**Project Name:** Fake News Detection

**Github Link:** <https://github.com/utkrisht2000/Fake-News-Detection>

**Why this project was created?**

Due to the recent increase in the use of social media and presence of people generating content, the popularity of fake news is at an all time high. As the findings of the surveys on fake news show, people are inconsistent, if not terribly bad, who spot fake news. Since then, many efforts have been made to automate the process of detecting fake news.

**What problem is it solving?**

The widespread dissemination of fake news has the potential to have very detrimental effects on people and society. As a result, the detection of fake news on social media has lately emerged as an area of research that is receiving a lot of interest. Social media brings unique characteristics and obstacles for fake news detection, rendering typical news media detection algorithms inefficient or irrelevant. We need to integrate auxiliary information, such as user social engagements on social media, to aid in making a conclusion because fake news is purposefully designed to deceive readers into believing false information, making it challenging and nontrivial to identify based on news content.

**Entire explanation of project**

* **PROPOSED APPROACH**

Inputs are gathered from a variety of sources, including newspapers and social media, and saved in the dataset. The dataset will be used to feed the system. Preprocessing removes extraneous information from data sets and, if required, changes the column data type. In the preceding phase, the Jupyter Notebook and Python libraries are employed. In the first phase, the vectorizing approach is applied. We need to use a dataset to train the machine to recognise bogus news. The data collection is separated into two datasets before being used to detect false news. The remaining 20% is utilised for exams and the remaining 80% is used for training. The model is trained using the train dataset utilising the Passive Aggressive Algorithm, Random Forest, and other clustering techniques. During testing, the test dataset is used as input and the anticipated output is used as output. The result confusion matrix is used to compare the expected and actual outcomes after the trial time. The confusion matrix provides information on the number that was properly anticipated. In the situation of genuine and false belief, and false. The test dataset's accuracy is calculated using the calculation Number of Correct Predictions / Total Input Size.

Algorithm for creating fake news detector model :

**Step 1:** Dataset is imported

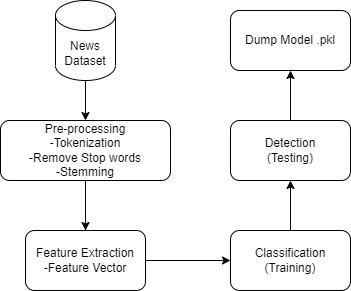
**Step 2:** The data is preprocessed, and the dataset is split into training and testing.

**Step 3:** The training data is converted to numerical using the vectorization approach.

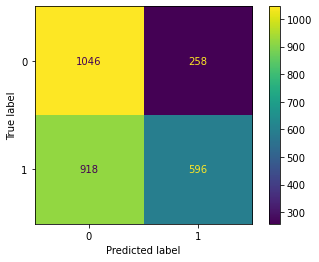
**Step 4:** The training dataset is used to develop the prediction model, which includes the Passive Aggressive Algorithm, Random Forest, and other classification methods.

**Step 7:** A confusion matrix is created.

* **DATA FLOW DIAGRAM**



* **RESULT**



* **CONCLUSION**

With the help of Machine Learning we are creating prediction models with different classification algorithms and choosing one which will give good accuracy and it covers all the latest available data. We intend to save the current queries and result in a database so that it can be utilised to train future models.