance of the post while mentioning its target and to show the extent of correctness of the same, sentiment and sarcasm detection of the same post have also been incorporated.

PROPOSED WORK AND METHODOLOGY

4.1 Proposed Work

To overcome the drawbacks of the methods that are reviewed above, a new model for stance detection is proposed. In this model, many techniques are combined to reach our final goal of emotion extraction.

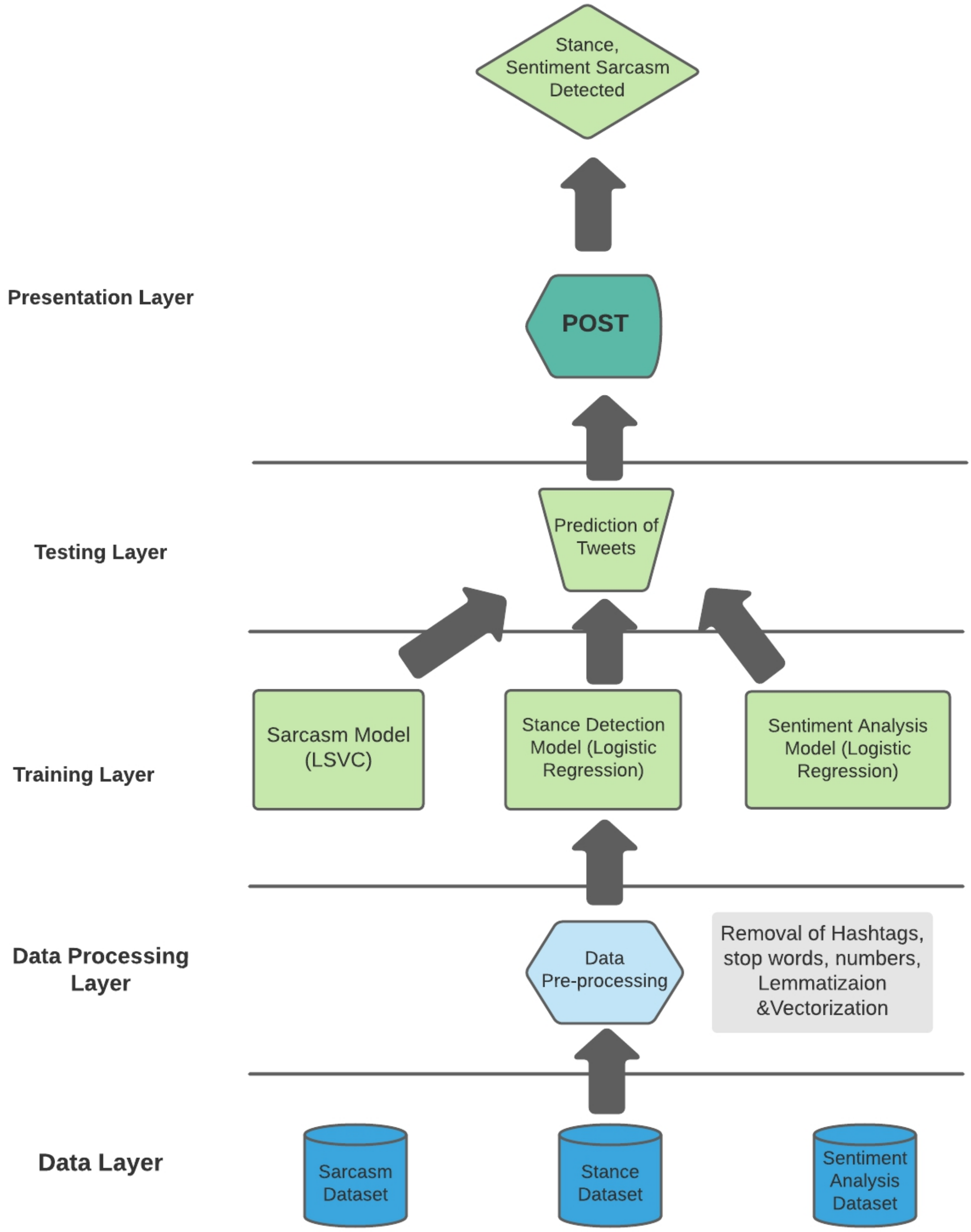

Fig. 1. Architecture of the project

The steps for the above process are documented below.

1. **Stance Detection Model:** A model for Stance Detection using attention-based framework and incorporating stance lexicon is proposed. The model will be using logistic regression and will calculate stance on the basis of the target in question.

