

Fake News Detection

A Social Internship Report

Submitted in the partial fulfillment of the requirements for the
award of the degree of

Bachelor of Technology

in

Department of Computer Science and Engineering

By

E. Pravallika (2010030046)

Tahseen Begum (2010030168)

N. Sowgna (2010030344)

P.Keerthana (2010030445)

under the supervision of

Mr. Giridhar Urkude



Department of Computer Science and Engineering

K L University Hyderabad,

Aziz Nagar, Moinabad Road, Hyderabad – 500 075, Telangana, India.

March, 20

Declaration

The Social Internship Report entitled “Fake News Detection” is a record of bonafide work of E.Pravallika (2010030046), Tahseen Begum (2010030168), N.Sowgna (2010030344), P.Keerthana (2010030445) submitted in partial fulfillment for the award of B.Tech in the Department of Computer Science and Engineering to the K L University, Hyderabad. The results embodied in this report have not been copied from any other Departments/universities/institutes.

Signature of the Students

Certificate

This is to certify that the Social Internship Report entitled “**Fake News Detection**” is being submitted by **E.Pravallika (2010030046)**, **Tahseen Begum(2010030168)**, **N. Sowgna (2010030344)**, **P.Keerthana (2010030445)** submitted in partial fulfillment for the award of B.Tech in **Mr.Giridhar Urkude(Asst. Professor)** to the K L University, Hyderabad is a record of bonafide work carried out under our guidance and supervision.

The results embodied in this report have not been copied from any other departments/universities/institutes.

Signature of the Supervisor

Signature of the HOD

Signature of the External Examin

ACKNOWLEDGEMENTS

It is great pleasure for me to express my gratitude to our honorable President **Sri. Koneru Satyanarayana**, for giving the opportunity and platform with facilities in accomplishing the project-based laboratory report.

I express my sincere gratitude to our Principal **Dr. L. Koteswara Rao** for his administration of our academic growth.

I express sincere gratitude to our Coordinator for her leadership and constant motivation provided in the successful completion of our academic semester. I record it as my privilege to deeply thank you for providing us with the efficient faculty and facilities to make our ideas into reality.

I express my sincere thanks to our project **supervisor Mr.Giridhar Urkude(Asst. Professor)** for her novel association of ideas, encouragement, appreciation, and intellectual zeal which motivated us to venture into this project successfully. Finally, it is pleased to acknowledge the indebtedness to all those who devoted themselves directly or indirectly to making this project report success.

ABSTRACT

Today the world relies upon electronic media for its everyday adventure. People have no time to be updated through Newspapers or watch or listen to the news on television or radios. The WWW (World Wide Web) is a huge, widely distributed, global information service center for Information services: news, advertisements, consumer information, financial management, education, government, e-commerce, etc. "24 Hours News Portal" is a service introduced to meet the above requirement and to make the people updated about the news, views, reviews, breaking news and latest headlines in different fields also the new inventions around the world.

In our modern era where the internet is ubiquitous, everyone relies on various online resources for news. Along with the increase in the use of social media platforms like Facebook, Twitter, etc. news spread rapidly among millions of users within a very short period. The spread of fake news has far-reaching consequences like the creation of biased opinions to sway election outcomes for the benefit of certain candidates. Moreover, spammers use appealing news headlines to generate revenue using advertisements via click-baits. In this paper, we aim to perform binary classification of various news articles available online with the help of concepts about Artificial Intelligence, Natural Language Processing, and Machine Learning. We aim to provide the user with the ability to classify the news as fake or real and also check the authenticity of the website publishing the news.

