Fake News Detection on Social Media

Harsh Dedhia, Garima Mahto, Jyoti Sharma and
Ashwini Deshmukh

November 28, 2021

Fake News Detection on Social Media

Harsh Dedhia
Shah & Anchor Kutchhi Engineering College
Chembur, India
harsh.dedhia@sakec.ac.in

Garima Mahto
Shah & Anchor Kutchhi Engineering College
Chembur, India
garima.mahto@sakec.ac.in

Jyoti Sharma
Shah & Anchor Kutchhi Engineering College
Chembur, India
jyoti.sharma@sakec.ac.in

Ashwini Deshmukh
Shah & Anchor Kutchhi Engineering College
Chembur, India
ashwini.deshmukh@sakec.ac.in

***Abstract*— Now-a-days in 21st century people are more interested to spend most of their free time on social media like Twitters, Facebook, Instagram, telegram all these app is now in trend to engage each and every people into gossips, although it connects people of different region to communicate with others, to make new friends all over the world. Still the question arises when you are surfing on internet is that all news that you see is generally true/false? the information that you find on these social media which has so many likes, so many shares are they really true? Looking towards these unexpected people's thoughts, hard to tackle it. Thus, there is a need for a solution to check on such problems which misleads people from true facts, from true news, in social media. After such encounter of situation many researchers came forward to look onto this matter and worked together to avoid spreading of such fake news.**

Keywords—fake news, mislead information, misinformation. false news, fake news detection.

I. INTRODUCTION

Detecting fake news has become an appearing research on social media now-a-days. The quality of news on social media is increasingly an important issue, but due to expert's ability to access correct content than inaccurate once's present in this platform. In this paper, we focus on automatic identification of fake contents in online news. We aim to perform binary classification of various news articles which are available online. we aim to provide users to classify the news as fake or real and check their real source by redirecting them to the correct path. This will not only help them to know the real news but also, they will get aware of such fake news available on social media are not just fake but some of their contents are also fake too if anybody read those contents too rather than viewing the title and captions.

Fake news is false or misleading information presented as news. It often has the aim on damaging the reputation of a person or entity, or making money through advertising revenue. However, the term does not have a fixed definition, and has been applied more broadly to include any type of false information, including unintentional and unconscious mechanisms, and also by high-profile individuals to apply to any news unfavorable to his/her personal perspectives.

In this system, we are using three datasets to look for the authenticity of the news on social media platform. The datasets are:

- Fake News Source
- LIARS
- ISOT Fake News

Mass media has enormous impact on society, this results in producing news article, which are not mostly true and that makes news almost an exclusive news when it's shared among many people. Through our system, for fake news detection supported the feedback for precise news

II. DATASET INFORMATION

The greatest challenges in machine-learning based approaches in general, and in automatic fake news detection in particular, is collecting a sufficiently large, rich, and reliably labelled dataset on which the algorithms can be trained and tested [1]. In our research, we collected data from the Kaggle website and then we train and tested our model with splitting the whole data in 70- 30%.

Fake News Source:

The dataset consisted of more than 56000 articles. This dataset consisted columns like headlines, summary, sources, fake/real.

LIARS:

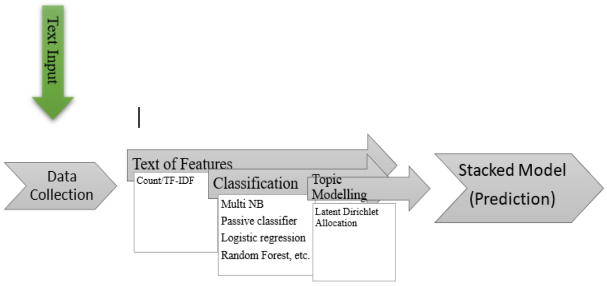
LIAR is a publicly accessible dataset for detecting fake news. POLITIFACT.COM, which includes full analysis reports and links to source materials for each case, collected 12.8K hand labelled brief remarks in diverse situations over the course of a decade. This dataset can also be used for fact-checking studies. This new dataset is an order of magnitude larger than prior publicly available false news datasets of a comparable nature. The LIAR dataset contains 12.8K human-labeled short statements from the POLITIFACT.COM API, with each statement being checked for truthfulness by a POLITIFACT.COM editor. This dataset had (10240, 8) articles with many columns like label as true/false, statements, subjects, location etc.