2. **Sarcasm Detection Model:** When the stance of a particular statement is considered, none of the models incorporate the fact that whether the statement is sarcastic or not. Therefore, the use of a sarcasm detection model using the LSVC model is proposed.

3. **Sentiment Analysis Model:** In order to incorporate sentiment analysis in the calculation of stance for a tweet towards a particular target, the use of a sentiment analysis model is proposed using logistic regression in order to depict the sentiments with which the stance was made.

In addition to this, Natural Language Processing (NLP) is also used to process textual data.

4.2 Methodology

A ML-based approach for stance detection combined with sentiment and sarcasm detection is used for the project as it has been proven that stance detection for Twitter can be improved by combining the two approaches: during the first stage sarcasm and sentiment of text is detected, during the second stage a machine learning model to determine the stance is applied that uses the aforementioned models as reference.

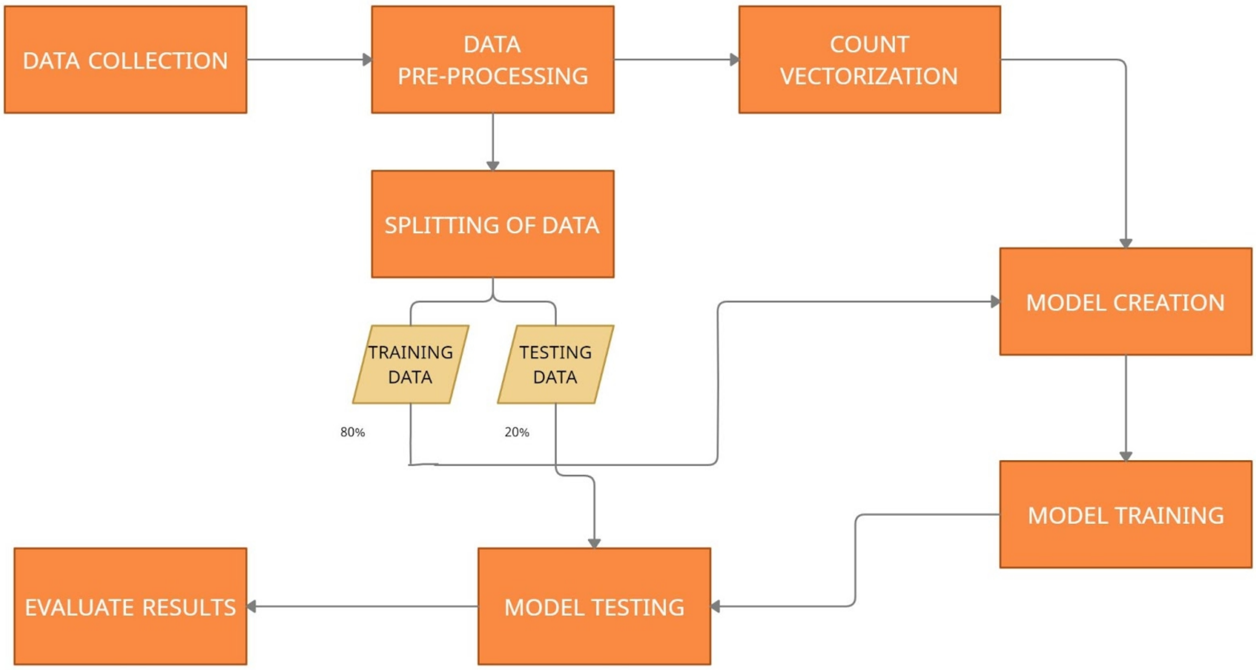

Fig. 2. Flow of model creation

1. Import certain libraries like pandas (for data analysis), numpy (for multi-dimensional arrays), nltk (natural language processing), pickle (serializing and deserializing objects) and sklearn (machine learning and statistical modelling).

2. **Data Gathering:** Getting the corpus of data that will be used to train and test the model aka the classifier. The project elects the use of manual retrieval of data mined from various sources. The dataset used is the SemEval-2016 dataset for stance detection, a twitter dataset for sentiment analysis and a headlines dataset for sarcasm detection.

3. **Pre-processing text:** Tokenizing the text, cleaning it out by removing stop words, numbers, punctuations, html tags, twitter handles, time stamps of the message, and embedded links and videos etc. Such information is largely irrelevant and may cause false results to be given by the system. The remaining words are also lemmatized in this step and any tweet data which does not have significant meaning and should not be used for analysis is eliminated.

   Pseudo Code for data preprocessing:

   ```
   1. Remove URLs
   2. Remove @ (user references) and # (hashtags)
   3. Remove emojis
   4. Convert text to lowercase
   5. Remove numbers and special symbols
   6. Remove stop words
   7. Apply lemmatization to words
   8. Return the cleaned dataset
   ```

4. **Count Vectorization:** Converting the textual words to their numeric representation (vectorizing text). This step is needed for training ML models. Simply put machines don't understand text, they get the numbers.

