

Fake News Detection Using A Deep Neural Network

Rohit Kumar Kaliyar
Computer Science & Engineering
Bennett University
Greater Noida, India
rk5370@bennett.edu.in

Abstract—The process of obtaining news from social media is like double edged weapon. On one hand, it is easy to access, less time consuming, user friendly, easily conveyable socially relevant news, possibility for obtaining various perspective of a single news and is being updated in every minute. On other hand, news is being manipulated by various networking sites based on private opinions or interest. Fake news is misinformation or manipulated news that is spread across the social media with an intention to damage a person, agency and organization. Due to the dissemination of fake news, there is need for computational methods to detect them. Fake news detection aims to help users to expose varieties of fabricated news. We can decide whether the news is solid or forged based on formerly witnessed fake or real news. We can use various models to access deceptive news in social media. Our contribution is bifold. First, we must introduce the datasets which contain both fake and real news and conduct various experiments to organize fake news detector. We use Natural Language Processing, Machine learning and deep learning techniques to classify the datasets. We yield a comprehensive audit of detecting fake news by including fake news categorization, existing algorithms from machine learning techniques.

Keywords—*Fake News, Opinion Mining, Machine Learning, Natural Language Processing*

I. INTRODUCTION

Counterfeit news or fake news is a bit of false data created for business enthusiasm to pick up consideration and produce promotion income or to spread scorn related violations to impact the world politically. News articles that imply to be truthful, however which contain purposeful misquotes of reality with the expectation to excite interests, draw in viewership, or cheat. As of late, there have been numerous examples of unsubstantiated or false data spreading quickly finished online informal organizations. For instance, there were ongoing reports about Russian hacking of an electrical matrix in Vermont and reports specifying that Emmanuel Macron's presidential battle is financed by Saudi Arabia. Such unconfirmed news has been spreading at a quick pace as of late and with the development of "enormous information" in these fields it is difficult to physically channel such news.

Even though the intricacy of imaginary news is not another debate, uncovering misleading news is accepted to be a confused obligation inured that people tend to acknowledge beguiling guideline and the nonappearance of control of the transmission of phony substance. Recognizing the validity and accurate rightness of a news article is a fascinating issue and can be actualized utilizing different philosophies.

II. OBJECTIVE AND MOTIVATION

The Overall Objective of the project 'Fake news detection' is to classify the news article or other documents into certain or not. We explore identification of fake news using various models and classifiers and predict the accuracy of different models and classifiers. Through this project, we examine which model will give more accuracy and classify the news into real or fake. In this project, we generate computational resources and models for the duty of fake news detection. We propose a dataset for Kaggle which combination of real and fake news is and implement that dataset using various technologies like machine learning, natural language processing and deep learning. Using these dataset, we handled diverse exploratory analyses to identify linguistic properties that are broadly present in deceptive content. To put our outcomes in perspective, we likewise associate the precision of our phony news recognition models with already anticipated exactness.

We use Natural Language Processing, machine learning and deep learning techniques to implement our models and compare which models will give more accuracy. We even make use of Nvidia DGX-1 supercomputer for easy analysis of results and examine which model will precisely classify the given dataset into real and fake news.

III. RELEVANCE

Fake news or forged news, refers to false instruction advertised under the disguise of being trustworthy news. According to New York Times 64% of adults in US said, "made up news" has always put them in confusion. The procreation of ambiguous Information in daily access media articles such as news feeds, blogs online newspaper, have been made challenging to identify realistic news sources, thus expanding the demand of computational tools which should be able to provide the vision into the authenticity of online content. The term 'Fake news' become relevant after US Election 2016. During the election, fake news was both widely shared and heavily tilted in favor of Donald Trump. Thereafter Fake news detection become emerging research area which is boosting interest to every people.

Since it is socially significant research area and by discovering methodologies to determine genuine news will provide appropriate conveying of knowledge to all people, thus we select this topic as our project. With the available resources (dataset, literature papers), we try to find the best model in machine learning and deep learning techniques that provide high accuracy in predicting and classifying the given combination set of real and fake news dataset into factual and wrong.

