Fake News Detection Using Machine Learning

Chinmay Chahar[1] Student, Prachi Saxena[2] Student,
Pritika Khurana[3] Student, Dr. Shikha Agarwal[4] *Faculty Mentor*
[1, 2,3,4] Department of Computer Science Indraprastha College for Women
University of Delhi

Abstract— **In today's time with social media around, it becomes difficult to identify and distinguish between real and fake news. Any user can post or spread information through these online platforms without being verified. Fake news can spread like a virus replicating from person to person, leading to a rise in false information. To solve this problem with the help of experimenting with a few datasets and models, we plan to resolve this issue using machine learning. Machine learning is particularly relevant in the perspective of fake messages in social media since it would be impossible for humans to detect fake news on such a widespread platform. Hence, there is a need for machine learning classifiers that can detect fake news automatically. In our research paper, we have trained a model to identify fake news and spam messages which are written in Hindi language.**

I. INTRODUCTION

What is Fake News?
"Fake news" is a term used to refer to fabricated news.News that is reported about an event, position, or point of view that are not necessarily substantiated with facts is what can be described as Fake News.

II. METHODOLOGY

The methodology of this Research Paper includes testing datasets using different machine learning algorithms.
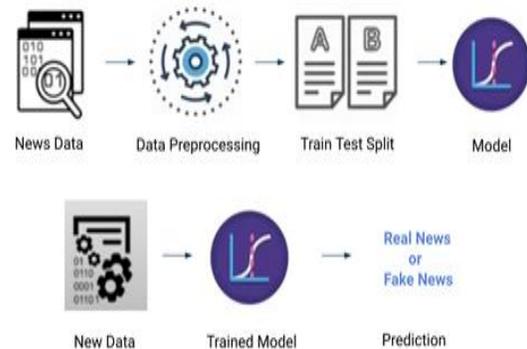
III. WORKFLOW

Fig. 1.

IV. DATASET USED

http://lcs2.iiitd.edu.in/CONSTRAINT-2021/

V. DATABASE

Fig. 2.

	Post	Labels
0	मेरे देश के हिन्दु बहुत निराले है। कुछ तो पक्क...	hate,offensive
1	सरकार हमेशा से किसानों की कमाई को बढ़ाने के लि...	non-hostile
2	सुशांत ने जो बिजनेस डील 9 जून को की थी, वो डील...	non-hostile
3	@prabhav218 साले जेएनयू छाप कमिने लोग हिन्दुओं...	defamation,offensive
4	#unlock4guidelines - अनलॉक-4 के लिए गाइडलाइन्स...	non-hostile
...	...	...
5723	उदितराज जी हिम्मत जुटा कर उस नेता के लिए कुछ ...	hate
5724	उप्र: पीएम रिपोर्ट में खुलासा: हार्टअटैक से हु...	non-hostile
5725	नौकरी गंवा चुके दोस्त की मदद: नगद के बजाए गिफ़...	non-hostile
5726	बंगाल में हिन्दु मरे हैं इसलिए मुझे कोई फर्क न...	fake
5727	रक्षा मंत्री @rajnathsingh ने पूर्व राष्ट्रपति...	non-hostile

5728 rows × 2 columns

- The Database has 5728 rows and 2 columns Post and Labels.
- Post contains the data of a news article.

- Labels contains characteristic label of the post, it can be Hate, non-hostile, defamation, offensive, Fake or a combination of these.

	Post	Labels	fake	hate	defamation	offensive	non-hostile	length
0	मेरे देश के हिन्दु बहुत निराले है। कुछ तो पक्क...	hate,offensive	0	1	0	1	0	100
1	सरकार हमेशा से किसानों की कमाई को बढ़ाने के लि...	non-hostile	0	0	0	0	1	137
2	सुशांत ने जो बिजनेस डील 9 जून को की थी, वो डील...	non-hostile	0	0	0	0	1	204
3	@prabhav218 साले जेएनयू छाप कमिने लोग हिन्दुओं...	defamation,offensive	0	0	1	1	0	144
4	#unlock4guidelines - अनलॉक-4 के लिए गाइडलाइन्स...	non-hostile	0	0	0	0	1	303
...	...	...	...	...	...	...	...	...
5723	उदितराज जी हिम्मत जुटा कर उस नेता के लिए कुछ ...	hate	0	1	0	0	0	187
5724	उप्र: पीएम रिपोर्ट में खुलासा: हार्टअटैक से हु...	non-hostile	0	0	0	0	1	212
5725	नौकरी गंवा चुके दोस्त की मदद: नगद के बजाए गिफ्...	non-hostile	0	0	0	0	1	231
5726	बंगाल में हिन्दू मरे हैं इसलिए मुझे कोई फर्क न...	fake	1	0	0	0	0	70
5727	रक्षा मंत्री @rajnathsingh ने पूर्व राष्ट्रपति...	non-hostile	0	0	0	0	1	123