   Pseudo Code for Count Vectorization:

   ```
   1. Tokenization of the text
   2. Vocabulary of known words created
   3. Return an encoded vector the length of vocabulary with integer
      count for occurrence of each word
   ```

5. **Training:** Transformed data at this stage is split into training and testing sets. The training set is used to train the ML classifier by providing both the features and labels as inputs. The key point here is "experimentation"; there is no one-size-fits-all algorithm. Some classifiers are good with sentences, some are better with words etc. The classification algorithms used are: Support Vector Machine Deep Learning (neural nets) and Logistic Regression.

Pseudo Code:

```
1. Extract features and labels (X and Y)
2. Split the preprocessed and vectorized dataset into 2 - training
   (80%) and testing (20%) data (X_train, Y_train, X_test, Y_test)
```

6. **Model Creation:** Create an attention-based framework along with sentiment analysis and sarcasm detection for stance detection.

Pseudo Code:

```
1. Create the ML model
2. Train the model with the X_train and Y_train data
```

7. **Testing:** The model will then be tested with the testing data-set.

Pseudo Code:

```
1. Make predictions on the X_test and Y_test data
2. Evaluate the predictions with a classification report and
   accuracy calculation.
3. Take user input and apply prediction using the model
```

RESULTS AND DISCUSSION

6.1 Linear Support Vector Classifier

SARCASM DETECTION

The project has used the LSVC method in the sarcasm detection model. The objective of a Linear SVC (Support Vector Classifier) is to fit to the data provided, returning a "best fit" hyperplane that divides, or categorizes, the data. From there, after getting the hyperplane, then one can then feed some features to the classifier to see what the "predicted" class is. The sarcasm detection model has an accuracy of 83%.

```
Accuracy:  82.99373837729578 %
```

Fig. 3. Accuracy of the sarcasm detection model

Past studies in **Sarcasm Detection** mostly make use of Twitter datasets collected using hashtag-based supervision but such datasets are noisy in terms of labels and language. Furthermore, many tweets are replies to other tweets and detecting sarcasm in these requires the availability of contextual tweets.

To overcome the limitations related to noise in Twitter datasets, this **News Headlines Dataset** for Sarcasm Detection is collected from The Onion website which aims at producing sarcastic versions of current events and HuffPost website which collects real news headlines.

This dataset has been resourced from Kaggle to train and test the model. Each record consists of three attributes:

- is_sarcastic: 1 if the record is sarcastic otherwise 0
- headline: the headline of the news article
- article_link: link to the original news article. Useful in collecting supplementary data

	article_link	headline	is_sarcastic
0	https://www.huffingtonpost.com/entry/versace-b...	former versace store clerk sues over secret 'b...	0
1	https://www.huffingtonpost.com/entry/roseanne-...	the 'roseanne' revival catches up to our thorn...	0
2	https://local.theonion.com/mom-starting-to-fea...	mom starting to fear son's web series closest ...	1
3	https://politics.theonion.com/boehner-just-wan...	boehner just wants wife to listen, not come up...	1
4	https://www.huffingtonpost.com/entry/jk-rowlin...	j.k. rowling wishes snape happy birthday in th...	0

Fig. 4. News Headlines Dataset before pre-processing

	article_link	headline	is_sarcastic
0	https://www.huffingtonpost.com/entry/versace-b...	former versace store clerk sues over secret b...	0
1	https://www.huffingtonpost.com/entry/roseanne-...	the roseanne revival catches up to our thorn...	0
2	https://local.theonion.com/mom-starting-to-fea...	mom starting to fear son s web series closest ...	1
3	https://politics.theonion.com/boehner-just-wan...	boehner just wants wife to listen not come up...	1
4	https://www.huffingtonpost.com/entry/jk-rowlin...	j k rowling wishes snape happy birthday in th...	0

Fig. 5. News Headlines Dataset after pre-processing

Here before finalizing a model, 4 different models were implemented. After applying all the models, the scores were as follows:

Table I. Testing scores of different Sarcasm Detection models

S.No.	Model Name	Testing Score
1.	Linear Support Vector Classifier	83.75
2.	Gaussian Naïve-Bayes	73.80
3.	Logistic Regression	83.08
4.	Random Forest Classifier	79.71