TABLE OF CONTENTS

S.NO	TITLE	PAGE NO
1	Introduction	03
2	Literature Survey	04
3	Methodology	05
4	Implementation	06-08
5	Results Discussion	09-10
6	Conclusion and Future Work	11
7	Social Contribution	12
8	References	13

INTRODUCTION

As an increasing amount of our lives is spent interacting online through social media platforms, more and more people tend to hunt out and consume news from social media instead of traditional news organizations. The explanations for this alteration in consumption behaviors are inherent within the nature of those social media platforms: (i) it's often more timely and less expensive to consume news on social media compared with traditional journalism, like newspapers or television; and (ii) it's easier to further share, discuss, and discuss the news with friends or other readers on social media. For instance, 62 percent of U.S. adults get news on social media in 2016, while in 2012; only 49 percent reported seeing news on social media. It had been also found that social media now outperforms television because a major news source. Despite the benefits provided by social media, the standard of stories on social media is less than traditional news organizations. However, because it's inexpensive to supply news online and far faster and easier to propagate through social media, large volumes of faux news, i.e., those news articles with intentionally false information, are produced online for a spread of purposes, like financial and political gain. It had been estimated that over 1 million tweets are associated with fake news "Pizzagate" by the top of the presidential election. Given the prevalence of this new phenomenon, "Fake news" was even named the word of the year by the Macquarie Dictionary in 2016. The extensive spread of faux news can have a significant negative impact on individuals and society. First, fake news can shatter the authenticity equilibrium of the news ecosystem for instance; it's evident that the most popular fake news was even more outspread on Facebook than the most accepted genuine mainstream news during the U.S. 2016 presidential election. Second, fake news intentionally persuades consumers to simply accept biased or false beliefs. Fake news is typically manipulated by propagandists to convey political messages or influence, for instance, some report shows that Russia has created fake accounts and social bots to spread false stories. Third, fake news changes the way people interpret and answer real news, for instance, some fake news was just created to trigger people's distrust and make them confused; impeding their abilities to differentiate what's true from what's not. To assist mitigate the negative effects caused by fake news (both to profit the general public and therefore the news ecosystem). We must build up methods to automatically detect fake news broadcasts on social media. Internet and social media have made access to news information much easier and more comfortable.

LITERATURE SURVEY

s/n o	Title	Author	Publishi ng	Characteristics	Pros	Cons
1.	Deep Learning and Natural Language Processing for fake news detection: A Survey	Aditya Chokshi, Rejo Mathew	2021	Fraudulent or false news may have a tendency to create destructive surrounding or mislead an individual or organization to do some unwanted threatening activity and harm the environment. So it is necessary to prevent its distribution. This paper describes several architectures and methodologies that can be employed in this field of fake news classification technique. Since artificial neural networks does not deal with long term dependencies it is necessary to involve the presence of LSTMs in the hidden layer.	Using long short term memory (LSTM) in the convolution neural network (CNN) can lead to the model yielding high accuracy along with using the drop out layers in hidden layer or the fully connected layer in the CNN also results in good predictions for the news content. Also using CNN for image classification of the news spreading on social media with LSTM can deliver good results. Also using bidirectional recurrent neural network with LSTMs as a memory element for handling the dependencies of large input and can be used as a model for prediction of the fabricated news and further soft max can be used as a activation function can be used for predicting the probabilities of how much the news is fake or real.	Latest victim to this field was during the Citizen Amendment act (CAA) in 2020 where a survey by a team of news reporters showed that almost 95% of the protestors didn't knew about that act and were indulged to think that their citizenships would be snatched and were a victim to a fake news.

s/no	Title	Author	Publishing	Characteristics	Pros	Cons
2.	Fake news detection based on statement conflict	Danchen Zhang, Jiawei Xu, Vladimir Zadorozhny & John Grant	Journal of Intelligent Information Systems (2022)	We propose two models: the first one uses only the agree and disagree classifications; the second uses a subjective opinions based model that can also handle the uncertain cases. Our experiments on a real-world dataset (the Fake News Challenge 1 dataset) and a simulated dataset validate that both proposed models achieve state-of-the-art performance.	Using this information, we focus on selecting the news articles that are most likely to be fake.	The immediate procedural success did not differ between SMVR and PMVR (92% vs. 93%, $P = 0.53$). Each patient is unique, and subtle clinical differences may have a high impact on outcome.

METHODOLOGY

The system is a Web application that helps the user to detect fake news. We have given the text box where the user has the option to paste the message or paste the URL link of the news and other message link and after that, it gives the reality of it. All the user gives data to the detector may save for further use to update the status of the model, and data analysis in the future. We also help users by giving some guidance on how to prevent such false events and how to stop such events from spreading.