IV. LITERATURE SURVEY

Through this section, we summarize some of the existing research works in the field of Machine learning/deep learning to analyses about Fake News Detection and build a model according to the existing applications.

Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, 52(1), 1-4.

- In this paper, the author mainly focusses on categorizing the news based on finding the degree of accuracy or correctness in the news.
- They focus on technologies adopted for fake news detection.
- Include mainly two categories for assessment-linguistic cue approach (with machine learning) and network analysis approach.
- Both approaches adopt machine learning techniques for training classifiers to suit the analysis.
- Future scope: -combination of linguistic cue and machine learning on network based behavioral data.
- Focus on structured dataset such as text
- These papers show the current development of correctness assessment methods, their goals and classes with the aim to propose new hybrid system for detection.

Wu, Liang, and Huan Liu. "Tracing Fake-News Footprints: Characterizing Social Media Messages by How They Propagate." (2018).

- The author focusses on classification of social media content for social media mining and text categorization problem mainly including text with hashtags and words.
- Through this paper, author focus on modelling the propagation of messages in social network.
- Here the author proposes an approach called TRACEMINER to infer the embedding of social media users in social network structure and to create a LSTM-RNN model to represent the path of messages.
- Using of trace miner approach, provide high classification accuracy and is better for interpreting real world dataset than traditional approaches.
- In Trace miner approach, they will take many messages as input and category as output. This approach is different from traditional approach as they directly concentrate in modelling the information and making predictions.
- They provide optimization methods to trace miner to guarantee the correctness and evaluate the performance of the real world social network data.

Granik, Mykhailo, and Volodymyr Mesyura. "Fake news detection using naive Bayes classifier." *Electrical*

and Computer Engineering (UKRCON), 2017 IEEE First Ukraine Conference on. IEEE, 2017.

- The author focusses on implementation of Fake News Detection using Naïve Bayes Classifiers. They tested the classifier against new post of Facebook dataset and obtain accuracy of about 74%.
- As the providence of accessing news information were given to users by social media and internet, more manipulation on news occurs. Basically, the manipulation occurs due to public opinion on certain matters.
- Most of the people believe that the problem of fake news can be solved by machine learning or artificial intelligence.
- By referring thorough the following paper, the author implemented the model – (Rada Mihalko, Carlo Strapparava, The lie detector: explorations in the automatic recognition of deceptive language, *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, August 04-04, 2009, Suntec, Singapore)
- <https://dl.acm.org/citation.cfm?id=1667679>
- Through this implementation, author focus on applying dataset on new datasets so that they focus on recent data.

Buntain, Cody, and Jennifer Golbeck. "Automatically Identifying Fake News in Popular Twitter Threads." *Smart Cloud (Smart Cloud), 2017 IEEE International Conference on. IEEE, 2017.*

- The objective of this paper was to automate fake news detection on twitter data by conducting accuracy assessment on two credible twitter datasets – CRED BANK and PHEME
- They apply this strategy to Twitter content sourced from Buzz Feed's fake news dataset and indicate models prepared against publicly supported specialists outflank models dependent on evaluation and models prepared on a pooled dataset. Each of the three datasets is adjusted into a uniform arrangement and highlight investigation is performed.
- The paper passes on around an automated system for arranging well known Twitter threads in a computerized component for characterizing prominent Twitter threads into genuine and counterfeit news stories, an investigation of the diverse highlights utilized by various individuals, for example, laborers, columnist in surveying precision in internet-based life stories, and an adjusted accumulation of three datasets that give exactness judgements crosswise over evident and false stories.

Internet and Social media provides facilities to the users to access news information and this can be a double-edged sword with both pros and cons, every day huge massive amount of news is generated and manipulated in social network media which can be used to analyze whether it is real or fake. We generally use machine learning and deep learning models to analyze the problem.