It can be seen that the best scorer is the **Linear Support Vector Classifier** with an **accuracy** of **83.75%**. And remaining models have accuracies nearer to Linear Support Vector Classifier. Therefore, the most accurate one is implemented.

```
Classification Report
              precision    recall  f1-score   support

           0       0.85      0.86      0.86       746
           1       0.82      0.81      0.82       590

    accuracy                           0.84      1336
   macro avg       0.84      0.83      0.84      1336
weighted avg       0.84      0.84      0.84      1336
```

Fig. 6. Precision, Recall and F-1 Score for Sarcasm Detection

```
Confusion Matrix
[[640 106]
 [111 479]]
```

Fig. 7. Confusion Matrix for Sarcasm Detection

```
['A man cleans his apartment once every relationship'] : 1
['Could hillary clinton have what it takes to defeat the democrats in 2008?'] : 0
```

Fig. 8. User Input tested for Sarcasm Detection Model

The target class variable is imbalanced, where "Is Not Sarcastic" values are more dominating than "Is Sarcastic".

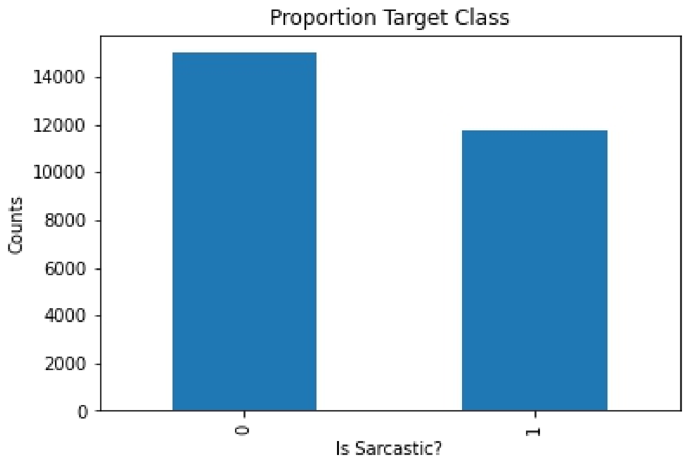

Fig. 9. Proportion Target Class for Sarcasm Detection Dataset

Fig. 10. Most common words in sarcastic headlines

Fig. 11. Most common words in non-sarcastic headlines

6.2 Logistic Regression

Logistic regression is named for the function used at the core of the method, the logistic function. Logistic Regression method is used in the stance detection and sentiment analysis model. Both have a decent accuracy and give precise outcomes.

Fig. 12. Accuracy of the sentiment analysis model

Fig. 13. Accuracy of the stance detection model

The logistic function, also called the sigmoid function was developed by statisticians to describe properties of population growth in ecology, rising quickly and maxing out at the carrying capacity of the environment. It’s an S-shaped curve that can take any real-valued number and map it into a value between 0 and 1, but never exactly at those limits.

$$1 / (1 + e^{-value})$$

Where e is the base of the natural logarithms (Euler’s number or the EXP() function in your spreadsheet) and value is the actual numerical value that you want to transform. Below is a plot of the numbers between -5 and 5 transformed into the range 0 and 1 using the logistic function.

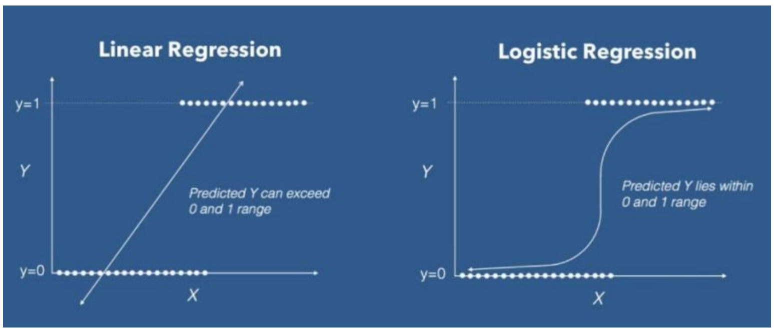

Fig. 14. Comparison between Linear and Logistic Regression

STANCE DETECTION

The dataset used to train the **stance detection** model is the SemEval-2016 dataset. It is a dataset of tweets manually annotated for stance towards given target, target of opinion (opinion towards), and sentiment (polarity). More than 4000 tweets are annotated for whether one can deduce favorable or unfavorable stance towards one of five targets ‘Atheism’, ‘Climate Change is a Real Concern’, ‘Feminist Movement’, ‘Hillary Clinton’, and ‘Legalization of Abortion’.

```
          Target                                               Tweet   Stance
ï»¿ID
\n1    Atheism  He who exalts himself shall      be humbled; a...  AGAINST
\n2    Atheism  RT @prayerbullets: I remove Nehushtan -previou...  AGAINST
\n3    Atheism  @Brainman365 @heidtjj @BenjaminLives I have so...  AGAINST
\n4    Atheism  #God is utterly powerless without Human interv...  AGAINST
\n5    Atheism  Morality is not derived from religion, it prec...  AGAINST
```