IMPLEMENTATION

DATA SET

LSTM: Long Short Term Memory

Techniques: Word Embedding: Word embedding is a set of language modeling and feature extraction techniques in Natural Language Processing (NLP). In word embedding, words from vocabulary are converted into vectors of real numbers. Word embedding is a type of word representation that allows words with similar meanings to have a similar representation.

Steps:

```
import pandas as pd
df=pd.read_csv('train.csv')
df.head()
####Drop Nan Values
df=df.dropna()
## Get the Independent Features
df_new=df.drop('PassengerId',axis=1)
## Get the Dependent features
y=df['PassengerId']
df_new.shape
y.shape
import tensorflow as tf
tf.__version__
from tensorflow.keras.layers import Embedding
from tensorflow.keras.preprocessing.sequence import pad_sequences
from tensorflow.keras.models import Sequential
from tensorflow.keras.preprocessing.text import one_hot
from tensorflow.keras.layers import LSTM
from tensorflow.keras.layers import Dense
#### Vocabulary size
voc_size=5000
```

Onehot Representation

```
messages=df_new.copy()
messages['Name'][:1]
messages.reset_index(inplace=True)
import nltk
import re
```

```

from nltk.corpus import stopwords
nltk.download('stopwords')
### Dataset Preprocessing
from nltk.stem.porter import PorterStemmer
ps = PorterStemmer()
corpus = []
for i in range(0, len(messages)):
    print(i)
    review = re.sub('[^a-zA-Z]', '', messages['Name'][i])
    review = review.lower()
    review = review.split()

    review = [ps.stem(word) for word in review if not word in stopwords.words('english')]
    review = ' '.join(review)
    corpus.append(review)

```

Embedding Representation

```

sent_length=20
embedded_docs=pad_sequences(onehot_repr,padding='pre',maxlen=sent_length)
print(embedded_docs)
embedded_docs[0]
## Creating model
embedding_vector_features=40
model=Sequential()
model.add(Embedding(voc_size,embedding_vector_features,input_length=sent_length))
model.add(LSTM(100))
model.add(Dense(1,activation='sigmoid'))
model.compile(loss='binary_crossentropy',optimizer='adam',metrics=['accuracy'])
print(model.summary())
len(embedded_docs),y.shape
import numpy as np
df_new_final=np.array(embedded_docs)
y_final=np.array(y)
df_new_final.shape,y_final.shape
from sklearn.model_selection import train_test_split
df_new_train, df_new_test, y_train, y_test = train_test_split(df_new_final, y_final, test_size=0.33, random_state=42)

```

Model Training

Finally Training

```
model.fit(df_new_train,y_train,validation_data=(df_new_test,y_test),epochs=10,batch_size=64)
```

Adding Dropout

```
from tensorflow.keras.layers import Dropout
```

Creating model

```
embedding_vector_features=40
```

```
model=Sequential()
```

```
model.add(Embedding(voc_size,embedding_vector_features,input_length=sent_length))
```

```
model.add(Dropout(0.3))
```

```
model.add(LSTM(100))
```

```
model.add(Dropout(0.3))
```

```
model.add(Dense(1,activation='sigmoid'))
```

```
model.compile(loss='binary_crossentropy',optimizer='adam',metrics=['accuracy'])
```

Performance Metrics And Accuracy

```
#y_pred=model.predict_classes(df_new_test)
```

```
predict_x=model.predict(df_new_test)
```

```
classes_x=np.argmax(predict_x,axis=1)
```

```
from sklearn.metrics import confusion_matrix
```

```
confusion_matrix(y_test,classes_x)
```

```
from sklearn.metrics import accuracy_score
```

```
accuracy_score(y_test,classes_x)
```

RESULTS DISCUSSION

- Fake news prediction using Machine Learning algorithms and Flask Framework.
- Enter your text or generate a random one from our dataset to try it.
- The text is first preprocessed and transformed as a vector. Then, the transformed vector is fed to the trained model to be classified as fake or not fake.