ISOT Fake News:

The dataset contained two types of articles fake and real News. The first file named "True.csv" contains more than 12,600 articles from reuter.com. The second file named "Fake.csv" contains more than 12,600 articles from different fake news outlet resources.

After all available dataset our team focused on these three datasets at the backend and trained our model to obtain the best result. After done with sorting of the dataset we worked on different tasks, we worked on preprocessing of all the three datasets, once done we did tokenization on whole data, as text transformation was needed, we worked on feature extraction of each and every article. Then finally moved onto the text classification which was the main task in order to find the best suite test classification algorithms for our model.

III. FLOWCHART



A. DATA PREPROCESSING:

Data preprocessing is the process of transforming raw data into an understandable format. It is also an important step in data mining as we cannot work with raw data. The quality of the data should be checked before applying machine learning or data mining algorithms.

Major Tasks in Data Pre-processing:

1. Data cleaning
2. Data integration
3. Data reduction
4. Data transformation

1) **Data Cleaning:** Cleaning data is an important step in every machine learning model, but it's especially important in NLP. Without the cleaning procedure, the dataset is frequently a jumble of words that the machine is unable to comprehend. We'll go over the procedures involved in cleaning data in a typical machine learning text pipeline.

1.1) **Punctuations**: There are various punctuations in the title text. Punctuation is rarely used because it adds no value or meaning to the NLP model. There are 32 punctuations in the "string" library. The punctuation is as follows:

```
'!"#$%&\'()*+,-./:;<=>?@[\\]^_`{|}~'
```

We made a function to remove the punctuation from our dataset and applied it to the dataset:

1.2) **Stop words:** Stop words are common terms that appear frequently in a text but do not contribute to the machine's comprehension of the content. These words should not exist in our data. As a result, these words are removed. All of these stop words are stored in several languages in the ntlk library.

1.3) **Tokenization:** Tokenization is the process of dividing text into a set of meaningful pieces. These pieces are called tokens. For example, we can divide a chunk of text into words, or we can divide it into sentences. Depending on the task at hand, we can define our own conditions to divide the input text into meaningful tokens.

1.4) **Feature Extraction:** Feature Extraction is a technique for reducing the number of features in a dataset by generating new ones from existing ones (and then discarding the original features) There are Two vectorizers which can be used to turn this text data into numerical data.

1.4.1) Count Vectorizer: To use textual data for predictive modelling, the text must first be parsed to remove specific terms, a process known as tokenization. These words must then be converted to integers or floating-point numbers in order to be used as inputs in machine learning methods. Feature extraction is the term for this procedure (or vectorization).

1.4.2) TF-IDF Vectorizer: Term Frequency — Inverse Document Frequency is abbreviated as TF-IDF. It is one of the most essential approaches for representing how relevant a certain word or phrase is to a given document in terms of information retrieval**.**

From the above two feature extraction technique we have used Count Vectorizer which helped us to get the highest accuracy.

1.5) **Modeling :**

The Text classification is a machine learning technique that assigns a set of predefined categories to open-ended text. Text classifiers can be used to organize, structure, and categorize pretty much any kind of text – from documents, medical studies and files, and all over the web.

For example, new articles can be organized by topics; support tickets can be organized by urgency; chat conversations can be organized by language; brand mentions can be organized by sentiment; and so on.

To Decide which classification model will be best for our problem First we split the data into train and test and then train the model to predict how accurate our model is

We have implemented eight models i.e., algorithms which are as follows:

1.5.1) Multinomial Naïve Bayes

1.5.2) Passive Aggressive Classifier

1.5.3) Logistic Regression

1.5.4) Decision Tree

1.5.5) Random Forest

1.5.6) XG Boost

1.5.7) Support Vector Machine

1.5.8) SGD

B. PERFORMANCE

The comparison table is stated below which shows the classification algorithms of all the three datasets.[1] Harsh Dedhia worked on LIARS dataset [2] Garima Mahto has worked on Fake News Source dataset [3] Jyoti has worked on ISOT Fake news dataset. Together we all worked on each dataset and obtained the following comparison.