Fig. 15. Stance Detection Dataset before pre-processing

```
          Target                                               Tweet   Stance
ï»¿ID
\n1    Atheism  exalts shall humbled humbles shall exalted mat...  AGAINST
\n2    Atheism  rt prayerbullets remove nehushtan previous mov...  AGAINST
\n3    Atheism  brainman heidtjj benjaminlives sought truth so...  AGAINST
\n4    Atheism  god utterly powerless without human interventi...  AGAINST
\n5    Atheism  morality derived religion precedes christopher...  AGAINST
```

Fig. 16. Stance Detection Dataset after pre-processing

M**ultinomial logistic regression is used here**. It is a classification method that generalizes logistic regression to multiclass problems, i.e. with more than two possible discrete outcomes. The performance reports for the stance detection model are as follows:

```
Classification Report
              precision    recall  f1-score   support

     AGAINST       0.63      0.74      0.68       291
       FAVOR       0.56      0.41      0.47       156
        NONE       0.49      0.46      0.47       129

    accuracy                           0.59       576
   macro avg       0.56      0.53      0.54       576
weighted avg       0.58      0.59      0.58       576
```

Fig. 17. Precision, Recall and F-1 Score for Stance Detection

```
Confusion Matrix
[[214  39  38]
 [ 69  64  23]
 [ 59  11  59]]
```

Fig. 18. Confusion Matrix for Stance Detection

```
['Trump is better than Hillary'] : AGAINST
```

Fig. 19. User Input tested for Stance Detection Model

The target class variable is imbalanced, where "AGAINST" values are more dominating than “FAVOR” and “NONE”.

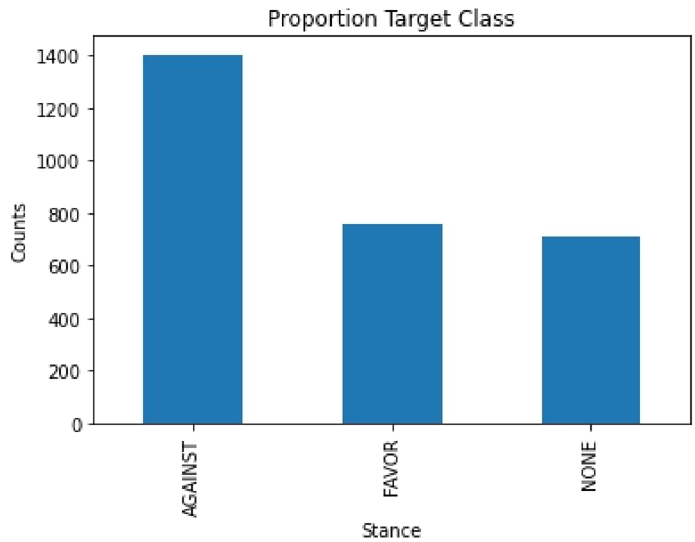

Fig. 20. Proportion Target Class for Stance Detection Dataset

Fig. 21. Most common words in Against tweets

Fig. 22. Most common words in Favor tweets

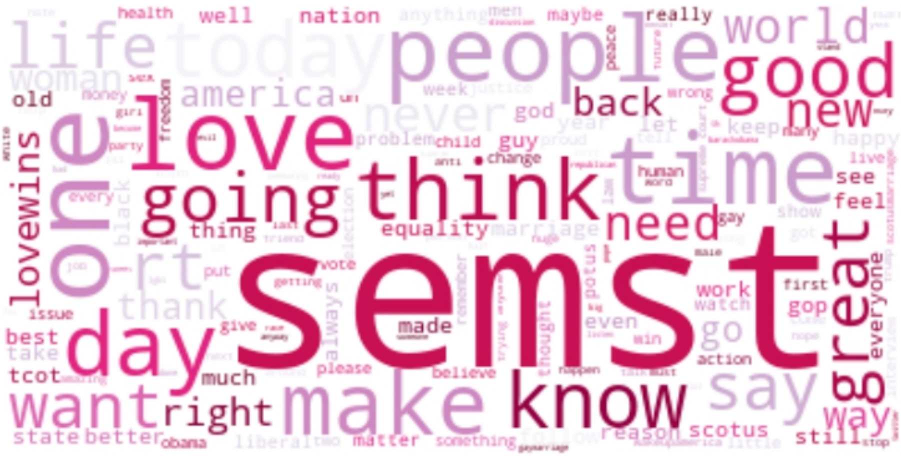

Fig. 23. Most common words in None tweets

SENTIMENT ANALYSIS

The dataset used for **sentiment analysis model** is the sentiment140 dataset from Kaggle. It contains over 1,600,000 tweets extracted using the twitter API. The tweets have been annotated (0 = negative, 1 = positive) and they can be used to detect sentiment.