Shu, Kai, et al. "Fake news detection on social media: A data mining perspective." ACM SIGKDD Explorations Newsletter 19.1 (2017): 22-36

- Problem definition- Fake news detection on social media present challenges and characteristics faced by the existing detecting algorithms on traditional approaches.
- In this paper, they present a survey of distinguishing counterfeit news via web-based networking media including counterfeit news portrayals, existing calculations from an information mining viewpoint, assessment measurements and delegate datasets.
- They focus on explaining the benefits of accessing news information on social media and conveys that quality of news is less than traditional approaches. They clarify what fake news and their qualities are. They give a diagram of existing fake news identification techniques by gathering agent strategies into different classes and they talk about a few open issues and give future headings of fake news location in web-based life.

BAJAJ, SAMIR. "“THE POPE HAS A NEW BABY!” FAKE NEWS DETECTION USING DEEP LEARNING.

- The objective of this paper was to build a classifier that can predict whether a piece of news is fake based only on its content using Natural language processing.
- Different model was explored for detecting fake news starting from Logistic Regression to CNN till RNN GRU. The focus of the research was using NLP techniques to detect fake news by exploiting the linguistic features of the fake news and real news.

Bhatt, Gaurav, et al. "Combining Neural, Statistical and External Features for Fake News Stance Identification." Companion of the Web Conference 2018 on the Web Conference 2018. International World Wide Web Conferences Steering Committee, 2018.

- The paper recognizes the veracity of a news article. It investigates a subtask to counterfeit news distinguishing proof, and that is position recognition. Given a news article, the goal is to decide the pertinence of the body and its case.
- They exhibited an original thought that joins the neural, factual and outer highlights to give an effective answer for this issue.
- They process the neural inserting from the profound repetitive model, factual highlights from the weighted n-gram pack of-words display and handmade outer highlights with the assistance of highlight designing.
- Finally, the features are combined, thereby classifying the headline-body news pair as agree, disagree, discuss, or unrelated. Through them

experiments, they found that the proposed model outperforms all the submissions to the fake news challenge.

Abbasi, Ahmed, Fatemeh Zahedi, and Siddharth Kaza. "Detecting fake medical web sites using recursive trust labelling." ACM Transactions on Information Systems (TOIS)30.4 (2012): 22

- This paper discusses the spread and construction of fake medical websites including the ways the to combat these fake sources using the techniques described below
- **RTL:** Recursive Trust labelling (type of adaptive learning algorithm). This technique uses linguistic features and graph-based classifier with recursive labelling mechanisms. The recursive marking component recursively grows the training dataset by choosing extra test examples amid every emphasis. Occasions that have the most grounded prediction scores (in view of the substance and chart classifiers) are chosen and included with class labels that based on the fundamental classifiers' predictions.
- **Content based methods:** use automatic content generation techniques to mass produce fake content which can be recognized by NLP techniques. The recursive naming component recursively grows the training dataset by choosing extra test occasions amid every cycle. Examples that have the most grounded expectation scores (considering the substance and chart classifiers) are chosen and included with class names that depend on the basic classifiers' predictions. Algorithms that are less subject to bigger training datasets are exceptionally attractive.
- Features used: **HTML** tags, URL text, image hashes, link information
- Accuracy: 68%
- **Neural Network, SVM, Naive Bayes:**
- Feature: Body text lexical measures and n-grams, anchor text
- Accuracy: **96%**
- Graph-based techniques designed to combat link farms rely on the assumption that good News link to good News, while bad News are likely to link to other bad ones.
- Rumor Gauge: Predicting the Veracity of Rumors on Twitter
- Type of model used: Hidden Markov Models (HMMs)
- Datasets extraction: Twitter historical API
- **Idea on which article is based:** even before trusted verification, there are signals, however weak, that are predictive of the veracity of rumors.