Fig. 3.

VI. PRE-PROCESSING

- 5 more columns of each unique label have been created containing a binary value. Here, 1 represents presence of that Label in Labels column and 0 represents the absence of the label in the Labels column.
- Column length contains the length of the of the Article in column Post.
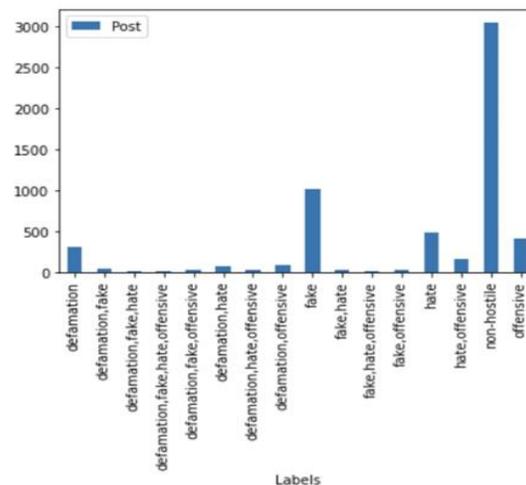
VII. BAR GRAPH REPRESENTATION

Fig. 4.

- The Database has over 3050 Posts labeled non-hostile.
- Over 1009 posts are labelled Fake.
- 478 posts are labelled hate.
- 405 posts are labelled offensive.
- 305 posts are labelled defamation.
- The rest of the posts are made of combinations.

VIII. FAKE NEWS DETECTION USING DIFFERENT MACHINE LEARNING ALGORITHMS

We employ Scikit-learn library for the implementation of various algorithms. we employed 4 best algorithms for binary data namely, Logistic regression, Naive-Bayes, Decision Tree and Passive-Aggressive Classifier. In order to detect all 5 parameters of our dataset one by one.

A. *Code for Logistic Regression*

```
from sklearn.linear_model import LogisticRegressio
logreg = LogisticRegression()
logreg.fit(x_train, y_train)

Accuracy     =     logreg.score(x_test,     y_test)
print(Accuracy*100)
```

B. *Code for Naive-Bayes*

```
from sklearn.naive_bayes import MultinomialNB NB
= MultinomialNB()
NB.fit(x_train, y_train)
Accuracy     =     NB.score(x_test,     y_test)
print(Accuracy*100)
```

C. *Code for Desicion Tree*

```
from sklearn.tree import DecisionTreeClassifier clf =
DecisionTreeClassifier()
clf.fit(x_train, y_train)
Accuracy     =     clf.score(x_test,     y_test)
print(Accuracy*100)
```

D. *Code for Passive-Aggressive Classifier*

```
from sklearn.metrics import accuracy_score
from sklearn.linear_model import PassiveAggressive
pac=PassiveAggressiveClassifier(max_iter=50)
pac.fit(x_train,y_train)

y_pred=pac.predict(x_test)
score=accuracy_score(y_test,y_pred)
```

```
print(f'Accuracy: {round(score*100,2)}%')
```
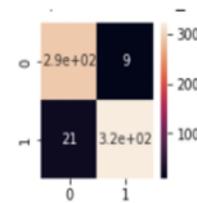
IX. PREDICTION

Fig. 5. Non- Hostile Logistic Regression Confusion Matrix

1) *Non-Hostile Prediction:*
- Logistic Regression: 95.37
- Naive-Bayes: 89.17
- Decision Tree: 92.38
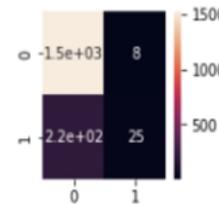- Passive-Aggressive Classifier:92.85

