Temporal analysis of misinformation tweets and interpreting its behavioral aspect

*

Sanjana S
Computer Science, B.Tech
PES University
PES2UG20CS549
Bangalore, Karnataka
sanjanasrinidhi1810@gmail.com

Varshini Jayasankar
Computer Science, B.Tech
PES University
PES2UG20CS567
Bangalore, Karnataka
varshinijais13@gmail.com

Rithvik Jayaram Kikkeri
Computer Science, B.Tech
PES University
PES2UG20CS583
Bangalore, Karnataka
rithvikjk@gmail.com

***Abstract*—The 2016 US Elections served as a focal point to understand how media could be used to persuade the majority of the population to solicit votes for a particular political party. The major concern around the issue was about handling the propagation of misinformation, which implicitly drove people to take uninformed decisions. Hence, in this work, we aim to build a misinformation classification model based on 2016 US elections that can correctly classify whether a given news is misinformation or not and thus keep the general public aware of the ground reality while making cardinal decisions. Through the extraction of linguistic features from textual data and training these features on various Machine Learning models, we have been able to achieve an accuracy of 98% with XGBoost Classifier. Comprehensive analysis of the results obtained show a strong correlation between misinformation and events, and misinformation and polarity of sentiments.**

I. INTRODUCTION

Ever since the internet has taken over the modern world, it has served as a platform for mass information sharing. The open access to the internet, connecting various people across the world, has had its fair share of pros and cons. While the internet has been used to save lives and bring about awareness and create enormous changes in the society, it has also been a platform for hate, negativity and misinformation. The repercussions of misinformation came into light after 2016 US elections when the company Cambridge Analytics harvested the data of up to 87 million Facebook profiles. Cambridge Analytica used the data to provide analytical assistance to the 2016 presidential campaigns of Ted Cruz and Donald Trump. This act of using people's data to solicit their votes using smart technology, without the knowledge of the general public, showed the extent to which the internet can be used for destruction. The availability of unauthentic data on media platforms has gained massive attention among researchers and become a hot-spot for sharing misinformation. Misinformation has been an important issue due to its tremendous negative impact, it has increased attention among researchers, journalists, politicians and the general public. Generally, misinformation can be defined as something that is written or published with the intent to mislead the people and to damage the image of an agency, entity, person, either for financial or political benefits. Due to its harmful consequences on people and society, in this work, we have developed a classifier that can classify misinformation from real news and alleviate its impact.

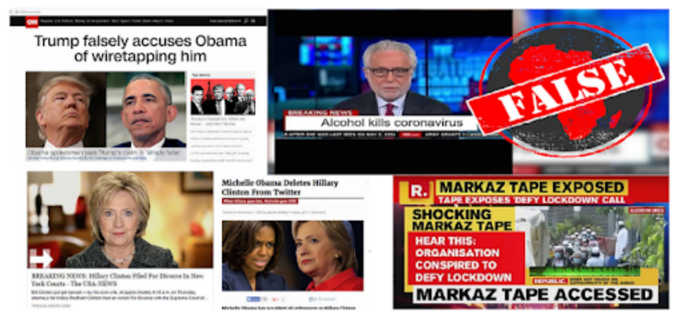


Fig. 1. Examples of misinformation news in the media

II. RELATED WORK

In literature, various methodologies have been proposed to analyze and detect fake news pertaining to elections. One such paper found that social media was the primary forum for the circulation of fake news.Voting patterns were found to be strongly correlated with the average daily fraction of users visiting websites serving fake news[1]. Their model was based on homophily in social networks to explain this linear association. We have also used temporal analysis on our data set,which was done in this paper. The paper by Richard Gunther, Paul A. Beck, and Erik C. Nisbet[2] claim that social media and fake news had an impact on the US elections of 2016. The authors had conducted a survey and had proved this correlation empirically.They have taken the voters who voted Obama in their survey. They had processed how the fake news had changed their stance from Clinton to Trump. Another methodology which is a BERT-based deep learning model using Convolutional Neural Network (CNN) called

