SUBMITTED BY-

Janvi Pangoriya

B.Tech (Computer Science)-2nd year

G.L.A University, Mathura Uttar Pradesh

Mentor -

Mr. Bipul Shahi

**ABSTRACT**

Recent political events have lead to an increase in the popularity and spread of fake news. As demonstrated by the widespread effects of the large onset of fake news, humans are inconsistent if not outright poor detectors of fake news. With this, efforts have been made to automate the process of fake news detection. The most popular of such attempts include “blacklists” of sources and authors that are unreliable. While these tools are useful, in order to create a more complete end to end solution, we need to account for more difficult cases where reliable sources and authors release fake news. As such, the goal of this project was to create a tool for detecting the language patterns that characterize fake and real news through the use of machine learning and natural language processing techniques. The results of this project demonstrate the ability for machine learning to be useful in this task. We have built a model that catches many intuitive indications of real and fake news as well as an application that aids in the visualization of the classification decision.

**ACKNOWLEDGEMENT**

I have taken efforts in this project. However, it would not have been possible without the kind support and help of your organizations. I would like to extend my sincere thanks to all of them.

I am highly indebted to Diginique Techlabs for their guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in completing the project.

I would like to express my gratitude towards my mentor Mr. Bipul Shahi of Diginique Techlabs for his kind co-operation and encouragement which help me in completion of this project.

I would like to express my special gratitude and thanks to industry persons for giving me such attention and time.

My thanks and appreciations also go to my colleague in developing the project and people who have willingly helped me out with their abilities.

I also declare to the best of my knowledge and belief that the Project Report has not been submitted anywhere else.

Thank You!! Janvi Pangoriya

**INTRODUCTION**

As an increasing amount of our lives is spent interacting online through social media platforms, more and more people tend to seek out and consume news from social media rather than traditional news organizations. The reasons for this change in consumption behaviors are inherent in the nature of these social media platforms: (i) it is often more timely and less expensive to consume news on social media compared with traditional news media, such as newspapers or television; and (ii) it is easier to further share, comment on, and discuss the news with friends or other readers on social media. The extensive spread of fake news can have a serious negative impact on individuals and society. First, fake news can break the authenticity balance of the news ecosystem. For example, it is evident that the most popular fake news was even more widely spread on Facebook than the most popular authentic mainstream news during the U.S. 2016 president election4. Second, fake news intentionally persuades consumers to accept biased or false beliefs. Fake news is usually manipulated by propagandists to convey political messages or influence. For example, some report shows that Russia has created fake accounts and social bots to spread false stories5. Third, fake news changes the way people interpret and respond to real news. For example, some fake news was just created to trigger people’s distrust and make them confused; impeding their abilities to differentiate what is true from what is not.



**Problem statement**

In this project, I am focus on the automatic identification of fake content in online news. In this we have describe the collection, annotation and validation process in detail and present several exploratory analysis on the identification of linguistic differences in fake and real news content. Second we conduct a set of learning experiments to build accurate .

In this paper,

we focus on the automatic identiﬁcation of

fake content in online news. Our contri-

bution is twofold. First, we introduce two

novel datasets for the task of fake news de-

tection, covering seven different news do-

mains. We describe the collection, anno-

tation, and validation process in detail and

present several exploratory analysis on the

identiﬁcation of linguistic differences in

fake and legitimate news content. Second,

we conduct a set of learning experiments

to build accurate fake news detectors. In

addition, we provide comparative analyses

of the automatic and manual identiﬁcation

of fake news

In this paper,

we focus on the automatic identiﬁcation of

fake content in online news. Our contri-

bution is twofold. First, we introduce two

novel datasets for the task of fake news de-

tection, covering seven different news do-

mains. We describe the collection, anno-

tation, and validation process in detail and

present several exploratory analysis on the

identiﬁcation of linguistic differences in

fake and legitimate news content. Second,

we conduct a set of learning experiments

to build accurate fake news detectors. In

addition, we provide comparative analyses

of the automatic and manual identiﬁcation

of fake news

In this paper,

we focus on the automatic identiﬁcation of

fake content in online news. Our contri-

bution is twofold. First, we introduce two

novel datasets for the task of fake news de-

tection, covering seven different news do-

mains. We describe the collection, anno-

tation, and validation process in detail and

present several exploratory analysis on the

identiﬁcation of linguistic differences in

fake and legitimate news content. Second,

we conduct a set of learning experiments

to build accurate fake news detectors. In

addition, we provide comparative analyses

of the automatic and manual identiﬁcation

of fake news.

