Fake News Detection in Political Artefacts

Swapnil Roy[1], Sunny Khatri[2], Merin Meleet[3]

[1,2]Student, Dept. of Information Science and Engineering, R.V. College of Engineering, Karnataka, India
[3]Assistant Professor, Dept. of Information Science and Engineering, R.V. College of Engineering, Karnataka, India

---***---

Abstract - *The world has witnessed an incredible amount of information relating to politics in the past few years which has led people to believe whether the information swaying all over the social media or any broadcasting platform is good for their welfare or their country. With this comes the fake news, which has created a sense of ambiguity on the authenticity of the information because the channels might tend to fabricate any political news available to provide uniqueness in the information as compared to its competitors. Viewers are getting attracted and further getting infected by the fake news easily, which has brought about an overwhelming distrust among the people and the government. An important goal in improving the trustworthiness of information in online/offline media networks is to identify the fake news from time to time. This paper aims at developing methodologies to detect the fake news present using a proper supervised model designed. This project addresses the need to detect the political news articles scrapped form websites. This paper introduces a fake news detection model capable of extracting explicit features extracted from the textual information, it then chooses the best algorithm by comparing various other algorithms under certain features to properly classify the news public shed in political articles. It would finally be used to predict any political news given to it manually and classify them as true or false.*

Key Words: **Fake news, Random Forest, Stochastic Gradient Classifier, t-SNE, Truncated SVD, Logistic Regression, Principal Component Analysis(PCA)**

1. INTRODUCTION

The news or the information that we receive throughout the broadcasting media cannot be relied upon. Over the last few years spreading fake news has increased so much that it has started affecting the political norms and societal issues. The most common platform that has been identified for spreading such malicious rumor is social media. Political parties now rely upon the power of social media for campaigning and has given rise to ever increasing rumors floating around the media, these rumors might include making false accusations on rival parties or making false statements on political reforms. Government in their respective countries are accountable for any such reformation done and be accountable to the people for making such statements. People tend to believe such news floating all around the media and try making false expectations from the government. Therefore, it requires people to be utterly careful before trusting the particular media for any such political news and plan their respective actions accordingly. People might also have the actions to report any such malicious rumors through social media or any other platform and take necessary actions towards the respective source of those information. Lack of constant supervision by the respective ruling government has allowed the media to run without any hesitation and spread false information.

Hence this paper focuses on the detection of fake news accurately so that it helps to report all such falsified political news available all over the media.

2. RELATED WORK

The studies relating to the BERT model method [1] gives more important insight towards the important comparison between the statements from the Fact Data Corpus built using NLP techniques and the given statements required prediction. There were other comparison methods done from the Fact Data Corpus [3] which took into account a Korean model for detection. The studies were based on Bidirectional Multi-Perspective Matching for Natural Language Sentences for the comparison between the statements. However, there were techniques providing more credible techniques to classify the statements [4] developed an infusive deep learning model that potentially identified the news with accuracy exceeding the BERT models but there were no evidence checking for the authenticity of the news but the algorithm found success ion huge collection of datasets. Sentiments of the statements played a major role for [2] which took an extra column in the data frame for sentiment and predicted the results by comparison of various features present in the words. [5] improved the techniques on feature extraction by giving more insights TextBlob natural language toolkits that were used to develop a novel fake news detector which used quoted attribution in a Bayesian machine learning system as a key feature to estimate the likelihood that a news article is fake. Therefore, this paper has developed a more insightful feature visualisation techniques that would be useful for the production use, and selecting algorithms based on the performance under considerate features to further build the model. The model would then be used to predict various real time inputs given to it manually.

3. PROPOSED SOLUTION

This section discusses the proposed solution for political fake news detection by combining feature visualisation steps with feature extraction. The proposed solution is shown in Fig. 1. It consists of various steps as below:

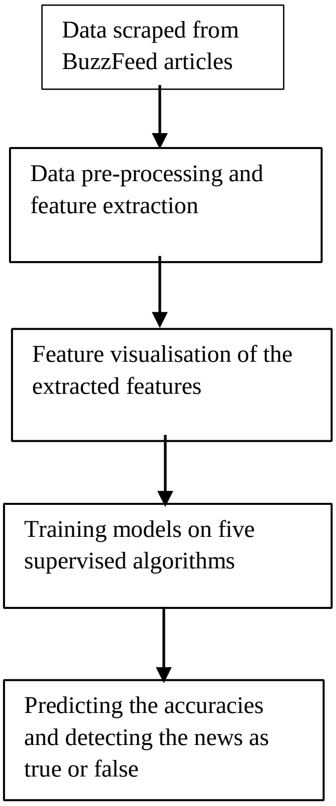


Fig-1. Proposed Solution

Step 1: The data set was prepared by scraping the political news articles from BuzzFeed site for all the fake and real news. The data set contained statement and label defined as true and false as the respective columns for the data frame formed.