FakeBERT,gave an accuracy of 98% in detecting fake news[3]. The model could handle ambiguity in the language.They utilize BERT as a sentence encoder, which can accurately get the context representation of a sentence. BERT was more powerful than Multinomial Naive Bayes using GloVe(accuracy: 89%). Another transformer based model called Faker, detected fake news early using temporal and spatial information[5]. Their model gave a 96% accuracy in detecting fake news.Researches show that a simple Naïve Bayesian model based on metadata mining and analysis can also be used to build a classifier that can detect fake news[4]. Feature extraction was bootstrapped and it proved to be a better method compared to the previous datasets. Using a dense neural network, a finely tuned TF-IDF model [6] gave an accuracy of 94.21%. They used Stance Detection, which is the process of automatically detecting the relationship between two pieces of text. The stance was predicted between the news headline and the article.

III. DATASET CONSTRUCTION

A. DATA EXTRACTION

Publicly available dataset Fake news classifier has been extracted from Kaggle. Data set was balanced with the distribution of of 21417 real news and 23481 of misinformation.

Feature	Description
Title	News Headline
Text	News Article
Subject	Category of news - politicsNews,worldnews,Government News, left-news,Middle-east
Date	Date the article was published
Label	Label indicating whether the news is misinformation or not.

TABLE I
TABLE:ATTRIBUTES FOR THE CONSTRUCTED DATASET DESCRIBING EACH COLUMN CORRESPONDING TO THE COLUMN HEADER.

B. DATA CLEANING AND PREPROCESSING

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset. When combining multiple data sources, there are many opportunities for data to be duplicated or mislabeled. Hence, the data set has been cleaned to remove redundant and inconsistent data. Categorical features, ‘title’ and ‘text’, have been cleaned to remove links and punctuations. Stop words have been removed from the categorical features in order to ensure better representation of the data. Training data set and testing data set have been created by splitting the dataset in the ratio of 70:30.

IV. FEATURE EXTRACTION

In order to represent categorical data in the form of meaningful vectors, TF-IDF transformer along with Count Vectorizer has been used for automatic feature extraction. The tf-idf value increases in proportion to the number of times a word appears in the document but is often offset by the frequency of the word in the corpus, which helps to adjust with respect to the fact that some words appear more frequently in general. TF-IDF uses two statistical methods, first is Term Frequency and the other is Inverse Document Frequency. Term frequency refers to the total number of times a given term t appears in the document doc against (per) the total number of all words in the document, which is computed using the Count Vectorizer and the inverse document frequency measure of how much information the word provides. It measures the weight of a given word in the entire document. IDF shows how common or rare a given word is across all documents.

idf(t) = log [n / df(t)] + 1
tf-idf(t, d) = tf(t, d) * idf(t)

	Train Data		Test Data	
Dataset Name	Misinformation news	Real news	Misinformation news	Real news
Fake news classifier	15945	14916	7067	6403

V. METHODS

A. MISINFORMATION CLASSIFICATION

Misinformation classification has been performed on the collected dataset by classifying a given news as one of the two labels.

Class	Description	News Article(Example)
0	Misinformation	Without Evidence, Trump Launches 59 Cruise Missiles, Destroying Syrian Air Force Base
1	Real	U.S. Republicans want to slash Obama’s security council

TABLE II
DESCRIPTION OF THE LABELS USED FOR MISINFORMATION CLASSIFICATION

Classification was done by training the data on various machine learning models including Decision trees, Random forests, K-nearest neighbors, Linear support vector machine, Logistic regression, Naive Bayes classifier, Passive aggressive classifier, Adaboost and XGboost. Among these models, XGBoost was able to classify the news into misinformation or real with highest accuracy score, followed by Logistic regression and K-Nearest neighbors.

- Misinformation classification with XGBoost
  - Features have been extracted from the training data by passing it to the pipeline of Count Vectorizer, TF-IDF Transformer and XGB Classifier respectively.
  - The parameters colsample_bytree and subsample, representing random subsample of columns when new tree is created and subsample ratio of the training instances respectively, were finetuned and it was observed that the value of 0.5 for both colsample_bytree and subsample resulted in convergence of the model, yielding the best results for classification.

