FAKE NEWS DETECTION USING MACHINE LEARNING USING SVM ALGORITHM

Dr.M.Rajeswari,

Assistant professor,

Department of B.com (Business Analytics),

PSGR Krishnammal College for women Coimbatore, India.

rajeswarim@psgrkcw.ac.in

R.S.Karthikaa,

UG Scholar,

Department of B.com (Business Analytics),

PSGR Krishnammal College for women Coimbatore, India.

292karthikaa@gmail.com

ABSTRACT:

There have been a great range of sources that we see in online to get any news we need, we find them at no time. Though we get a lot of options and wide range of sources the genuineness of the news fails many times. That's where the many researchers arrive with their idea to find out a way to get rid of that fake news, a lot of research papers focus on the reasons behind the spread of fake news, and a lot of researchers have tried and came up with many technical ways to detect fake news. The aim of this work is to create a system or model that can use the data of past news reports and predict the chances of a news report being fake or not. Various researchers have attempted solving this challenge in a multitude of ways to test which method works and get desirable results.

KEY WORDS: fake news, social media, machine language, python, colab, SVM algorithm.

I. INTRODUCTION

We live in the 21th century, where we live by data. Yes, data drive our lives. It's both the way advantage and disadvantage to us. Data are used in many way that we cannot separate it from our daily lives. Though we have access to data in our finger tip we still cannot find the true and fake news that we seek out. Technology is in a way plays it's pro here and it acts also con to the user when there is spreadation of fake news. There are different online platforms where the person can spread the fake news. This includes the Facebook, Twitter etc. Machine learning is the part of artificial intelligence that helps in making the systems that can learn and perform different actions.

A variety of machine learning algorithms are available that include the supervised, unsupervised, reinforcement machine learning algorithms. The algorithms first have to be trained with a data set called train data set. After the training, these algorithms can be used to perform different tasks. Machine learning is using in different sectors to perform different tasks. Most of the time machine learning algorithms are used for prediction purpose or to detect something that is hidden. Due to the majority of society opinion impact

changes, fake news detection is an important challenge to researchers. The detection of misinformation is not an easy task for anyone, but quite is a complex for people. Here, we analyze the different fake news detection approaches followed in current scenario and compute the detection process through machine learning and deep leaning algorithms for better accuracy.

II.OBJECTIVE

The main objective of this project is to detect fake news using SVM (Support Vector Machine). Here we used SVM algorithm to visualise the values of the predicting data with the fake news detection.

III. RELATED WORKS

Though we have various mediums to spread fake news, still there are social media platforms that educate or help detect fake news.[7]At the time of writing, Facebook uses machine learning algorithms to identify false or sensational claims used in advertising for alternative cures, they place potential fake news articles lower in the news feed, and they provide users with tips on how to identify fake news themselves. Some approaches detect fake news by using metadata such as a comparison of release time of the article and timelines of spreading the article as well where the story spread. The spreading of false political information have increased due to the emergence of streamline media environments.[4]

A number of studies have primarily focused on detection and classification of fake news on social media platforms such as Facebook and Twitter. At conceptual level, fake news has been classified into different types; the knowledge is then expanded to generalize machine learning (ML) models for multiple domains. The study by Ahmed et al included extracting linguistic features such as n-grams from textual articles and training multiple ML models including K-nearest neighbor(KNN), support vector machine (SVM), logistic regression (LR), linear support vector machine (LSVM), decision tree (DT), and stochastic gradient descent (SGD), achieving the highest accuracy (92%) with SVM and logistic regression.[8]

Marco L. Della Vedova et. al. first proposed a novel ML fake news detection method which, by combining news content and social context features. Second, they implemented their method within a Facebook Messenger Chabot and validate it with a real-world application. Their goal was to classify a news item as reliable or fake; they first described the datasets they used for their test, then presented the content-based approach they implemented and the method they proposed to combine it with a social-based approach available in the literature.[2]

Shivam B. Parikh et. al. aims to present an insight of characterization of news story in the modern diaspora combined with the differential content types of news story and its impact on readers. Subsequently, we dive into existing fake news detection approaches that are heavily based on text-based analysis, and also describe popular fake news datasets. We conclude the paper by identifying 4 key open research challenges that can guide future research. It is a theoretical Approach which gives Illustrations of fake news detection by analyzing the psychological factor.[10]

These models are based on knowledge such as writing style, and social context such as stance and propagation. Different researchers are working for the detection of fake news. The use of Machine learning is proving helpful in this regard. Researchers are using different algorithms to detect the false news. Researchers in (Wang, 2017) said that fake news detection is big challenge. They haveused the machine learning for detecting fake news.[6]