ï»¿	id	label	tweet
0	1.0	0.0	is upset that he can't update his Facebook by ...
1	2.0	0.0	@Kenichan I dived many times for the ball. Man...
2	3.0	0.0	my whole body feels itchy and like its on fire
3	4.0	0.0	@nationwideclass no, it's not behaving at all....
4	5.0	0.0	@Kwesidei not the whole crew

Fig. 24. Sentiment140 Dataset before pre-processing

ï»¿	id	label	tweet
0	1.0	0.0	upset update facebook texting might cry result...
1	2.0	0.0	kenichan dived many time ball managed save res...
2	3.0	0.0	whole body feel itchy like fire
3	4.0	0.0	nationwideclass behaving mad see
4	5.0	0.0	kwesidei whole crew

Fig. 25. Sentiment140 Dataset after pre-processing

Here too before finalizing a model, 5 different models were used and tested. After applying all the models, the scores were as follows:

Table II. Accuracy scores of different Sentiment Analysis models

S.No.	Model Name	Accuracy
1.	Naïve Bayes	76.87%
2.	Logistic Regression	78.56%
3.	Linear Support Vector Classifier	61.28%
4.	Ada Boosting	56.69%
5.	Random Forest Classifier	70.12%

It can be seen that the best scorer is **Logistic Regression** with an **accuracy** of **78.56%**. Therefore, the most accurate one is chosen.

Binomial or binary logistic regression is used here. It deals with situations in which the observed outcome for a dependent variable can have only two possible types, "0" and "1". The performance reports for the sentiment analysis model are as follows:

```
Classification Report
              precision    recall  f1-score   support

         0.0       0.80      0.81      0.81     61075
         1.0       0.76      0.76      0.76     49787

    accuracy                           0.79    110862
   macro avg       0.78      0.78      0.78    110862
weighted avg       0.79      0.79      0.79    110862
```

Fig. 26. Precision, Recall and F-1 Score for Sentiment Analysis

```
Confusion Matrix
[[49340 11735]
 [12033 37754]]
```

Fig. 27. Confusion Matrix for Sentiment Analysis

On testing the model for user inputs:

```
['Trump is a good candidate'] : 1.0
['Could hillary clinton have what it takes to defeat the democrats in 2008?'] : 0.0
```

Fig. 28. User Input tested for Sentiment Analysis Model

The target class variable is imbalanced, where "Negative" values are more dominating than “Positive”.

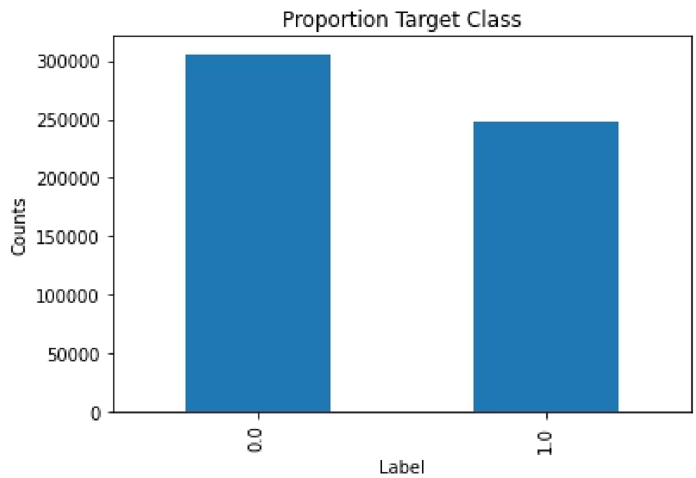

Fig. 29. Proportion Target Class for Sentiment140 Dataset

Fig. 30. Most common words in Negative tweets

Fig. 31. Most common words in Positive tweets

WORKING MODEL

In this project, a web application is created which enables the users to predict the stance, presence of sarcasm and the sentiment behind a tweet. The machine learning models employed have a decent accuracy and provide instantaneous results.

A platform has been provided where one can post a tweet on a collection of target-entities and the system provides a stance, sentiment and sarcasm of the tweet. It also constituted a feature which calculates and displays the overall stance of all targets.

