



**IJIRCCCE**

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

**Volume 10, Issue 4, April 2022**

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 8.165**



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

# Fake News Detection Using Machine Learning

Anamika Choudhary, Rutuja Gandhe, Aditya Dagda, Prof. Pooja Ojha

Department of Computer Engineering, JSCOE, Hadapsar, Pune, India

**ABSTRACT:** In today's world internet is the most used platform. Everyone depends on different online resources for news. Due to major use of social media platforms like Facebook, Twitter etc. news spread swiftly among millions of people in no time. Because, the Fake news spreads rapidly, it has various consequences like biased election results which benefits the candidates.

Alluring headlines are used by the spammers for generating capital using advertisements. In this paper, we aim to classify different news available by using machine learning. We want help the user to be able to check whether the news is fake or real.

## I. INTRODUCTION

As we know people spend a lot of time online whether it is social media or for interacting. 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. This happens mostly because it is less time consuming and low cost rather than spending time and money on traditional sources like newspaper and other ones. Even though getting news on social media is easy it is not as standard as newspaper. Most of the time the news that is passaged through social media is fake because maintaining the standard of news online is far more costly than the traditional source because updating the news online means spending more money internet and many more attributes.

## III. LITERATURE SURVEY

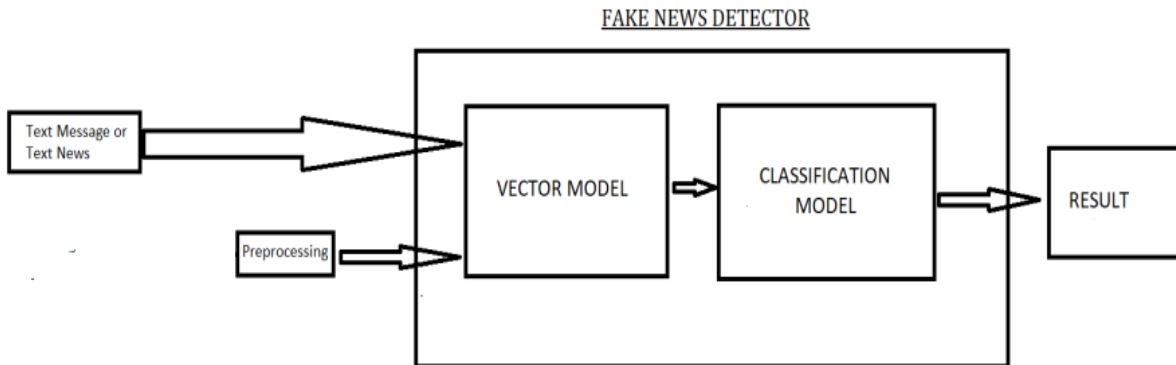
- ❑ MykhailoGranik et. al. in their paper shows a **simple approach for fake news detection using naive Bayes classifier**. This **approach was implemented as a software system and tested against a data set of Facebook news posts**. They were collected from three large Facebook pages each from the right and from the left, as well as three large mainstream political news pages (Politico, CNN, ABC News). **They achieved classification accuracy of approximately 74%**. Classification accuracy for fake news is slightly worse. This may be caused by the skewness of the dataset: only 4.9% of it is fake news.
- ❑ Himank Gupta et. al. gave a **framework based on different machine learning approach that deals with various problems including accuracy shortage**, time lag (BotMaker) and high processing time to handle thousands of tweets in 1 sec. **Firstly, they have collected 400,000 tweets from HSpam14 dataset. Then they further characterize the 150,000 spam tweets and 250,000 non-spam tweets**. They also derived some lightweight features along with the Top-30 words that are providing highest information gain from Bag-of-Words model. **4. They were able to achieve an accuracy of 91.65% and surpassed the existing solution by approximately 18%**.
- ❑ 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 textbased 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 analysing the psychological factors**.

