Original Research Article

Evaluating Fake News Detection Models

Dr. S. Nagarajan[1*], S.P. Sudha[2]

[1]Associate Professor, Department of Computer Science Engineering, Swarnandhra College of Engineering and Technology, Narsapur, Andhra Pradesh, India
[2]Assistant Professor, Department of Artificial Intelligence and Machine Learning, Swarnandhra College of Engineering and Technology, Narsapur, Andhra Pradesh, India

*Corresponding Author: Dr. S. Nagarajan
Associate Professor, Department of Computer Science Engineering, Swarnandhra College of Engineering and Technology, Narsapur, Andhra Pradesh, India

Abstract: Fake news is untrue information presented as news. It often the aim of damaging the reputation of a person or entity or making money through advertising like a picture, short film, song, etc., revenue. Due to exponential growth of information in online, it is very impossible to figure out the true from the false. Some of the social media networks such as Twitter, Instagram and Facebook are affected by their user through fake news, and it is very hard to catch the fake stories before they go viral. Today most of the people prefer to search and absorb media news instead conventional one. Now days the widespread distribution of fake news may affect people and society. So we decide to build a web app to detect the fake news. The credibility and trust in the news media are at an all-time low with an object that is made to look real or valuable in order to deceive people. It is becoming increasingly difficult to determine which news is real and which is fake various machine learning methods have been used to separate real news from fake ones like experts revealed that the painting was a fake. There are some of the aspects that has to be kept in mind considering the factas something is believable, it seems possible, real, or true that fake news detection is not only a simple web interface but also a quite complex thing that includes a lot of backend work to identify the information or reports about recent events.

Keywords: Fake, News, Information, Online, Fake news, detection.

INTRODUCTION

Fake news detection system

In the world of rapidly growing technology, information facts about a situation, person, event, etc,. Sharing has become an easy task. There is no doubt especially about how good or true it is found in real life that internet has made our lives easier and access to lots of information. This is an evolution in human history, but at the same time it to not direct your attention or your efforts towards the right activity, subject, or problem; to direct your attention, etc. in the wrong way the line between true media and maliciously forget media. Today any one can publish content credible ornate that can be consumed by the storage and retrieval of information through. World Wide Web, Sadly, fake news accumulates a great deal of attention over the internet, especially on social media. People get deceived and don't think twice before circulating such miss-informative pieces to the world.

The motive of this project is to increase the accuracy of detecting fake news more than the present results that are available. By fabricating this new model which will judge the news articles on the base on the title of the news. We will be working on different fake news data set in which we will apply different machine learning algorithms to train the data and test it to find which news is the real new or which one is the fake news. In recent years, due to the huge development of online social networks, fake news for various commercial and political purposes has been appearing in large numbers and widespread in the online world. With deceptive words, online social network users can get infected by this online fake news easily. So it is very important to build a fake news detection system.

SYSTEM DEVELOPMENT

Problem statement

The spread of fake news has become a significant challenge in the digital age; leading to miss information, manipulation, and social unrest. Traditional methods of news verification are time-consuming and often insufficient to keep up with the rapid dissemination of information on social media platforms and online news outlets. To combat this issue, there is a need to develop an effective and efficient automated system for detecting fake news using machine learning techniques. The problem at hand is to design and implement a machine learning model that can accurately classify news articles as either real or fake. The model should be capable of analyzing various textual features, including language patterns, writing style, source credibility, and contextual information, to make informed predictions.

Existing system

Existing systems for fake news detection utilize various approaches and techniques to identify and classify misleading or false information. The system gathers a diverse range of news articles, social media posts, and other textual content from various sources. This data forms the basis for training and testing the fake news detection model. During preprocessing the textual data is processed to remove noise and irrelevant information. This involves tasks such as tokenization, stop word removal. Additionally, the data may undergo cleaning to eliminate special characters, and punctuation. Machine learning algorithms or deep learning models are employed to classify news articles into fake or genuine categories. The classification model is trained using labeled datasets that consist of news articles labeled as fake or real.