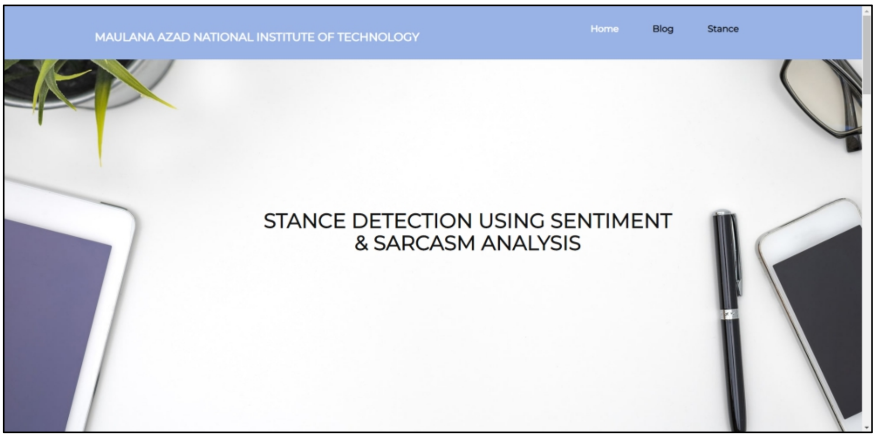

Fig. 32. Snapshot of Home Page

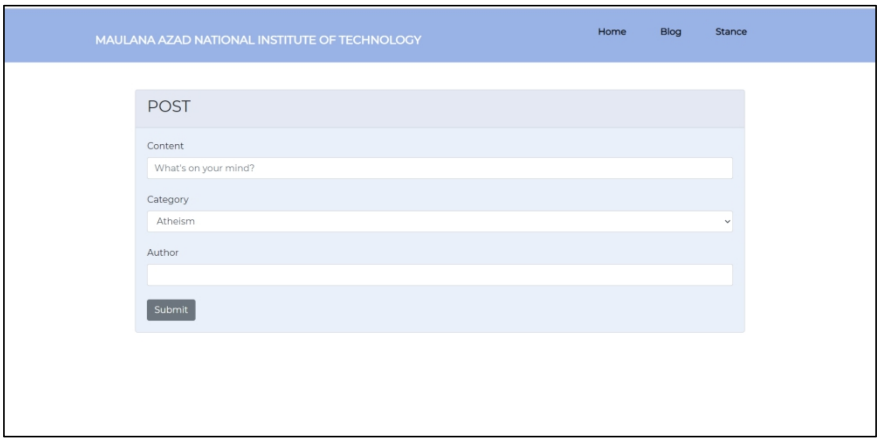

Fig. 33. Snapshot of Post Form

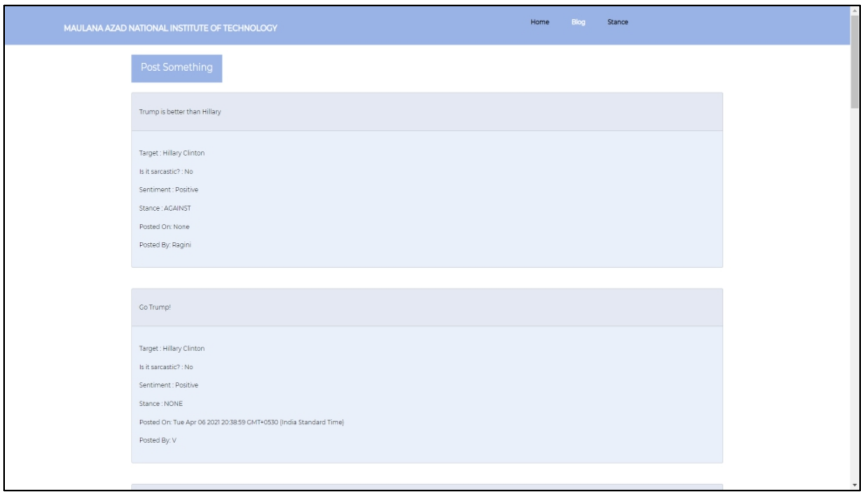

Fig. 34. Snapshot of Blog

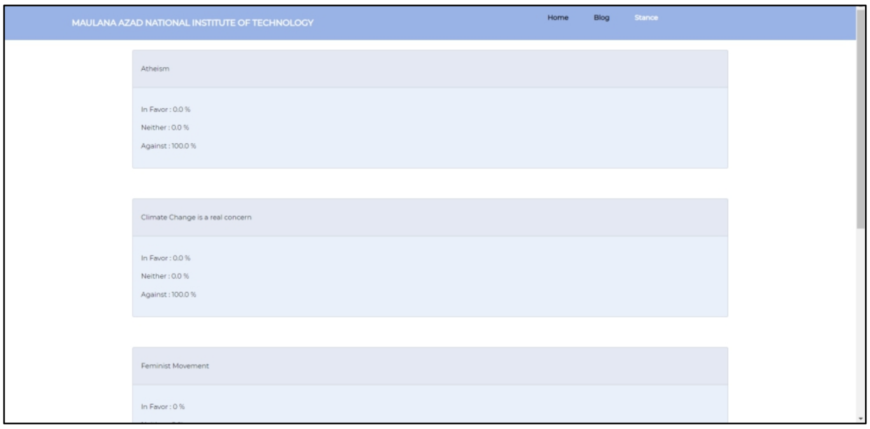

Fig. 35. Snapshot of Overall Stance of a target

CONCLUSION AND FUTURE SCOPE

From the research done, a conclusion was made that stance detection will remain an important feature in this social media age. The different models have depicted that there is room from improvement and experimentation in stance detection.

One can often detect from a person's utterances whether he/she is in favor of or against a given target entity— their stance towards the target. However, a person may express the same stance towards a target by using negative or positive language. Therefore, this project shows that while knowing the sentiment expressed by a tweet is beneficial for stance classification, it alone is not sufficient. The sentiment analysis has a minuscule effect on detecting the accurate stance due to the complexity of stance modelling on social media. Stance can be in favor, against or none towards a target and it can have any sentiment positive negative or neutral.

The intention is to present an exhaustive review of stance detection techniques on social media, including the task definition, different types of targets in stance detection, features set used, and various machine learning approaches.

To sum up, a multi-level framework based on ML models and stance, sentiment analysis and sarcasm detection as factors has been employed. By determining whether text is sarcastic or not, a better insight is provided for stance detection. The project built has provided with fairly accurate results and gives an overall stance of the target-entity in real time.

The different models have shown that there is room from improvement and experimentation in stance detection. The future scope of the project would be to remove the restriction of target and dynamically update the target entities without requiring retraining of the models' multiple times. Also, the project could be improved if the output of the sentiment analysis and sarcasm detection models be used as a layer for detection of stance.