There are two main categories of state of the art that are interesting for this work: previous work on fake news detection and on general text classification. Works on fake news detection is almost inexistent and mainly focus in 2016 US presidential elections or does not use the same features. That is, when this work focus on automatic features extraction using machine learning and deep learning, other works make use of hand-crafted

features[13, 14] such as psycholinguistic features[15] which are not the goal here. Current research focus mostly on using social features and speaker information in order to improve the quality of classifications.

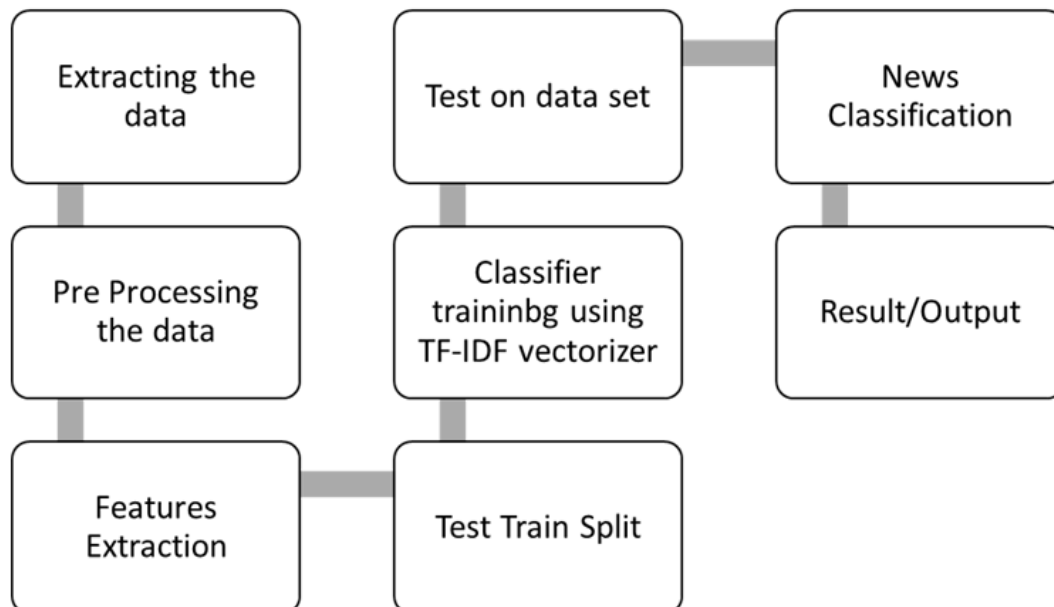
**IV. METHODOLOGY**

i) Proposed model:



As above figure shows, our system is a web application which helps user to detect fake news. A text box is displayed for the user to copy and paste the news to detect whether it is fake or real.