Web design

The screenshot displays a web application titled "FAKE NEWS" with a dark theme. The navigation bar at the top includes links for "Home", "Dataset", "NoteBook", "Code Source", and two buttons: "Sign Up" and "Login". The main heading is "Fake News - Machine Learning", with "Fake News" highlighted in a red box. Below the heading, a subtitle reads: "Fake news prediction using Machine Learning algorithms. Enter your text or generate a random one from our dataset to try it." A large text input area is provided with the placeholder text "Enter text here..". At the bottom of the input area, there are two buttons: "Generate Text" and "Predict".

How It Works

Fake news prediction using Machine Learning algorithms. The text is first preprocessed and transformed as a vector. Then, the vector is feeded to the trained model to be classified.



Original Text

Bodies with bound hands, close-range gunshot wounds and signs of torture lay scattered in a city on the outskirts of Kyiv after Russian soldiers withdrew from the area. Ukrainian authorities accused the departing forces on Sunday of committing war crimes and leaving behind a "scene from a horror movie." As images of the bodies emerged from Bucha, European leaders condemned the atrocities and called for tougher sanctions against



Preprocessed Text

bodi bound hand close rang gunshot wound sign tortur lay scatter citi outskirt kyiv russian soldier withdrew area ukrainian author accus depart forc sunday commit war crime leav behind scene horror movi imag bodi emerg bucha european leader condemn atroc call tougher sanction moscow sign horrif report sh read http economicitm indiatim com news intern world news ukrain accus russia massacr citi strewn bodi articleshow cm utrn sourc



Transform & Predict

Text predicted as: **FAKE**



Fake News - Machine Learning

Fake news prediction using Machine Learning algorithms. Enter your text or generate a random one from our dataset to try it.

Bodies with bound hands, close-range gunshot wounds and signs of torture lay scattered in a city on the outskirts of Kyiv after Russian soldiers withdrew from the area.

Ukrainian authorities accused the departing forces on Sunday of committing war crimes and leaving behind a "scene from a horror movie."

As images of the bodies emerged from Bucha, European leaders condemned the atrocities and called for tougher sanctions against Moscow. In a sign of how the horrific reports sh...

Read more at:

[https://www.bbc.com/news/health-60888888](#)

Generate Text

Predict

Prediction : **FAKE** [See result details](#)

CONCLUSION & FUTURE WORK

With the help of Machine Learning, we have created 5 prediction model which gives an accuracy above 90% and it covers all the latest political covid 19 news. Also with some pre-trained models, we have covered news related to history and sport. We are intent to build our dataset which will be kept up to date according to the latest news in the future.

This project can be further enhanced to provide greater flexibility and performance with certain modifications whenever necessary. Deep fake learning can help detect fake images. Deep learning machine learning to get more accurate results.

SOCIAL CONTRIBUTION

Since fake news attempts to spread false claims in news content, the most straightforward means of detecting it is to check the truthfulness of major claims in a news article to decide the news veracity. With the increasing popularity of social media, more and more people consume news from social media instead of traditional news media. However, social media has also been used to spread fake news, which has strong negative impacts on individual users and broader society.

REFERENCES

1. Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu, "Fake News Detection on Social Media: A Data Mining Perspective" arXiv:1708.01967v3 [cs.SI], 3 Sep 2017
<https://arxiv.org/pdf/1708.01967.pdf>
2. M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kyiv, 2017, pp. 900-903.
https://www.researchgate.net/publication/330151342_FAKE_STATISTICS_DETECTION_WITH_ENSEMBLE_OF_MACHINE_LEARNING_ALGORITHMS
3. Fake news websites. (n.d.) Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Fake_news_website. Accessed Feb. 6, 2017
https://en.wikipedia.org/wiki/Fake_news_website
4. Cade Metz. (2016, Dec. 16). The bittersweet sweepstakes to build an AI that destroys fake news.
<https://www.ijert.org/fake-news-detection-using-machine-learning-algorithms>
5. Conroy, N., Rubin, V. and Chen, Y. (2015). "Automatic deception detection: Methods for finding fake news" at Proceedings of the Association for Information Science and Technology, 52(1), pp.1-4
<https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/pa2.2015.145052010082>