B. SENTIMENT ANALYSIS

We have used VADER in order to analyze sentiments in the news articles to understand the polarity of sentiments in misinformation, which may have substantially impacted the election results. VADER uses a combination of sentiment lexicon which are generally labeled according to their semantic orientation as either positive or negative. VADER outputs a probability distribution of the news being neutral, positive and negative. In this work, we consider the ‘compound’ score given by VADER, which is a metric that calculates the sum of all the lexicon ratings which have been normalized between -1(most extreme negative) and +1 (most extreme positive).

The news is labeled as:

- positive sentiment, if compound score >= 0.05
- neutral sentiment, if compound score >-0.05 and compound score <0.05
- negative sentiment, if (compound score <= -0.05)

Class	News headline	Compound score	Overall sentiment
0	Presidents Bush and Clinton: Be humble in victory, responsible with power	0.9902	Positive
1	Trump Just Openly Called For A Third World War	-0.9763	Negative
1	Senate to vote on tax plan this week, No. 2 Republican says	0	Neutral

TABLE III
DESCRIPTION OF THE SENTIMENTS USED FOR EMOTION CLASSIFICATION

C. TEMPORAL ANALYSIS

Time series data have been plotted and analyzed in order to infer the impact of misinformation on elections. In this work, we have considered time series for misinformation over the period of time from March 2015 to February 2018 over coarser monthly granularity and finer daily granularity. Time series data for emotions has been analyzed to check for the presence of correlation between polarity of sentiments and propagation of misinformation. Further, the time series plots have been assessed with reference to events that have occurred during the time, which could be an indicative of a possible stimulus to the misinformation rate.

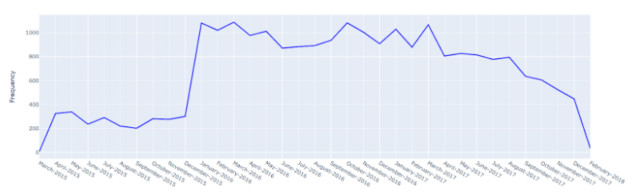


Fig. 2. Time series of frequency of misinformation news per month plotted from March 2015 to February 2018.

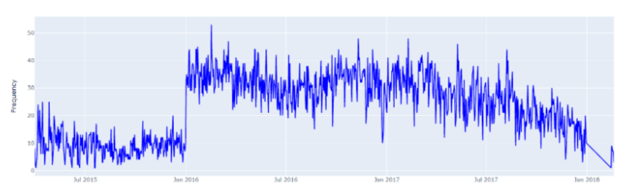


Fig. 3. Time series of frequency of misinformation news per day plotted from March 2015 to February 2018.

VI. DISCUSSION

The peer review held for this work helped in gaining various insights into the intricate details of the project and its application in the real world. There were notable improvements suggested and worked on, whose results helped in gaining a better understanding of the data set as well as the modeling aspect. With respect to the methods, one of the suggested models was passive aggressive classifier, which is an incremental supervised machine learning model that retains the model if the classification is correct and updates the model only when the classification made by it is incorrect. The use of passive aggressive classifiers in our work yielded a result of 49.61%. Since incremental learning does not have a big picture about the data as they work their way through each instance, the result will be affected by the order of presentation of samples. This drawback of passive aggressive classifiers led to low accuracy in the misinformation classification task. However, analyzing the working and drawback of passive aggressive classifiers for classification led to development of models based on ensemble learning. Boosting is an iterative ensemble learning method that combines a set of weak learners into a strong learner to minimize training errors.In each iteration, a new classifier is trained to try to correctly classify instances that the previous classifier got wrong. XGBoost algorithm was our best performing model which could classify the news correctly with an accuracy of 98.3%. The trend of misinformation changing with time is also addressed in this work by performing time-series analysis. Lead-lag analysis of time series can be useful in forecasting how various factors change with time. Further, for better learning of classification task with changing time and data, the model must be trained on new data constantly to ensure that it learns changing trends in elections and the news related to elections. This work has been performed for a dataset that comprises news regarding the 2016 US elections. The model can be validated for election data across other countries in order to generalize the results obtained.