The approach utilized: used to prepare two HMMs, one for genuine and one for false bits of gossip. It is important to prepare two separate HMMs for false and genuine bits of gossip since it is our theory that the fleeting elements of our highlights contrast among false and genuine true rumors.

V. IMPLEMENTATION & METHODOLOGY

A. Platform and Technologies

Jupyter Notebook

- The IPython scratch pad is otherwise called Jupyter note pad.
- It is an open source aggregate computational condition which consolidate codes, science, plots and content.
- It is a user-friendly platform.
- We upload dataset and python files in notebook and implement codes of machine learning/deep learning models. We get result in form of confusion matrix or graph.

Nvidia DGX

- Nvidia DGX build a platform for deep learning.
- Mostly they work by selecting a server and GPUs.
- To implement AI and Machine learning/deep learning models in our project commit carefully recruiting and incorporating complex software with hardware.
- NVIDIA DGX tracks out our initiative with a solution that works correct out of the box, so that we can implement codes in hours than in weeks.

B. Dataset and collection

- A dataset is a collection of data which mainly comprises of single statistical data matrix, database table where every row corresponds to each member in datasets and each column represents variable.
- The dataset list values for each variable such as title, id, author, label etc.
- We commonly collect various datasets from <https://www.kaggle.com>.
- Exclusively fake news article datasets contain news article are forged.
<https://www.kaggle.com/mrisdal/fake-news>
- Exclusively real news Data: - This dataset contains news articles which are certain.
<http://research.signalmedia.co/newsir16/signal-dataset.html>
- Mixed Data of fake and real news: - his dataset contains the combination of real and fake news where (row*column) is (2080 *5)
<https://www.kaggle.com/c/fakenews/data>

- This dataset also contains combination of real and fake news where (row*column) is (6335*3)

C. Features and Factors

• TERM FREQUENCY: -

Term frequency is the type of path that appropriate the counts of words present in the document and figure out the inequality between the documents. It is the statistical measured term used to measure the document from the data sets. Each document is characterized in a vector that contains the word count. Each word count is transformed into the possibility of such word current in the documents.

• TERM FREQUENCY-INVERTED DOCUMENT FREQUENCY: -

The Term frequency-inverted document frequency is type of filling metric used in recapturing and natural processing. It is the statistical metric term used to regulate the document from the data sets. The term consequence increases with the number of times a word appears in the document. The important feature if IDF is it weighs down the term frequency while mounting up the occasional ones.

• CONFUSION MATRIX: -

Confusion matrix is a type of statistical classification also known as error matrix. It grants visualization and usually called as matching matrix. Each row of matrix represents immediate prediction and column represent instances of a real class.

• HASING VECTORIZER: -

Other type used in text classification is Hashing vectorizer crave a less memory and faster because they use Hashes rather than tokens. It gives an improved accuracy results than the TF-IDF vectorizer using Multinomial Naïve Bayes.

• PRECISION: -

Precision also called positive predicted value is the fraction of significant instances among the retrieved instances.

• ACCURACY: -

The capacity of the framework to precisely characterize information depends to a vast degree on the illustrations that you give.

• RECALL: -

Recall also known as sensitivity is the fraction of significant instances that have been retrieved over the total amount of relevant instances.

• F1-SCORE: -

F₁ score also **F-score** or **F-measure** is a measure of a test's accuracy for binary classification.

D. Methodologies

Naïve Bayes Model

- It uses probabilistic approaches and are based on Bayes theorem. They deal with probability distribution of variables in the dataset and predicting the response variable of value
- They are generally used for text classification and used in medical diagnosis.
- Bayes theorem
 $P(a|b) = p(b|a)p(b)/p(a)$
- There are mainly 3 types of naïve base models. Gaussian Naïve Bayes, Multinomial naïve Bayes and Bernoulli Naïve Bayes. We have used Multinomial Naïve Bayes model for our project to detect fake news.
- Naive Bayes classifier is the straightforward procedure of creating classifier models that select class names to issue example, communicated has a vector of highlight esteems, where the class names are drawn from some limited set.
- Naive Bayes classifier model have worked well in many complicated real-world situations.
- An advantage of naïve Bayes classifier is that only requires less bulk of training data to access the parameters necessary for classification.