Proposed system

Introduction the proposed system that we are going to implement has interactive user inter face to use. It is very useful for the people who want to know whether the news is fake or real. We use machine learning to implement the fake news detection model. We used dataset which contains numerous news title and description. The model is used the stop words removal pre-processing technique. The purpose of removing stop words is to reduce the dimensionality of the text data and to eliminate noise that can negatively impact the performance of machine learning algorithms. By removing these words, we can focus more on the important and meaningful words that carry the essence of the text. We used TF-IDF vectories for feature extraction. The TF-IDF vectorizeris a feature extraction method that converts a collection of documents into a matrix of TF-IDF features. We used a Passive Aggressive Classifier to classify the data set. The Passive Aggressive classifiers a machine learning algorithm that is used for binary classification tasks. It is very useful when dealing with large scale data. This classifier classifies the input which is given by the user in the web page and predict the output whether the news is fake or real. In extent we also provided envelop mechanisms for users to provide feedback on the accuracy of the system predictions. The feedback can be used to continuously update the model, making it more effective overtime.

Advantages

- Fake News detection system is help to controlling the spread of fake news over social media.
- This way, we can help the people to make informed decisions, and they are made to think about what others are trying to manipulate to believe.
- A Fake news detection system reduces the burden to check the authenticity of the news manually and saves lots of time.

Architecture design

An architectural diagram is a diagram of a system that is used to abstract the overall outline of the software system and the relationships, constraints and boundaries between components. It is an important tool as it provides an overall view of the physical deployment of the software system and its evolution roadmap.

METHODOLOGY

Dataset

The data set is simple. It contains the titles of the news, the body text and a label field, which, if the news is authentic, shows REAL and if inauthentic, shows FAKE. There are 3 main segments of the methodology:

- The Machine Learning model.
- The web interface.
- The common platform that brings the model and the interface together.

The machine learning model

Machine Learning is a part of our life that can help us in predicting. Machine learning algorithms analyse and interpret large amounts of data, identifying patterns, trends that can be used to make predications, classify objects.

➢ **Data Loading:** We are loading a CSV file for the data sorting and training-testing part of the model. The CSV file is turned into an array for easier work purpose.

➤ **Vectorization:** Vectorization is needed for determining the frequency of the words present in a passage. This is needed to determine which words are used often.
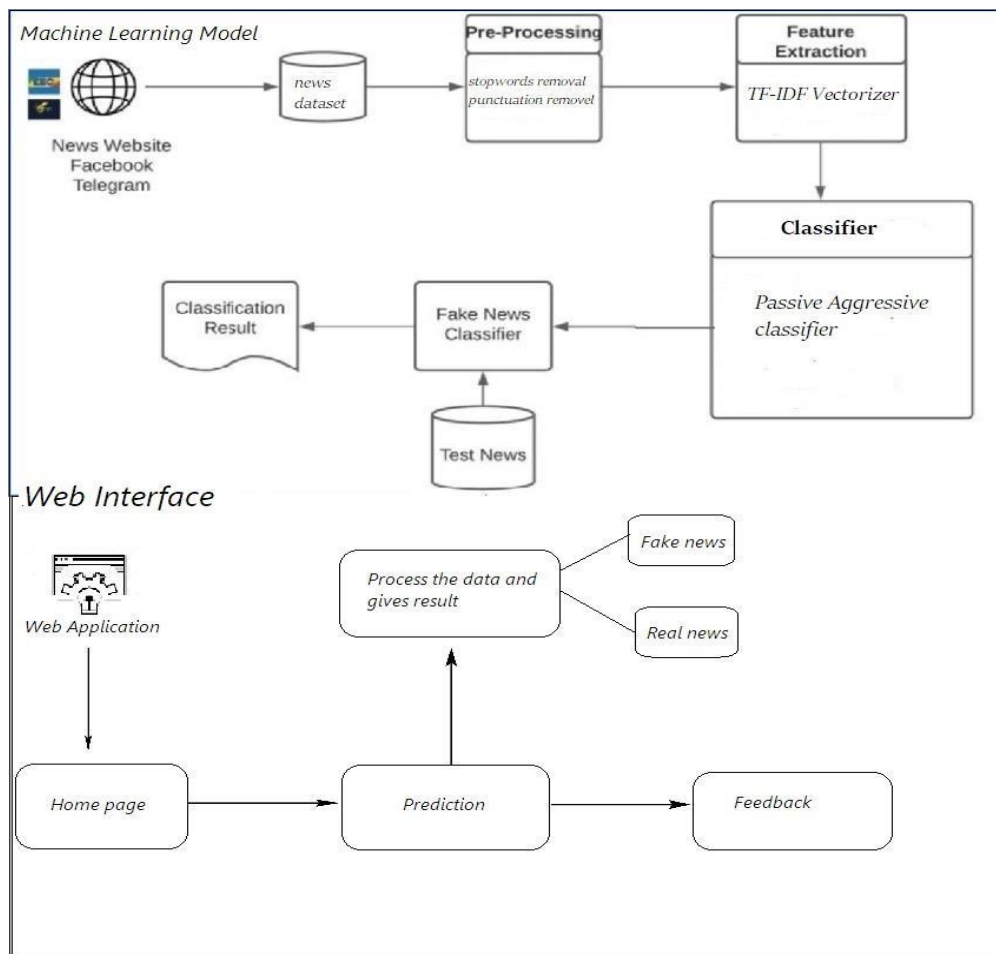

