Fake News Prediction

By: Hariharan Manickam

Abstract

The dataset I have used is a **news** dataset.

The question is “Can we use a **classifier machine learning model** to **predict** if a news is real or fake? ”.

We used a **Passive Aggressive Classifier**, which is updated **step-by-step**, as opposed to **batch learning**, where the entire training dataset is used at once.

Post analysis, we found that our model was **92.69%** accurate.

Motivation

We can all agree that news **spreads** across the globe **instantaneously** through the internet **regardless** of it being real/fake. Fake news can cause **harm** to people.

It is important that we, as budding data scientists, devise a method to verify the **credibility** of the news, and **prevent** fake news from **spreading**. People who consume news will benefit as they will no longer consume fake news. Problems that arise due to fake news will no longer be a problem.

Through **machine learning**, I will try to **predict** and **classify** news as real/fake and stop the spreading of news in real time.

Example: XYZ is the cure for coronavirus. But in reality, there has been no cure for coronavirus. People who read and believe it will use XYZ to cure themselves of the virus regardless of it not being medically proven effective. Machine Learning model should label this as fake.

Dataset

The dataset I have used is a **news** dataset.

It contains **4** columns:

1. (unnamed column): Random numbers which is most likely a userid
2. 'title' : The title of the news.
3. 'text' : A short summary of the news.
4. 'label' : A label which says if the news is Fake/Real.

It contains **6335** news records.

Link for the dataset:
https://drive.google.com/file/d/1er9NJTLUA3qnRuyhfzuN0XUsoIC4a-_q/view

Data Preparation and Cleaning

I had to **create** a new dataframe that was the **Y variable** for the Machine Learning algorithm.

I created **training** and **testing datasets** from the original dataset to perform classification. Then a **Tfidf Vectorizer** was used to **remove** words which have a document frequency higher than **0.7**.

Kindly go through the Jupyter Notebook to better understand how the Tfidf Vectorizer works as the process has been described in a **step-by-step** manner.

Research Question

Based on the dataset that I have selected, the research question I want to ask is:

Can we use a **classifier machine learning model** to **predict** if a news is real or fake?

Methods

A **Passive Aggressive Classifier** is generally used for **large-scale** learning. It is one of the few **online-learning** algorithms. In online ML algorithms, the input data comes in **sequential** order and the ML model is updated **step-by-step**, as opposed to **batch learning**, where the entire training dataset is used at once.

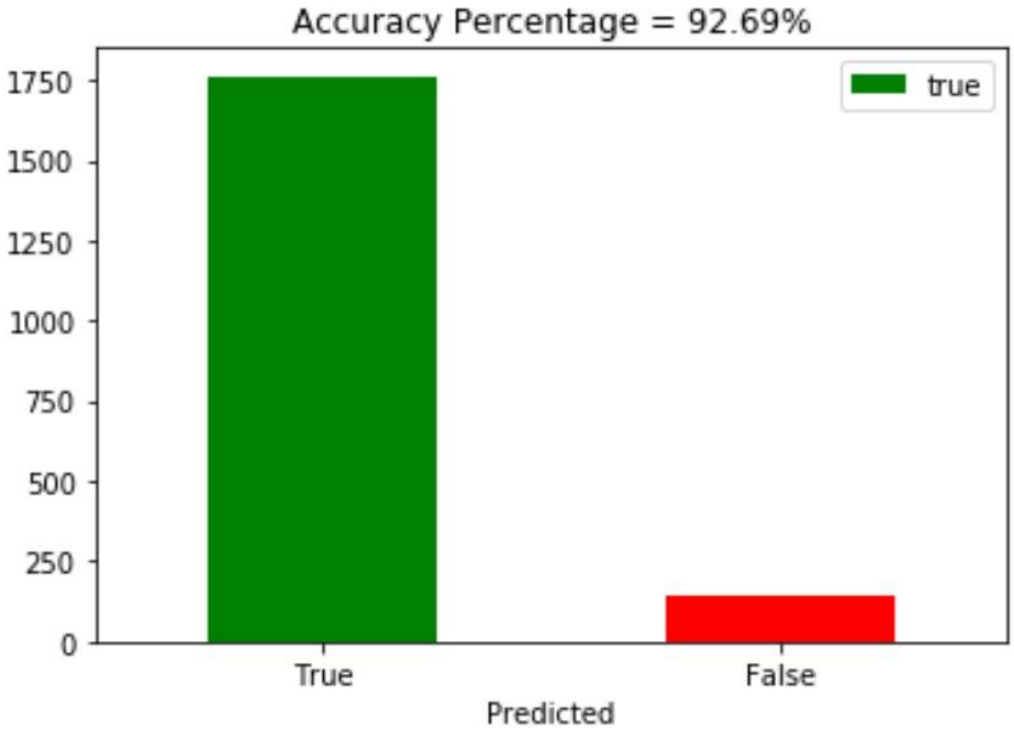
This is very useful in situations where there is a huge amount of data and it is **computationally infeasible** to train the entire dataset because of the sheer size of the data. We can simply say that an online-learning algorithm will **get** a training example, **update** the classifier, and then **throw away** the example.

A very good example of this would be to detect fake news on a social media website like Twitter, where new data is being added every second. To **dynamically** read data from Twitter continuously, the data would be huge, and using an online-learning algorithm would be the ideal choice.

Findings

As we can see, the model had a pretty good accuracy score of **92.69%**.

Seperating the true/false predictions into green and red respectively makes it more elegant and accessible.



Limitations

The idea is that right before a news is uploaded on the internet, a **short summary** of it has to go through a machine learning model which will then label the news real/fake accordingly. Since the model is **not 100% accurate**, **labelling** the news is preferred **instead** of **censoring/permanently deleting** the news which is predicted fake.

The **limitation** is that the model currently **does not** have **access** to all the news which is being posted around the world in **real time**. Additionally, for the **Passive Aggressive Classifier** to work **optimally**, there should be a **standard format**(length of text, example twitter's character limit of 280) of text in which the summary is sent to the machine learning model.

Conclusions

I conclude by saying that “**Yes**, through machine learning, we can predict if a news is real/fake with an **accuracy** of **92.69%** .”

Although it is **not 100% accurate**, when you consider the **sensitivity** of the use case 92.69% accuracy means the model is doing a pretty good job. With time and **more training** data, I believe the model will **increase** its accuracy even further.

Acknowledgements

I would like to thank the **course instructors** at **UC San Diego** for providing me the **opportunity** to exhibit my learnings in the means of a project.

I got the **dataset** on the **internet** while exploring various news datasets.

I would also like to thank my **parents** who **commented** on the **social impact** of the model. They opined that a label of Real/Fake for the news would definitely **benefit** news readers.

References

I had used the following page as a guidance to create my project.

https://www.pantechsolutions.net/fake-news-detection-using-machine-learning