Decision Tree:

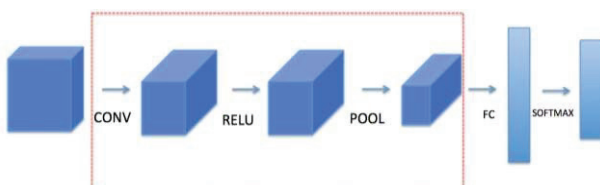
- It is used to optically serve decisions and decision making.
- It uses tree like model of both classification and regression.
- They are commonly used for data mining that support machine learning.

Random forest

- It is a combination of decision trees. Here each tree will build a random subset of a training dataset.
- In each decision tree model, a random subset of variables is used to partition the data set at each node.

K nearest neighbor

- They store the entire data set for the implementation. It is an instant based learning which is done by approximation.
- Their data values are arranged in a feature space. They depend on the value of 'k'.
- The data value or the feature unknown to us is found out by using the value of 'k'. That is, the nearest 'k' neighbors is taken and most occurring feature is observed.



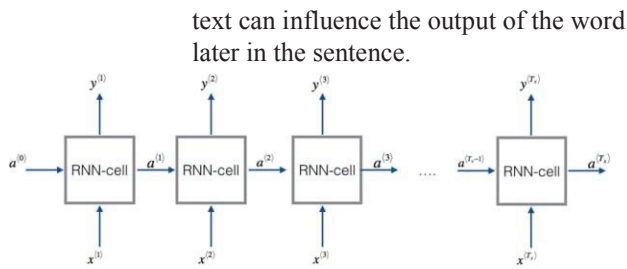
- CNN has been successful in various text classification tasks. In “Kim Y. Convolutional Neural Networks for Sentence Classification. 2014”, Kim showed that a simple CNN with some hyper parameter tuning and static vectors like Glove vector and word2vec CNN can achieve excellent results on text classification.
- The use of word vectors allows CNN to use its state of art parameter sharing for reduction in calculation and govern the overall text features.

Main functions in CNN are:

- **ReLU:** stands for Rectified Linear Unit for a non-linear operation. The output is $f(x) = \max(0, x)$. REL U's use is to introduce non-linearity in our CNN.
- **Pooling:** The pooling layer reduces the height and width of the input to the layer. It helps reduce computation, as well as helps make feature detectors more invariant to its position in the input
- **Padding:** It allows you to use a CONV layer without necessarily shrinking the height and width of the volumes. This is important for building deeper networks, since otherwise the height/width would shrink as you go to deeper layers.

LSTM

- ❖ Traditional neural networks cannot remember or keep the record of what all is passed before they are executed this stops the desired influence of words that comes in the sentence before to have any influence on the ending words, and it seems like a major shortcoming. For example, imagine you want to classify what kind of event is happening at every point in a movie. It's unclear how a traditional neural network could use its reasoning about previous events in the film to inform later ones.
- ❖ LSTM address this issue. They are networks with loops in them, allowing information to persist.
- ❖ Long short-term memory (LSTM) units are a building blocks for the layers of a recurrent neural network (RNN). A LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell is responsible for "remembering" values over a vast time interval so that the relation of the word in the starting of the



VI. RESULTS

CNN & LSTM

- We use a shallow CNN based model the results are shown below, accuracy of the model after 2 epochs is 91.3%
- WE then increased the depth of the network by a few layers to check whether the performance is increased or not. The accuracy of the model was increased to 98.3%.
- We also implemented a combination of CNN and LSTM. Accuracy was reduced by a bit to 97.3% but precision and recall was effectively improved.