Step 2: The statements from the dataset were then tokenized and the tokens were lemmatized and stemmed separately for extracting features through both bag of words and tf-idf techniques. The tf-idf technique extracted the best features in the n-gram range from one to four.

Step 3: The features were visualised separately. The count features were visualised using cosine similarity between the top five words through Word2Vec model including both Continuous Bag of Words(CBOW) and Skip Gram methods. The tf-idf features were visualised on t-SNE, Truncated SVD methods. These features were topologically analyzed using keppler mapper technique giving us the insight on the clusters of fake and real news with respect to tf-idf values for each statement. Finally adding sentiment as the new column in our dataset improved the accuracy further using Textblob toolkits. The new columns include polarity and subjectivity for each statement in the dataset.

Step 4: The extracted features were then trained under five algorithms including Naïve Bayes, Support Vector Machine(SVM), Logistic Regression, Random Forest, Stochastic Gradient Classifier. The count features and tf-idf features were trained separately and the two best performing algorithms were then tuned by GridSearchCV method of hyperparameter tuning. The best performing algorithm was then stored permanently onto the local hard disk present in the system.

Step 5: The algorithm was then saved as a file through onto the local hard disk using pickle operation. The user is then able to add any manual input to the model and predict the label of the news.

3.1 Dataset

The dataset is formed by scraping the news article published online on the site of BuzzFeed. The dataset consists of three columns before the dimensionality reduction procedure. The three columns include the serial number, the top headline present and the label of that particular news. The dimensionality reduction steps further adds two more columns in the previous dataset which includes the first two components of TruncatedSVD operations which is used to give more comprehensive and linear results between the distribution of labels. The dataset is further modified by adding two columns in the dataset which is that of sentiment operation in TextBlob library namely polarity and subjectivity. Polarity helps in defining the emotions relating to the particular statement whereas subjectivity differentiates the statements from personal opinion and factual information. Polarity varies from [-1,1] whereas subjectivity varies in the range [0,1].

4. MACHINE LEARNING ALGORITHMS

Machine Learning is the area of study that helps in computers the power or the capability of learning without actually being explicitly programmed and being a subset of AI, it is one of the most exciting technologies that the information domain has ever come across. It gives the computer the thought process similar to that of humans and that what makes it more similar to them. Machine learning is currently in active use, perhaps in many more places than one would expect. ML techniques come across three categories which clearly specify them based on predicting outputs with certain input given.

1) Supervised learning: It is based upon the fact that whenever the input observations are fed into the algorithm in form a training data it generates a function that is clearly able to predict the output based on the testing data that is given to the system to predict.

2) Unsupervised learning: The machine is not given an input observation beforehand and therefore it is forced to learn from an unlabeled set of dataset to predict the output.

3) Reinforcement learning: The learning process in this category iterates over time and all the system states eventually learn input observations over a period of time.

A. Naïve Bayes Classifier

Naive Bayes classifiers are nothing but all the collection of algorithm based on classification on Bayes Theorem. It is a family of all algorithms and not to be confused by a single algorithm. It's utmost important principle is that all the features which are classified are actually independent of each other. In this case all the headlines are independent of each other. Basically, it implies that all the classification of news whether true or false don't rely upon the fact that what might be the label of other set of news present.

B. Logistic Regression Classifier

The algorithms take into account the categorical nature of the label predicted or the dependent variable. The algorithm is mostly common in biological sciences in almost all areas of research. The algorithm differs from in linear algorithm in the sense that it is not very sensitive to the threshold set because of the sigmoid function associated with it. The operation of this classifier is very similar to that of the linear regression, except for the last stage where the finally predicted value is passed onto the sigmoid function, so that the threshold not sensitive to the set value. The value can then be decided between 0 or 1.

C. Support Vector Machine

The algorithm is best used for two class classification problem. SVM is similar to linear regression but differing in way such that the there is a hyper plane differentiating the two classes The algorithm performs best in text classification problem of Natural Language Processing. The algorithms take into account the features for the dataset and forms a hyperplane such that it succeeds tremendously in maintaining a proper distance from the nearest entity of each class and successfully developing two classes of separate features.