Fig. 6. OffensiveLogistic Regression Confusion Matrix

2) *Offensive Prediction:*
- Logistic Regression: 88.10
- Naive-Bayes: 87.36
- Decision Tree: 84.37
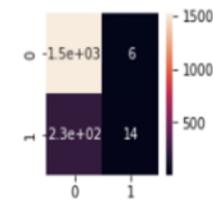- Passive-Aggressive Classifier:84.81

Fig. 7. Defamation- Logistic Regression Confusion Matrix

3) *Defamation Prediction:*
- Logistic Regression: 88.75
- Naive-Bayes: 87.79
- Decision Tree: 85.71
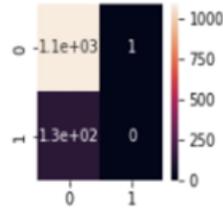- Passive-Aggressive Classifier:84.74

Fig. 8. Hate - Naive-Bayes Confusion Matrix

4) *Hate Prediction:*
- Logistic Regression: 89.35
- Naive-Bayes: 89.35
- Decision Tree: 86.23
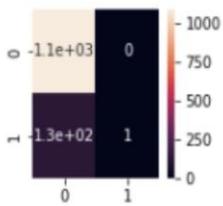- Passive-Aggressive Classifier:87.29

Fig. 9. Fake - Logistic Regression Confusion Matrix

5) *Fake Prediction:*
- Logistic Regression: 86.60
- Naive-Bayes: 85.01
- Decision Tree: 82.09
- Passive-Aggressive Classifier:81.93

X. LITERATURE REVIEW

AI calculations can be utilized to distinguish fake news. This is accomplished through utilizing various kinds of preparing datasets to refine the calculations. Datasets are utilized to prepare the calculations to distinguish counterfeit news. An ML approach called the talk distinguishing proof structure has been fostered that legitimizes signs of uncertain posts so an individual can undoubtedly recognize counterfeit news (Sivasangari et al. 2018). Some of these popular classifiers are given below that are used for this purpose.

Support Vector Machine: This algorithm is a supervised machine learning algorithm that learns from the labeled data set, mostly used for classification. Researchers in (Singh et al., 2017) used various classifiers of machine learning and therefore the support vector machine have given them the most effective leads to detecting the fake news.

Random Forests: During this classifier, there are different random forests that provides a value and a value with more votes is the actual result of this classifier.

Decision Tree: This supervised algorithm of machine learning can help to detect the fake news. It breaks down the dataset into different smaller subsets.

Naïve Bayes: This calculation is for the most part utilized for order. This is a managed AI calculation that utilizes different classifiers of AI and the help vector machine has given them the best outcomes in recognizing the fake news.

Logistic Regression: This classifier is used when the value to be predicted is categorical. For example, it can predict or give the end in true or false.

K-Nearest Neighbor: This is a supervised algorithm of ma- chine learning that is used for solving classification problems. It stores the data about all the cases to classify the new case on the basis of similarity.

XI. CONCLUSION

The best accuracy was achieved by using logistic regression while detecting non-hostile news. Therefore, it can be concluded that it is easier to predict the regular patterns of non-hostile news as compared to the irregular fake news. As depicted in the table below:

Algorithm	Hate	Fake	Offensive	Defamation	Non-hostile
Logistic Regression	89.35%	86.60%	88.10%	88.75%	95.37%
Naïve- Bayes	89.35%	85.01%	87.36%	87.79%	89.17%
Decision Tree	86.23%	82.09%	84.37%	85.71%	92.38%
Passive Aggressive Classifier	87.29%	81.93%	84.81%	84.74%	92.85%

Fig. 10.

XII. ROAD AHEAD

We aim to increase our accuracy by:
- Testing more pre-defined datasets against our algorithms.
- Try combinations of different algorithms for predicting Fake news.
- Scrape the web to create and test against more datasets.

REFERENCES

[1]   https://www.tutorialspoint.com/machine$_l$ earning$_w$ith$_p$ython/classification$_a$lgorithm s$_d$ecision$_t$ree.htm

[2]   https://www.geeksforgeeks.org/passive-aggressive-classifiers/

[3]   https://arxiv.org/pdf/2011.03588.pdf