In this paper,

we focus on the automatic identiﬁcation of

fake content in online news. Our contri-

bution is twofold. First, we introduce two

novel datasets for the task of fake news de-

tection, covering seven different news do-

mains. We describe the collection, anno-

tation, and validation process in detail and

present several exploratory analysis on the

identiﬁcation of linguistic differences in

fake and legitimate news content. Second,

we conduct a set of learning experiments

to build accurate fake news detectors. In

addition, we provide comparative analyses

of the automatic and manual identiﬁcation

of fake news.

In this paper,

we focus on the automatic identiﬁcation of

fake content in online news. Our contri-

bution is twofold. First, we introduce two

novel datasets for the task of fake news de-

tection, covering seven different news do-

mains. We describe the collection, anno-

tation, and validation process in detail and

present several exploratory analysis on the

identiﬁcation of linguistic differences in

fake and legitimate news content. Second,

we conduct a set of learning experiments

to build accurate fake news detectors. In

addition, we provide comparative analyses

of the automatic and manual identiﬁcation

of fake news.

Technology and Concepts

Machine Learning

Machine Learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

Machine learning algorithms are often categorized as supervised or unsupervised*.*

* Supervised learning
* Unsupervised learning
* Reinforced learning

Supervised learning: Supervised machine learning can take what it has learned in the past and apply that to new data using labeled examples to predict future patterns and events. It learns by explicit example. Supervised learning requires that the algorithm’s possible outputs are already known and that the data used to train the algorithm is already labeled with correct answers. It’s like teaching a child that 2+2=4 or showing an image of a dog and teaching the child that it is called a dog. The approach to supervised machine learning is essentially the same – it is presented with all the information it needs to reach pre-determined conclusions. It learns how to reach the conclusion, just like a child would learn how to reach the total of ‘5’ and the few, pre-determined ways to get there, for example, 2+3 and 1+4. If you were to present 6+3 as a way to get to 5, that would be determined as incorrect. Errors would be found and adjusted. Supervised learning is further divided into:

1. Classification :   refers to a predictive modeling problem where a class label is predicted for a given **example** of input data. **Examples** of **classification** problems include: Given an **example**, **classify** if it is spam or not. Given a handwritten character, **classify** it as one of the known characters.
2. Regression : models are used to predict a continuous value. Predicting prices of a house given the features of house like size, price etc is one of the common **examples** of **Regression**.

Unsupervised learning :Supervised learning tasks find patterns where we have a dataset of “right answers” to learn from. **Unsupervised learning tasks find patterns where we don’t.** This may be because the “right answers” are unobservable, or infeasible to obtain, or maybe for a given problem, there isn’t even a “right answer” per se. Unsupervised learning is used against data without any historical labels. The system is not given a pre-determined set of outputs or correlations between inputs and outputs or a “correct answer.” The algorithm must figure out what it is seeing by itself, it has no storage of reference points. The goal is to explore the data and find some sort of patterns of structure. Unsupervised learning works well when the data is transactional. For example, identifying pockets of customers with similar characteristics who can then be targeted in marketing campaigns. Unsupervised machine learning is a more complex process and has been used far fewer times than supervised machine learning. But it’s exactly for this reason that there is so much buzz around the future of AI. Advances in unsupervised ML are seen as the future of AI because it moves away from narrow AI and closer to AGI (‘artificial general intelligence’ that we discussed a few paragraphs earlier). If you’ve ever heard someone talking about computers teaching themselves, this is essentially what they are referring to.

### **Reinforcement learning :** Reinforcement learning is a type of dynamic programming that trains algorithms using ****a** system of reward and punishment**.**** A reinforcement learning algorithm, or agent, learns by interacting with its environment. It receives rewards by performing correctly and penalties for doing so incorrectly. Therefore, it learns without having to be directly taught by a human –**it learns by seeking the greatest reward and minimizing penalty.** This learning is tied to a context because what may lead to maximum reward in one situation may be directly associated with a penalty in another.

DATASET DESCRIPTION

The dataset we’ll use for this python project- we’ll call it news.csv. This dataset has a shape of 6335×4. The first column identifies the news, the second and third are the title and text, and the fourth column has labels denoting whether the news is REAL or FAKE.



As shown above are the names of the columns in our dataset.