D. Stochastic Gradient Descent Classifier

Stochastic Gradient Descent (SGD) is an efficient approach with immense simplicity to which uses a discriminative method of learning the classifiers like Logistic Regression classifiers and Support Vector machine. SGD has been increasingly used in recent time because of its ability to scale large amount amounts of input data very efficiently, due to this nature it performs very efficiently on sparse datasets with minimal time and computational complexities.

E. Radom Forest Classifier

A random forest classifier uses the decision based on multiple decision trees that branches out to predict the correct label of the features present. The label prediction is based upon the fact that when the decision tree which predicts the labels with corrects set of features classified, that tree is used up for the model prediction. Due to the computational complexities of large number of trees also known as ensemble the performance lags behind and the time take to predict the results correctly. But on a bright side, this classifier is said to have a tremendous accuracy rate for sparse dataset.

F. Comparison of algorithms

The performance of the algorithms are based upon the fact that the which algorithm might be able to perform better on sparse dataset, also what on what features including Count and TF-IDF have greater impact on each algorithms. The metrics to compare also include accuracies of the algorithms in case the weighted F1-score of two classifier clashes.

	TF-IDF features	Count features
Naïve Bayes classifier	0.593	0.607
Logistic Regression Classifier	0.619	0.598
Support Vector Machine	0.617	0.572
Stochastic Gradient Descent	0.541	0.549
Random Forest Classifier	0.603	0.619

Table-1 Comparison of accuracies of the model

	TF-IDF features	Count features
Naïve Bayes classifier	0.593	0.646
Logistic Regression Classifier	0.669	0.610
Support Vector Machine	0.607	0.662

Stochastic Gradient Descent	0.541	0.662
Random Forest Classifier	0.623	0.701

Table-2 Comparison of F1-score of the models

5. RESULTS AND DISCUSSIONS

5.1. Dimensionality Reduction

The dimensionality reduction method has been performed on TF-IDF features present after the feature extraction steps because of the highly unvarying set of words present in each statement of the dataset, also the unrelated values had no contribution as such to the label of each statement. These were build using two techniques including Truncated SVD and t-SNE techniques. Since the former provided better linearity and congruency among the data labels its first two components was later added in the dataset.

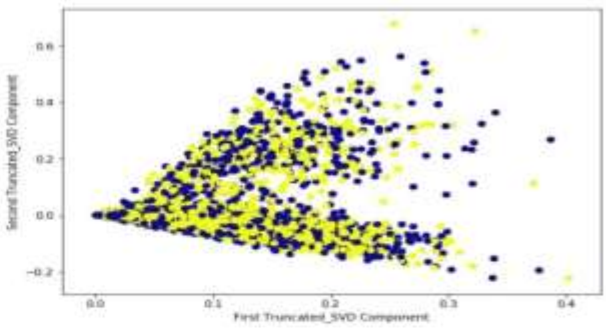


Fig-3 Scattermap representation of TruncatedSVD components

5.2 Selection of candidate models and the final model

The selection of candidate model was based on the selection of highest weighted average F1-score. This was calculated by knowing the average of the performances of each algorithm on respective features. The score for Logistic Regression model was known to be 0.6395 whereas the score for Random forest model was known to be 0.662. SVM model also performed quit well with a score of 0.645. Based on the scores, Logistic Regression and Random Forest were chosen to be the two best performing candidate models.

The final model selection was based upon the fact of optimization of hyperparameters of the two classifiers. The performance of the two optimized classifiers on grid operations were quite the same and hence accuracy of the two optimized model were used as a tie breaker. The accuracy of the optimized Logistic Regression model was found to be 0.641 whereas the accuracy of Random Forest model was found to be 0.543.

Logistic Regression model was the finally selected model used for classification.

5.3 Final results of the proposed model

The final results of the model gave us the label of the given news entered manually by the user. Along with it truth percentage value was given for a better understanding of the project. The truth percentage value above fifty (>50%) was considered as true whereas anything less than fifty (<50%) was considered as false.