IV. METHODOLOGY

A. WORK FLOW

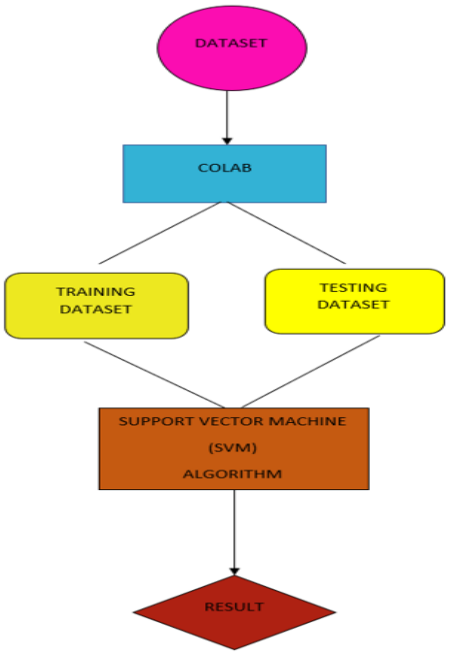

FIG 4.1

B. ALOGRITHM

SVM algorithm, Support Vector Machine is the algorithm that is mostly used for classifications and regression problems. It uses a technique called the kernel trick to transform your data and then based on these transformations it finds an optimal boundary between the possible outputs. SVMs are different from other classification algorithms because of the way they choose the decision boundary that maximizes the distance from the nearest data points of all the classes.

V. IMPLEMENTATION

```
from sklearn import svm

#Create a svm Classifier
clf = svm.SVC(kernel='linear') # Linear Kernel

pipe = Pipeline([('vect', CountVectorizer()),
                 ('tfidf', TfidfTransformer()),
                 ('model', clf)])

model = pipe.fit(X_train, y_train)
prediction = model.predict(X_test)
print("accuracy: {}%".format(round(accuracy_score(y_test, prediction)*100,2)))
dct['SVM'] = round(accuracy_score(y_test, prediction)*100,2)
```

```
accuracy: 99.07%
```

FIG 5.1

The fig5.1 shows the accuracy level of fake news deduction using SVM algorithm.

FIG 5.2

The fig5.2 shows the visualization of fake and real news in confusion matrix

VI. CONCLUSION

Since we have ample way to get news at our hands, as already told it is advantageous and disadvantage too. Fake news detection has many open issues that require attention of researchers. For instance, in order to reduce the spread of fake news, identifying key elements involved in the spread of news is an important step. And again technology has developed so great that we have so many methods of detecting fake news. One such way is the above mentioned machine learning method.

REFERENCES

[1]. Douglas, "News consumption and the new electronic media," The International Journal of Press/Politics, vol. 11, no. 1, pp. 29–52, 2006.

[2]. J. Wong, "Almost all the traffic to fake news sites is from facebook, new data show," 2016.

[3]. M. J. Lazer, M. A. Baum, Y. Benkler et al., "The science of fake news," Science, vol. 359, no. 6380, pp. 1094–1096, 2018.

[4]. S. A. García, G. G. García, M. S. Prieto, A. J. M. Guerrero, and C. R. Jiménez, "The impact of term fake news on the scientific community scientific performance and mapping in web of science," Social Sciences, vol. 9, no. 5, 2020.

[5]. S. Kogan, T. J. Moskowitz, and M. Niessner, "Fake News: Evidence from Financial Markets," 2019, https://ssrn.com/abstract=3237763.

[6]. J. Soll, "The long and brutal history of fake news," Politico Magazine, vol. 18, no. 12, 2016.

[7]. J. Hua and R. Shaw, "Corona virus (covid-19) "infodemic" and emerging issues through a data lens: the case of China," International Journal of Environmental Research and Public Health, vol. 17, no. 7, p. 2309, 2020.

[8]. F. T. Asr and M. Taboada, "Misinfotext: a collection of news articles, with false and true labels," 2019.

[9]. H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," Journal of Economic Perspectives, vol. 31, no. 2, pp. 211–236, 2017.View at: Publisher Site.

[10]. V. L. Rubin, N. Conroy, Y. Chen, and S. Cornwell, "Fake news or truth? using satirical cues to detect potentially misleading news," in Proceedings of the Second Workshop on Computational Approaches to Deception Detection, pp. 7–17, San Diego, CA, USA, 2016.