**DATA CLEANING**

Languages we speak and write are made up of several words often derived from one another and can contain words which don’t add meaning or context. In order to clean the data, we implemented 5 approaches.

Lemmatization: Lemmatization is the process of grouping together the different inflected forms of a word so they can be analyzed as a single item.

Stop words removal: Stop words are usually articles or prepositions which do not help us to find the context or the true meaning of a sentence. For example, ‘is’ and ‘are’.

Stemming: With stemming, words are reduced to their word stems by cutting off the end or the beginning of the word, taking into account a list of common prefixes and suffixes that

can be found in an inflected word.

Greek characters: We next removed special characters and numbers.

Slang words: Slang words include informal short forms of words which are usually used in speech.

EXTRACTING FEATURES

From the below techniques i.e. Term Frequency Inverse Document Frequency , we will extract the features from the text and label present in our datasets.

TFIDF (Term Frequency Inverse Document Frequency) :

TFIDF is a numerical statistic that is intended to reflect how important a word is to a document in a collection. It is used as a weighting factor in searches of information retrieval, text mining, and user modeling.TF\*IDF is an information retrieval technique that weighs a term’s frequency (TF) and its inverse document frequency (IDF). Each word or term has its respective TF and IDF score. The product of the TF and IDF scores of a term is called the TF\*IDF weight of that term. Put simply, the higher the TF\*IDF score (weight), the rarer the term and vice versa. The TF\*IDF algorithm is used to weigh a keyword in any content and assign the importance to that keyword based on the number of times it appears in the document. More importantly, it checks how relevant the keyword is throughout the web, which is referred to as corpus. For a term t in a document d, the weight Wt, d of term t in document d is given by: Wt, d = TFt, d log (N/ DFt) Where:

● TFt, d is the number of occurrences of t in document d.

● DFt is the number of documents containing the term t.

● N is the total number of documents in the corpus.

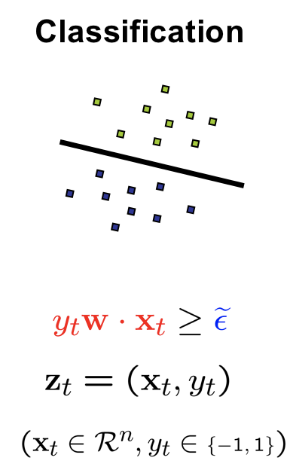
The TF (term frequency) of a word is the frequency of a word (i.e. number of times it appears) in a document. When you know it, you’re able to see if you’re using a term too much

or too little.

For example, when a 1000-word document contains the term “cat” 12 times, the TF for the word ‘cat’ is TFcat = 12/1000 i.e. 0.012. The IDF (inverse document frequency) of a word is the measure of how significant that term is in the whole corpus.

Algorithm:

Passive Aggressive Classifier : In this we have instances in R^n and labels. The goal is to find a hyperplane that separates the positive instances from the negative instances, that is, we would like that y**t**(w) \* x**t** will have a large positive value. By “large” we mean at least epsilon e we see in figure. We denote the instance-label pair by z.



Term frequency vectorizer (tf):

The term frequency is often divided by the document length (the total number of terms in the document) as a way of normalization:

**TF**(t) = Number of times term t appears in a document

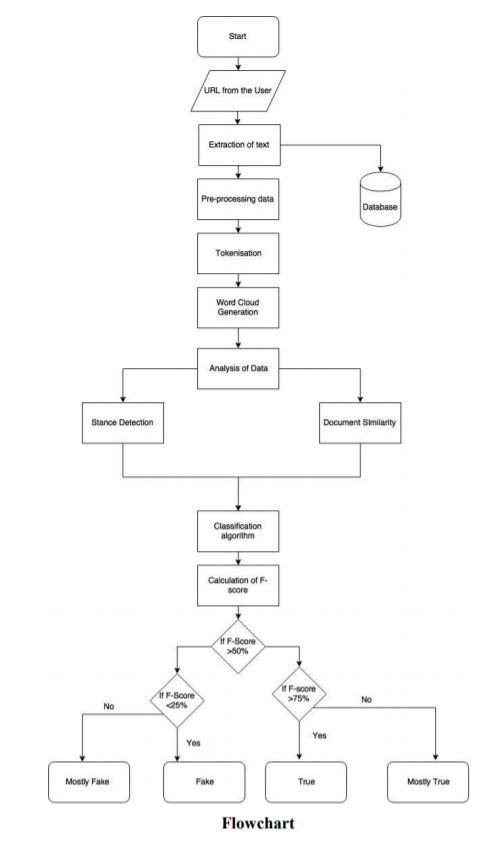
(Total number of terms in the document).