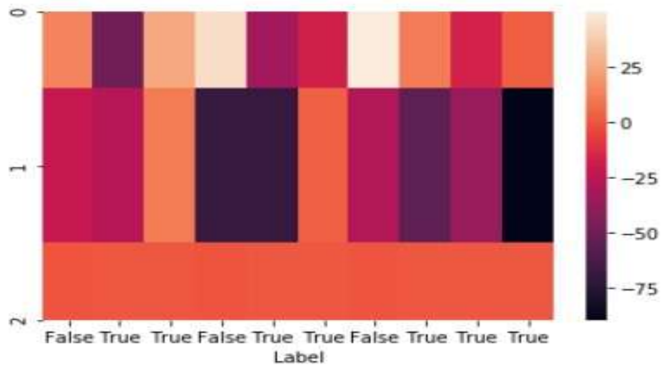
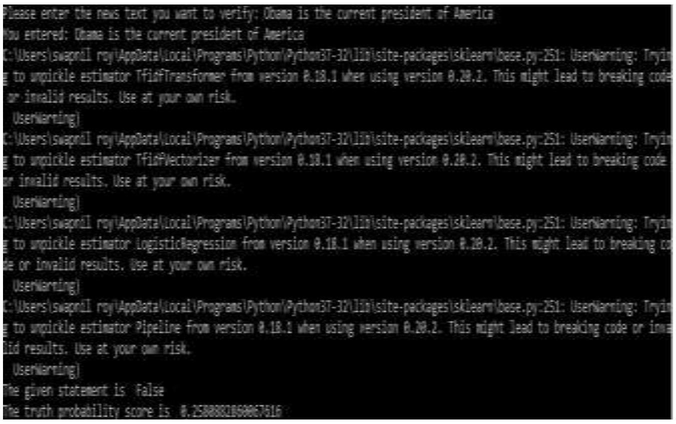


Fig-2. Heatmap representation of t-SNE components

Fig-4 Prediction of news giving true



Fig,-5 Prediction of news giving false

6. CONCLUSIONS

The project came into the picture because of the ever increasing amounts of fake news flooding all over the internet especially relating to political news when the ruling party announced a scheme and therefore there was need to overcome the limitations faced by the existing Fake News Detectors which instead of red flagging the instance of that news goes onto red flag all the sources or all the important considerations or facts of the news. These limitations were translated into new requirements and a new project “Fake News Detection in Political Artefacts” was designed and implemented, so as to overcome the limitations faced by the existing Detector models.

The current model of fake news detection takes into account all the possible advantages that the current detector machine fail to prevent. The first objective that stands out in this matter is giving the truth percentage value. The truth percentage value helps in finding the relevance of the news headlines with the real world that helps in turns find out how the authenticity of the news is important to the users. The second objective of the Detector model is to take into the account only the sentiment of the news as the reason for checking the authenticity of information, by not collecting the factual information of the news headlines including the author of the particular news articles the model stand out while comparison to other models.

REFERENCES

[1] Bhavika Bhutani, Neha Rastogi, Priyanshu Sehgal, Archana Purwar. "Fake News Detection Using Sentiment Analysis", 2019 Twelfth International Conference on Contemporary Computing (IC3), 2019.

[2] Kyeong-Hwan Kim, Chang-Sung Jeong. "Fake News Detection System using Article Abstraction", 2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2019.

[3] Terry Traylor, Jeremy Straub, Gurmeet, Nicholas Snell. "Classifying Fake News Articles Using Natural Language
Processing to Identify In-Article Attribution as a Supervised Learning Estimator", 2019 IEEE 13th International Conference on Semantic Computing (ICSC), 2019.

[4] Jiawei Zhang, BowenDong, Philip S. Yu. "DeepDiffusive Neural Network based Fake News Detection from Heterogeneous Social Networks", 2019 IEEE International Conference on Big Data (Big Data), 2019.

[5] C. H. P. Ferreira, D. M. R. de Medeiros, F. O. de França, "DCDistance: A Supervised Text Document Feature extraction based on class labels", CoRR, vol. abs/1801.04554, 2018.

[6] R. J. Sethi, "Crowdsourcing the verification of fake news and alternative facts", Proceedings of the 28th ACM Conference on Hypertext and Social Media HT ‘17, pp. 315316, 2017.

[7] John Dickerson, Vadim Kagan, VS Subrahmanian, "Using Sentiment to Detect Bots on Twitter: Are Humans more Opinionated Than Bots?", Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis, pp. 620-627, 2014.

[8] Manish Gupta, Peixiang Zhao, Jiawei Han, "Evaluating Event Credibility on Twitter", Proceedings of the 12th SIAM SDM, pp. 153-164, Apr. 2012.

[9] Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, Rada Mihalcea, "Automatic Detection of Fake News", Proceedings of the 27th International Conference on Computational Linguistics, pp. 3391-3401, August 20-26, 2018.

[10] N. J. Conroy, V. L. Rubin, and Y. Chen, “Automatic deception detection: Methods for finding fake news,” in Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community, p. 82, American Society for Information Science, 2015.

[11] V. L. Rubin, Y. Chen, and N. J. Conroy, “Deception detection for news: three types of fakes,” in Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community, p. 83, American Society for Information Science, 2015.