Figure 1: Architecture Design

➤ **Classifier:** Passive-aggressive algorithms are a family of great learning algorithms. Passive is used when the prediction is correct and there is no change in the model. But if there is any kind of change in the model, that is if the prediction is not correct then the aggressive part is called, which changes the model accordingly. The aggressive part of the model changes the model according to its wish on the backend.

➤ **Model Building:** The model is built through the train and test of the dataset.

Web interface
This was the simplest part.
1. **HTML**: HTML makes the structure of the web application and also there are some of the functions that can be achieved best with HTML only.
2. **CSS**: The CSS part is for designing only because it will give a more beautiful aspect to the website.

COMMON PLATFORM: FLASK
This acts as a common platform and takes the input with the pickle module and passes it to the machine learning model afterwards the prediction is shown on the screen with the HTML and CSS website.
- Building functions for taking input.
- Passing input values through the ML model.
- Using the Pickle module for serializing and de-serializing the dataset.
- Providing output.

Hardware requirements
- RAM: 4GB
- Storage: 500GB
- CPU: 2GHz or faster

- Architecture: 32-bit or 64-bit

Software requirements
- Coding language: Pycharm & Jupyternotebook
- Frontend: HTML, CSS, BOOTSTRAP5

CONCLUSION

Many people consume news from social media instead of traditional news media. However, social media has also been used to spread fake news, which has negative impacts on individual people and society. In this paper, an innovative model for fake news detection using machine learning algorithms has been presented. We believe the web interface provides an easier way for the average person to check the authenticity of news. Projects like this one with more advanced features should be integrated on social media to prevent the spread off news.

Future work

In future we can propose fake news detection model which is used to predict the real time news. Now every fake news detection system is based on text. In future we can propose fake news detection by using images, Urls and audios. I don't mean that my project is the best or that I have used the best technology available it just a simple and humble venture is easy to understand. We add develop mechanisms for users to provide feedback on the accuracy of the system predictions. But we can store the feedback from details only in the list; it can be viewed by admin. In extent we can add the feedback form for the users and add for min to the database.

REFERENCES

1. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, *19*(1), 22-36.
2. Kaurand, H., Singh, U. (2019). Fake News Detection using Machine Learning: A Systematic Literature Review". *International Conference on Sustainable Computing and Intelligent Systems (ICSCIS).*
3. Zellers, R., Holtzman, A., Rashkin, H., Bisk, Y., Farhadi, A., Roesner, F., & Choi, Y. (2019). Defending against neural fake news. *Advances in neural information processing systems*, *32*.
4. Fariha, N., Salman, M., Qureshi, S., and Qadir, M.A. (2020). Fake News Detection using Machine Learning Techniques: A Systematic Review". *International Conference on Computing, Electronics & Communications Engineering(iCCECE).*
5. Tacchini, E., Ballarin, G., Vedova, M.L.D., Moret, S., and deAlfaro, L. (2017). Automated Fake News Detection in Social Networks arXiv:1704.07506v1[cs. LG].