VII. RESULTS

A. MISINFORMATION CLASSIFICATION

We have investigated and analyzed the results with several classifiers having different types of learning paradigms. Classification results demonstrate that the capability of automatic feature extraction plays an essential role in the accurate detection of misinformation. Our proposed model, XGBoost, was able to classify news with better accuracy of 98.3% as compared to existing benchmarks. Logistic regression was the second best performing model with an accuracy of 72.08%. Other models which were trained on the data included Decision trees, Random forests, K-nearest neighbors, Linear support vector machine, Naive Bayes classifier, Passive aggressive classifier and Adaboost. One of the primary reasons most of the Machine Learning models could not provide results as per the benchmarks was because of the poor feature extraction from the categorical data which led to low accuracy scores. This was overcome through bootstrapping the data, which penalized incorrect classifications, thereby, providing a better representation of the linguistic features.

B. TIME SERIES ANALYSIS OF MISINFORMATION

The time series plot of misinformation from the time March 2015 to February 2018 shows evident peaks. In order to

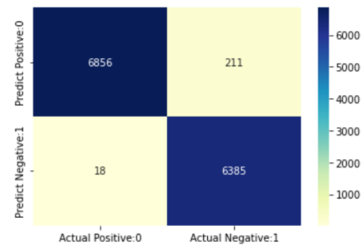


Fig. 4. Confusion matrix for XGBoost classifier.

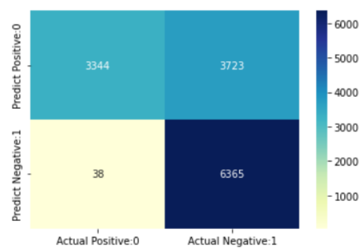


Fig. 5. Confusion matrix for Logistic Regression

Model	Accuracy	Precision	Recall	F1-score
XGBoost	0.98	0.98	0.98	0.98
Logistic Regression	0.72	0.82	0.72	0.70
K-Nearest Neighbors	0.64	0.77	0.64	0.61
Decision trees	0.55	0.76	0.55	0.45
Random Forests	0.54	0.77	0.54	0.44
Linear Support Vector Machine	0.53	0.76	0.53	0.42
Adaboost	0.50	0.75	0.50	0.36
Passive Aggressive Classifier	0.50	0.75	0.50	0.35
Naive Bayes Classifier	0.50	0.75	0.50	0.35

Fig. 6. Classification report for all models.

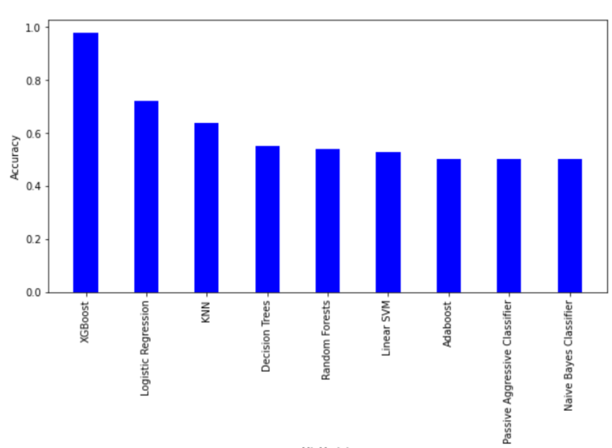
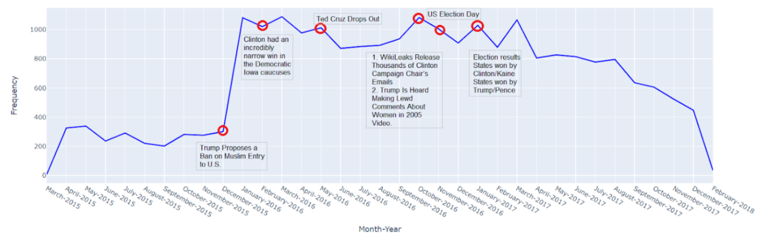


Fig. 7. Accuracy scores for all models. XGBoost performed the best with an accuracy of 98% followed by Logistic Regression with an accuracy of 72%

understand the propagation of misinformation and the cause for the fall and rise of misinformation through the years, we have identified the timeline of elections whose data was scraped[7][8]. The results have been indicative of a direct relationship between certain events and misinformation rate.