REFERENCES

[1] Yingjie Li and Cornelia Caragea. 2016. Multi-Task Stance Detection with Sentiment and Stance Lexicons. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP).Pages 6299-6303

[2] Saif M. Mohammad, Parinaz Sobhani and Svetlana Kiritchenko. 2016. Stance and Sentiment in Tweets. ACM Transactions on Internet Technology. Pages 0-0.8h

[3] Linh Vu and Thanh Le. 2017. A lexicon-based method for Sentiment Analysis using social network data. Proceedings of the International Conference on Information and Knowledge Engineering (IKE); Athens, (2017). Pages 10-11

[4] Abeer Aldayel and Walid Magdy. 2020. Stance Detection on Social Media: State of the Art and Trends. eprint arXiv:2006.03644. Pages 1-16.

[5] Olga Kolchyna, Tharsis T. P. Souza, Philip C. Treleaven and Tomaso Aste. 2015. Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination. arXiv:1507.00955v3 [cs.CL] 18 Sep 2015. Pages 1-20.

[6] Dilek Küçük and Fazli Can. 2018. Stance Detection on Tweets: An SVM-based Approach. arXiv preprint arXiv:1803.08910, 2018. Pages 1-8.

[7] Swapna Somasundaran and Janyce Wiebe, 2010. Recognizing Stances in Online Debates. Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP. Pages 1-9.

[8] 2019. Sentiment Analysis – The Lexicon Based Approach.

https://alphabold.com/sentiment-analysis-the-lexicon-based-approach/

[9] 2019. Sentiment Analysis – Machine Learning Classification Use Case.

https://alphabold.com/sentiment-analysis-machine-learning-classification-use-case/

[10] Sarcasm Detection: Step towards Sentiment Analysis
https://towardsdatascience.com/sarcasm-detection-step-towards-sentiment-analysis-84cb013bb6db

[11] News Headlines Dataset for Sarcasm Detection

https://www.kaggle.com/rmisra/news-headlines-dataset-for-sarcasm-detection

[12] The SemEval-2016 Stance Dataset
https://www.saifmohammad.com/WebPages/StanceDataset.htm#:~:text=SemEval%2D2016%20Stance%20Dataset&text=A%20dataset%20of%20tweets%20manually,id%20information%20is%20available%20here.

[13] Sentiment140 dataset with 1.6 million tweets

https://www.kaggle.com/kazanova/sentiment140