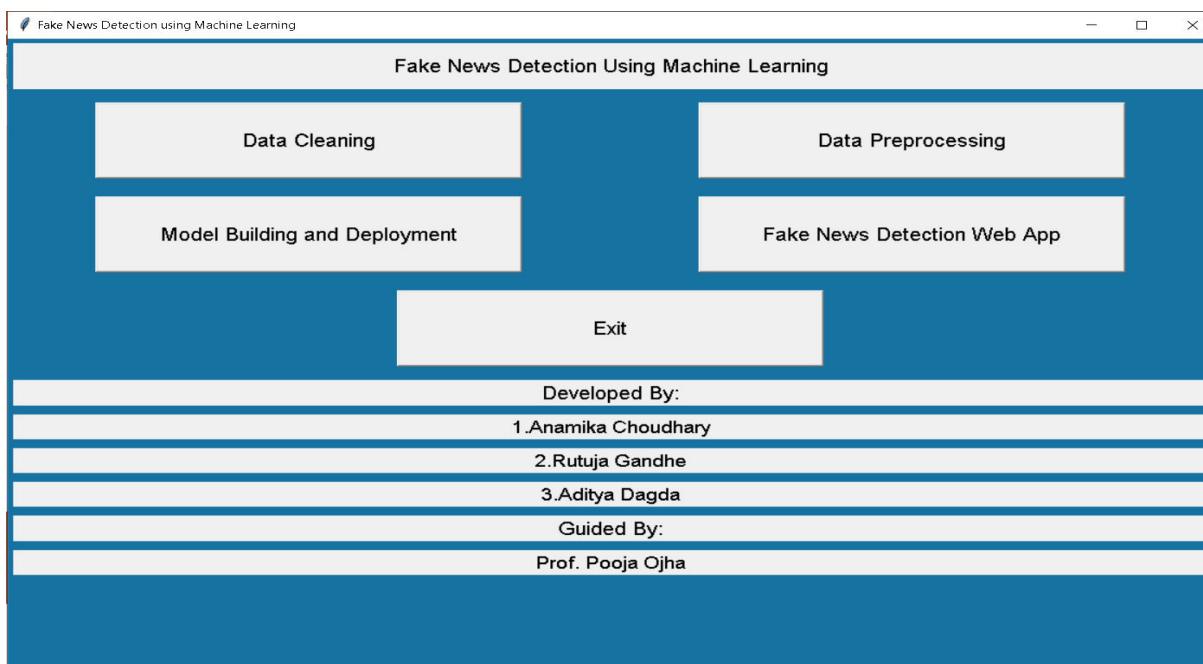
ii) System Architecture:



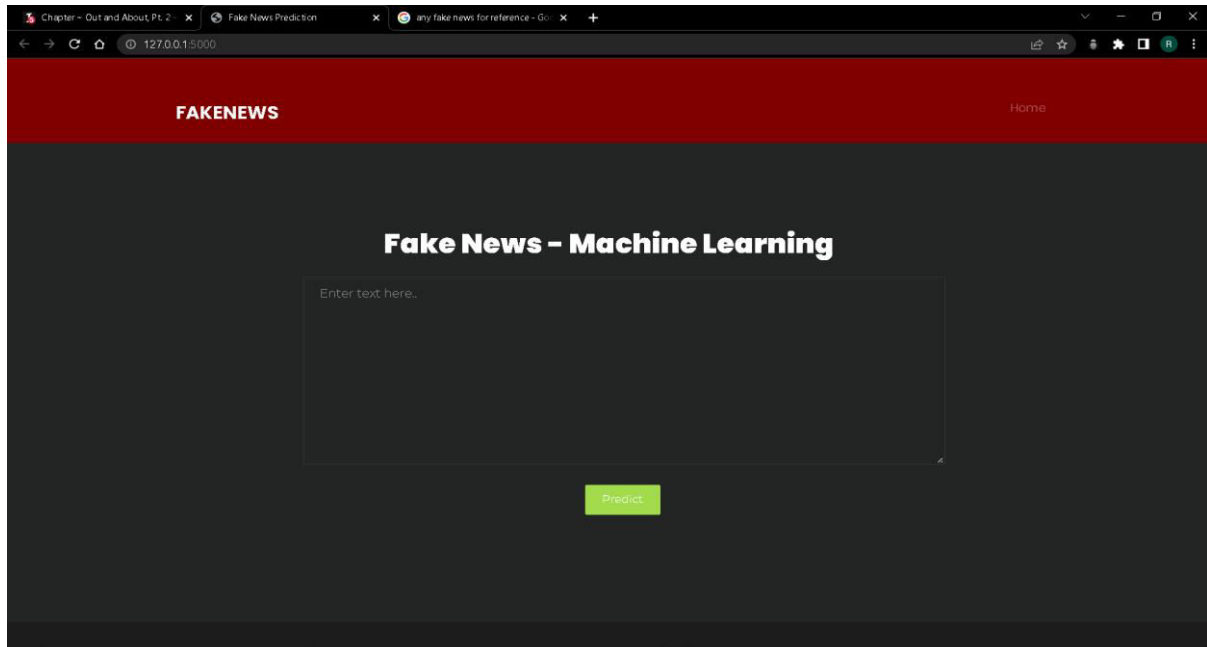
### V. IMPLEMENTATION STEPS:

1. Download Dataset from Kaggle website.
2. Libraries required for the project.
  - NumPy: NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.
  - pandas: Pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series.
  - Sklearn: Scikit-learn is a free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms.
  - nlTK: The Natural Language Toolkit, or more commonly NLTK, is a suite of libraries and programs for symbolic and statistical natural language processing for English written in the Python programming language.
3. Remove all non-words from news column.
4. Remove all stop-words using nltk
5. Then create features using tf-idf vectorizer.
6. Then create X as input variable having feature and Y as output variable having label 0 for fake and 1 for real.
7. Then split data in training and testing.
8. Then pass training data to Passive Aggressive Classifier algorithm. Passive Aggressive Classifier belongs to the category of online learning algorithms in machine learning. It works by responding as passive for correct classifications and responding as aggressive for any miscalculation.
  - Passive: If the prediction is correct, keep the model and do not make any changes. i.e., the data in the example is not enough to cause any changes in the model.
  - Aggressive: If the prediction is incorrect, make changes to the model. i.e., some change to the model may correct it.
9. Then we will check accuracy of model and deploy it on flask web framework.
10. Our web application will take a news as input and will classify it as Fake or Real according to model.

### VI.SNAPSHOTS OF THE SYSTEM







## VII. RESULT

Hence, in the above paper we discussed the components to detect fake news, and its significance to find the accuracy of the news available on internet.

## 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
- [2] Julio CS Reis, Andr'e Correia, Fabr'icioMurai, Adriano Veloso, Fabr'icioBenevenuto, and Erik Cambria. Supervised learning for fake news detection. IEEE Intelligent Systems, 34(2):76–81, 2019.
- [3] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kiev, 2017, pp. 900-903.
- [4] Fake news websites. (n.d.) Wikipedia. [Online]. Available: [https://en.wikipedia.org/wiki/Fake\\_news\\_website](https://en.wikipedia.org/wiki/Fake_news_website). Accessed Feb. 6, 2017
- [5] Cade Metz. (2016, Dec. 16). The bittersweet sweepstakes to build an AI that destroys fake news.
- [6] 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.
- [7] Markines, B., Cattuto, C., &Menczer, F. (2009, April). —Social spam detection|. In Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web
- [8] Rada Mihalcea , Carlo Strapparava, The lie detector: explorations in the automatic recognition of deceptive language, Proceedings of the ACL-IJCNLP
- [9] Kushal Agarwalla, Shubham Nandan, Varun Anil Nair, D. Deva Hema, —Fake News Detection using Machine Learning and Natural Language Processing,| International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6, March 2019
- [10] H. Gupta, M. S. Jamal, S. Madisetty and M. S. Desarkar, "A framework for real-time spam detection in Twitter," 2018 10th International Conference on Communication Systems & Networks (COMSNETS), Bengaluru, 2018, pp. 380-383.
- [11] M. L. Della Vedova, E. Tacchini, S. Moret, G. Ballarin, M. DiPierro and L. de Alfaro, "Automatic Online Fake News Detection Combining Content and Social Signals," 2018 22nd Conference of Open Innovations Association (FRUCT), Jyvaskyla, 2018, pp. 272-279.
- [12] C. Buntain and J. Golbeck, "Automatically Identifying Fake News in Popular Twitter Threads," 2017 IEEE International Conference on Smart Cloud (SmartCloud), New York, NY, 2017, pp. 208-215.



INNO  SPACE  
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

 **doi**<sup>®</sup>  
**cross** **ref**

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  [ijircce@gmail.com](mailto:ijircce@gmail.com)



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details