Naïve Bayes model

- We use fake_or_real_news dataset for implementation.
- By implementing this model on my data set, we've got an accuracy of 89%
- **Precision, Recall, f1-Score** got the same value (90%).
- it is shown below: Precision recall f1-score support

Type of News	Precision	recall	F1-score
FAKE	0.90964	0.90	0.90
REAL	0.90171	0.90	0.90
Total	0.901935	0.90	0.90

Decision Tree

We use train_mixed.csv data set file in naïve Bayes for this model. Decision Tree Regressor is used to classify the data set into smaller subsets and at the same time a decision tree is incrementally developed.

- By implementing this model on my data set, we've got an accuracy of (73%)
- (Precision=75%), (Recall=74%), (f1-Score=73%)
- It is shown below:

Type of News	Precision	recall	F1-score
FAKE	0.82	0.60	0.70
REAL	0.69	0.87	0.77
Total	0.75	0.74	0.73

Random Forest

- Here also we use train_mixed.csv dataset for the Implementation.
- Random Forest Regressor is used to classify the data set by constructing multiple number of Decision trees at training time and obtaining mode of classes as outputs.
- By Implementing this model on my data set, we've got an accuracy of 71%
- (Precision=72%), (Recall=71%), (f1-Score=71%)
- It is shown below:

Type of News	Precision	recall	F1-score
FAKE	0.74	0.65	0.70
REAL	0.69	0.69	0.77
Total	0.72	0.71	0.71

K Nearest Neighbors

- Here also we use train_mixed.csv dataset for the Implementation.
- K-Neighbors classifier is used to classify the data set.
- By Implementing this model on my data set, we've got an accuracy of 53%
- (Precision=.54%), (Recall=53%), (f1-Score=50%).

Type of News	Precision	recall	F1-score
FAKE	0.59	0.27	0.37
REAL	0.52	0.81	0.63
Total	0.55	0.54	0.50

VI. CONCLUSION

Fake News Detection is the analysis of socially relevant data to distinguish whether it is real or fake. In this project we explored different Machine learning models like Naïve Bayes, K nearest neighbors, Decision tree, Random forest and Deep Learning networks like Shallow Convolutional Neural Networks (CNN), Very Deep Convolutional Neural Network (VDCNN), Long Short-Term Memory Network (LSTM), Gated Recurrent Unit Network (GRU), Combination of Convolutional Neural Network with Long Short-Term Memory (CNN-LSTM) and Convolutional Neural Network with Gated Recurrent Unit (CNN-LSTM). We also explored the benefit of feature extraction, features like n-gram, TF-IDF features were extracted and used in our model. We also explored the effective of word embedding's and word2vec features in Deep Neural networks. We also explored the use of select best and chi2 for feature extraction in Machine learning model.

REFERENCES

- [1] Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, 52(1), 1-4.
- [2] Wu, Liang, and Huan Liu. "Tracing Fake-News Footprints: Characterizing Social Media Messages by How They Propagate." (2018).
- [3] Granik, Mykhailo, and Volodymyr Mesyura. "Fake news detection using naive Bayes classifier." *Electrical and Computer Engineering (UKRCON), 2017 IEEE First Ukraine Conference on*. IEEE, 2017.
- [4] Buntain, Cody, and Jennifer Golbeck. "Automatically Identifying Fake News in Popular Twitter Threads." *Smart Cloud (Smart Cloud), 2017 IEEE International Conference on*. IEEE, 2017.
- [5] Shu, Kai, et al. "Fake news detection on social media: A data mining perspective." *ACM SIGKDD Explorations Newsletter* 19.1 (2017): 22-36.
- [6] Bhatt, Gaurav, et al. "Combining Neural, Statistical and External Features for Fake News Stance Identification." *Companion of the Web Conference 2018 on the Web Conference 2018*. International World Wide Web Conferences Steering Committee, 2018.
- [7] Abbasi, Ahmed, Fatemeh Zahedi, and Siddharth Kaza. "Detecting fake medical web sites using recursive trust labelling." *ACM Transactions on Information Systems (TOIS)* 30.4 (2012): 22.