DATASET	MODEL	ACC	True/False	PRECISION	RECALL	F-SCORE
LIARS	MNB	0.59	0/1	0.64/0.51	0.65/0.50	0.65/0.51
	PAC	0.55	0/1	0.60/0.46	0.62/0.45	0.61/0.46
	LR	0.59	0/1	0.63/0.51	0.69/0.45	0.66/0.48
	DT	0.59	0/1	0.59/0.63	0.97/0.07	0.73/0.13
	KNN	0.50	0/1	0.62/0.44	0.32/0.73	0.42/0.55
	RF	0.59	0/1	0.60/0.56	0.88/0.20	0.71/0.30
	XGBOOT	0.59	0/1	0.61/0.53	0.81/0.30	0.70/0.38
	SGD	0.54	0/1	0.61/0.46	0.59/0.48	0.60/0.47

DATASET	MODEL	ACC	True/False	PRECISION	RECALL	F-SCORE
	MNB	0.92	0/1	0.93/0.90	0.94/0.89	0.94/0.90
	PAC	0.94	0/1	0.94/0.96	0.98/0.89	0.96/0.92
	LR	1.00	0/1	1.00/1.00	1.00/1.00	1.00/1.00
FakeNews Source	DT	1.00	0/1	1.00/1.00	1.00/1.00	1.00/1.00
	GBC	1.00	0/1	1.00/1.00	1.00/1.00	1.00/1.00
	RF	1.00	0/1	1.00/1.00	1.00/1.00	1.00/1.00
	SVM	1.00	0/1	1.00/1.00	1.00/1.00	1.00/1.00
	XGBOOT	1.00	0/1	1.00/1.00	1.00/1.00	1.00/1.00
	SGD	1.00	0/1	1.00/1.00	1.00/1.00	1.00/1.00

DATASET	MODEL	ACC	True/False	PRECISION	RECALL	F-SCORE
ISOT Fake News	MNB	0.92	0/1	0.93/0.94	0.94/0.91	0.93/0.92
	PAC	0.93	0/1	0.92/0.93	0.95/0.91	0.94/0.93
	LR	0.95	0/1	0.94/0.96	0.94/0.95	0.95/0.95
	DT	0.95	0/1	1.00/1.00	1.00/1.00	1.00/1.00
	GBC	0.89	0/1	0.84/0.96	0.89/0.96	0.89/0.90
	RF	0.94	0/1	0.95/0.96	0.95/0.95	0.950.95
	SVM	0.96	0/1	0.97/0.95	0.96/0.97	0.96/0.96
	XGBOOT	0.89	0/1	0.96/0.85	0.85/0.96	0.90/0.90
	SGD	0.95	0/1	0.96/0.94	0.95/0.96	0.95/0.95

Fake News Source	100%
LIARS	59%
ISOT Fake news	95%

Here, the red colour marked depicts the best suite algorithm while the blue colour indicates the lower performance of the algorithm.

From the above classification Algorithms **Decision Tree** performed the best on the train set and gave us the best accuracy for all the datasets

C. MANUAL TESTING

Manual Testing is something that lets the researchers work on the model from backend in order to test the working of model from the dataset obtained.

In our model we tried to incorporate manual testing of data to confirm that our trained model is able to classify/ detect the truthiness of the news

.

```
When did the decline of coal start? It started when
dministration.


MNB Prediction: It's Fake News
PAC Prediction: It's Fake News
LR Prediction: It's Fake News
DT Prediction: It's Real News
RF Prediction: It's Real News
```

D. CROSS VALIDATION

We proposed that existing ensemble techniques be combined with textual features as feature input to improve overall accuracy when determining whether an item is true or untrue. Because more than one model is trained using a particular technique to minimise the total error rate and increase the model's performance, ensemble learners have higher accuracies. The rationale underpinning ensemble modelling is similar to that which we are already used to in our daily lives, such as obtaining numerous expert opinions before making a decision in order to reduce the risk of making a bad judgement or having an unfavourable outcome. A classification algorithm, for example, can be trained on a specific dataset with a specific set of parameters to produce a decision boundary that roughly fits the data. The output of that algorithm is influenced not only by the parameters used to train the model, but also by the type of training data used. The model may overfit and generate biased findings over unseen data if the training data contains less variance or homogenous data. As a result, methods like as cross validation are utilised to reduce the danger of overfitting. To establish numerous decision limits using randomly picked data points as training data, a variety of models can be trained on different sets of parameters. As a result, these issues can be handled and mitigated utilising ensemble learning approaches by training many algorithms and combining their results for a near-optimal output.

E. MULTILABELING CLASSIFICATION

The rapid dissemination of disinformation is becoming a rising global problem, as it has the potential to have a significant impact on individual reputation and society behaviour. The implications of unregulated disinformation spread can range from political to economical, and can have a long-term impact on world opinion. As a result, spotting fake news is crucial but difficult, as even humans have limited ability to effectively categorise specific information as real or false. Furthermore, fake news is a mix of true and fraudulent information, making proper categorization much more difficult. We propose a unique multilabel multiclass false news detection approach based on relabelling the dataset and learning iteratively in this research. The suggested method beats the benchmark, and our results show that the source of information's profile plays the most important role in detecting news.

