# Fake News Detection Using Machine Learning

**Prof. C. P. Lachake[1], Ritesh Paliwal[2], Akshay Patil[3], Tejas Chaudhari[4], Harshal Borse[5]**

Assistant Professor, Department of Computer Engineering, SKN Sinhgad institute of technology and science,

Lonavala, India[1]

Student, Department of Computer Engineering, SKN Sinhgad institute of technology and science,

Lonavala, India[2,3,4,5]

**Abstract:** The phenomenon of Fake news is experiencing a rapid and growing progress with the evolution of the means of communication and Social media. Fake news detection is an emerging research area which is gaining big interest. It faces however some challenges due to the limited resources such as datasets and processing and analysing techniques.In this work, we propose a system for Fake news detection that uses machine learning techniques. We used term frequency-inverse document frequency (TF-IDF) of bag of words and n-grams as feature extraction technique, and Support Vector Machine (SVM) as a classifier. We propose also a dataset of fake and true news to train the proposed system. Obtained results show the efficiency of the system.Index Terms—Fake news, Social media, Web Mining, Machine Learning, Support Vector Machine, TF-IDF.

## I. INTRODUCTION

In the last decade, Fake News phenomenon has experienced a very significant spread, favored by social networks. This fake news can be broadcasted for different purposes. Some are made only to increase the number of clicks and visitors on a site. Others, to influence public opinion on political decisions or on financial markets. For example, by impacting the reputation of companies and institutions on the Web. Fake news concerning health on social media represents a risk to global health.

The WHO warned in February 2020 that the COVID-19 outbreak had been accompanied by a massive 'infodemic', or an overabundance of information—some of which was accurate and some of which was not which made it difficult for people to find reliable sources and trustworthy information when they needed it. The consequences of disinformation overload are the spread of uncertainty, fear, anxiety and racism on a scale not seen in previous epidemics [11].In this paper, we present a novel method and tool for detecting fake news that uses: • Text preprocessing: consisting of steaming and ana-lyzing the text by removing stop words and special characters. • Encoding of the text: using bag of words and N-gramthen TF-IDF• Extraction of the characteristics: this allows a precise identification of false information. We use the source of a news, its author, the date and the feeling given by the text as features of a news. • Support vector machine: a supervised machine learning algorithm that allows the classification of new information.

## II. RELATED WORK

In literature, many works are interested to fake new detection.Authors of [3] propose a typology of several methods of truth assessment emerging from two main categories: linguistic cue approaches with machine learning and network analysis approaches, for detecting fake news. In [5], authors present a simple approach to fake news detection using a naive Bayesian classifier. This approach is tested on a set of data extracted from Facebook news posts.They claim to be able to achieve an accuracy of 74%. The rate of this model is good but not the best, as many other works have achieved a better rate using other classifiers. We discuss these works in the following. Authors of [1] propose a fake news detection model that uses n-gram analysis and machine learning techniques by comparing two different feature extraction techniques and six different classification techniques. The experiments carried out show that the best performances are obtained by using the so-called features extraction method (TF-IDF). The used the Linear Support Vector Machine (LSVM) classifier that gives an accuracy of 92%. This model uses LSVM that is limited to treat only the case of two linearly separated classes.Authors of [6] describe how users of social networks can ensure the truth of information. They also describe the mechanisms that allow their validation and the role of journalists or what to expect from researchers and official institutions. This work helps people see a little bit of the truth behind the news on social media and not believe anything.Authors of [9] propose several strategies and types of indices relating to different modalities (text, image, social information). They also explore the value of combining and merging these approaches to assess and verify shared information.
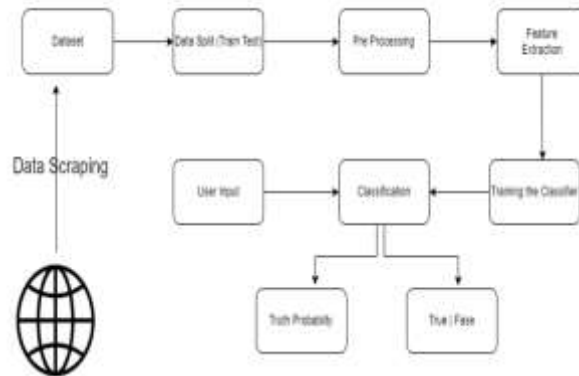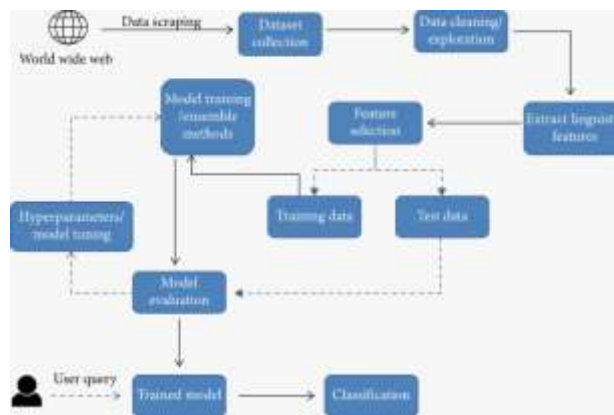
Fig.1. Architecture Diagram



Fig.2. Use-Case Diagram

## III. CONCLUSION

Thus we aim to improve Accuracy of news which is shared on social media that claims the true news which you can read and share with your friends without any hesitation.
This will also make people to read and share the true news on social media.

## REFERENCES

[1] Hadeer Ahmed, Issa Traore, and Sherif Saad. Detection of online fake news using n-gram analysis and machine learning techniques. In International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, pages 127–138. Springer, 2017.
[2] Chih-Chung Chang and Chih-Jen Lin. LIBSVM – A Library for Support Vector Machines, July 15, 2018.
[3] Niall J Conroy, Victoria L Rubin, and Yimin Chen. Automatic deception detection: Methods for finding fake news. Proceedings of the Association for Information Science and Technology, 52(1):1–4, 2015.
[4] Chris Faloutsos. Access methods for text. ACM Computing Surveys (CSUR), 17(1):49–74, 1985.
[5] Mykhailo Granik and Volodymyr Mesyura. Fake news detection using naive bayes classifier. In 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), pages 900–903. IEEE, 2017.
[6]Florian Sauvageau. Les fausses nouvelles, nouveaux visages, nouveaux défis. Comment déterminer la valeur de l'information dans les sociétés démocratiques? Presses de l'Université Laval, 2018.