C. TIME SERIES ANALYSIS OF SENTIMENTS

In order to relate the negative effects of misinformation with elections, sentiment analysis has been performed using VADER and compound score for each news is extracted.

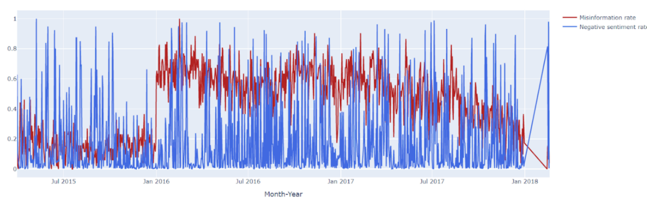


Fig. 8. Time series of misinformation rate with rate of negative sentiment taken everyday from March 2015 to February 2018.

Lead-lag analysis done on the data showed that misinformation lags sentiment by 6 days.

VIII. CONCLUSION

In this work, we have built a model for misinformation classification using various Machine Learning algorithms among which bootstrapping techniques with XGBoost achieved an accuracy of 98% which is on par with the current benchmarks. We have also analyzed and interpreted the effects of misinformation on temporal and behavioral aspects that have an adverse effect on elections. We plan on creating an election dataset pertaining to other countries as well through which we can validate and generalize the results we have obtained in this paper. Further, deep learning models have proved to be better at feature extraction which might extract better results and accuracies.

CONTRIBUTIONS

1) Sanjana S(PES2UG20CS549) - Modeling of XGBoost, Logistic regression, Decision trees, Random forests, Sentiment analysis, Temporal analysis of sentiments.
2) Varshini Jayasankar(PES2UG20CS567) - Modeling of Passive aggressive classifier, Linear support vector machine and Adaboost, Temporal analysis of misinformation.
3) Rithvik Jayaram Kikkeri(PES2UG20CS583) - Data extraction, Data pre-processing, Modeling of Naive Bayes classifier and K-nearest neighbors.

REFERENCES

[1] Adam Fourney, Miklos Z. Racz, Gireeja Ranade, Markus Mobius, and Eric Horvitz. 2017. Geographic and Temporal Trends in Fake News Consumption During the 2016 US Presidential Election. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM '17). Association for Computing Machinery, New York, NY, USA, 2071–2074. https://doi.org/10.1145/3132847.3133147

[2] Fake News Did Have a Significant Impact on the Vote in the 2016 Election by Richard Gunther, Paul A. Beck, and Erik C. Nisbet.

[3] Kaliyar, Rohit Kumar et al. "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach." Multimedia Tools and Applications 80 (2021): 11765 - 11788.

[4] Chaudhary, Pratik and Pokharna, Shrenik and Shinde, Soumil and Panjeta, Manisha, Fake News Detection During 2016 U.S. Elections Using Bootstrapped Metadata-Based Naïve Bayesian Classifier (March 31, 2022). Available at SSRN: https://ssrn.com/abstract=4074884

[5] Raza, S. (2021). Automatic Fake News Detection in Political Platforms - A Transformer-based Approach. Proceedings of the 4th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE 2021).

[6] Thota, Aswini; Tilak, Priyanka; Ahluwalia, Simrat; and Lohia, Nibrat (2018) "Fake News Detection: A Deep Learning Approach," SMU Data Science Review: Vol. 1: No. 3, Article 10. Available at: https://scholar.smu.edu/datasciencereview/vol1/iss3/10

[7] https://abcnews.go.com/Politics/key-moments-2016-election/story?id=43289663

[8] https://en.wikipedia.org/wiki/Timeline_of_the_2016_United_States_presidential_electionDecember_2015