IV. BACKGROUND WORK

The present state of knowledge, as well as substantive results, theoretical and methodological contributions to this subject, are all included in a literature review [1] F. Monti, F. Fresca, D. Eynard, D. Mannion, M. M. Bronstein described geometric deep learning and obtained results as propagation-content-based approaches. Here, they separated the real and fake stories content to train their model. The next research paper involved [2] J. Zhang, B. Dong, S. Y. Philip, Fake detector showcase the detection of fake news on news articles, their subjects and creators. Their content was based on hybrid feature extraction process then the build on deep diffusive model as GDU for multiple inputs. [3] Adelani DI, Mai H, Fang F, Nguyen HH, Yamagishi J, Echizen I evaluated the paper description as in sentiment preserving online fake reviews, they worked on neural language model and performed BERT classifier to get best accuracy for their paper. They worked on online review. [4] An Empirical study on Pretrained Embeddings and Language Models for bot detection stated that this paper used a classification task to validate whether the improvement that transfer learning approaches based on fine-tuning pre-trained language models have brought to NLP tasks can be also achieved with social media text. The challenge for these models is that they have been learned from corpora like Wikipedia, News, or Books, where text is well written, grammatically correct and contextualized.[5] Pre-training of deep bidirectional transformers for language understanding the author J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert argued that current techniques restrict the power of the pre-trained representations, especially for the fine-tuning approaches. The major limitation is that standard language models are unidirectional, and this limits the choice of architectures that can be used during pre-training. Their major contribution was further generalizing these findings to deep bidirectional architectures, allowing the same pre-trained model to successfully tackle a broad set of NLP tasks. [6] Yang wang argued that when interpersonal organizations advance, fake news for various purposes will emerge. In recent years, there has been an increase in the number of people using the internet for business and political purposes. huge numbers and have been gotten across the board in the web world Individuals might be corrupted by deceptive language in all ways by false news effectively and will share They haven't done any reality checks on them.

V. PROPOSED MODEL

Throughout the whole journey, our team finally decided to work on such a model which will in future predict the news as fake or deal depending on the F1- score obtained from the different dataset best suite classification algorithm.

This F1-score basically will help in understanding the level of percentage obtained while training the model. Thus, based on percentage we would be able to decide how much the news is fake or real?

Once the truth is found about the news our model will redirect the user to the new page in the window where the user can look onto the news that's been viral on social media, and which they should not intend to forward it to other unless and until it is verified that it is real or fake by reading the whole article in detail that they had searched for it.

VI. Conclusion

Continuous increase in the fake news made the things go in wrong direction which led to increase in misinformation in people's mind. This is something that to be controlled, it can't be stopped at all, but preventing them to get spread that can help out as a solution. In short, our model will try to keep aware of such fake news and help the people to know the truth of news, so that they can be aware of false spreading of such news articles. Our whole team tried best to train such a model, which will be helpful in future generation.

.

References

[1] F. Monti, F. Fresca, D. Eynard, D. Mannion, M. M. Bronstein, Fake news detection on social media using geometric deep learning, arXiv preprint arXiv:1902.06673.

[2] J. Zhang, B. Dong, S. Y. Philip, Fakedetector: Effective fake news detection with deep diffusive neural network, in: 2020 IEEE 36th International Conference on Data Engineering (ICDE), IEEE, 2020, pp. 1826–1829.

[3] Adelani DI, Mai H, Fang F, Nguyen HH, Yamagishi J, Echizen I. Generating sentiment-preserving fake online reviews using neural language models and their human-and machine-based detection. In: International Conference on Advanced Information Networking and Applications. Springer; 2019. p. 1341–1354.

[4] Garcia-Silva A, Berrio C, Gómez-Pérez JM. An Empirical Study on Pre-trained Embeddings and Language Models for Bot Detection. In: Proceedings of the 4th Workshop on Representation Learning for NLP (RepL4NLP-2019). Association for Computational Linguistics; 2019. p. 148–155.

[5] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understading, arXiv preprint arXiv:1810.04805.

[6] William Yang Wang. 2017. "liar, liar pants on fire": A new benchmark dataset for fake news detection. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pages 422–426. Association for Computational Linguistics.