Inverse document frequency vectorizer (idf):

**IDF** are word frequency scores that try to highlight words that are more interesting, e.g. frequent in a document but not across documents.

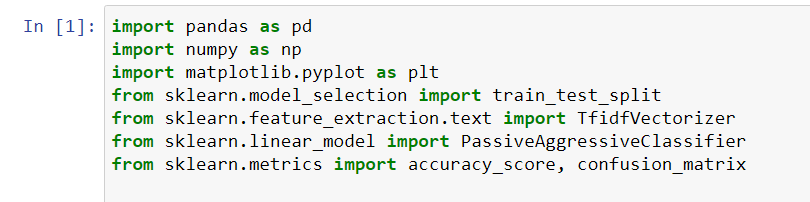
**The TfidfVectorizer will tokenize documents, learn the vocabulary and inverse document frequency weightings, and allow you to encode new documents.**

Steps to be followed…..



Steps:

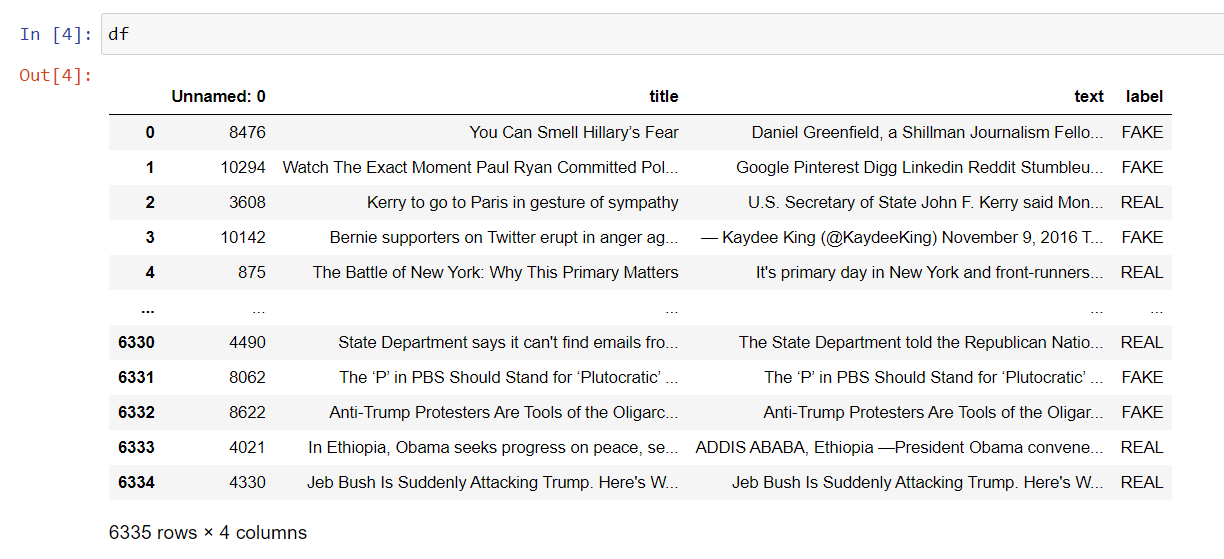
* Importing all the necessary libraries required for this project



* Importing the dataset which consist of features and labels on which we are going to train our model.



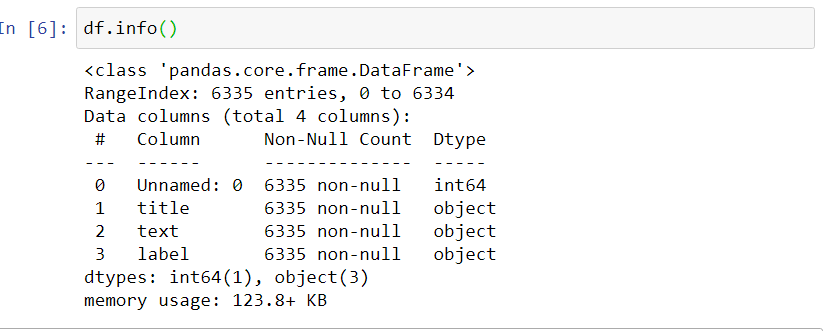
* Displaying the dataset which gives us idea how it looks like



* Checking the characteristics of the dataset



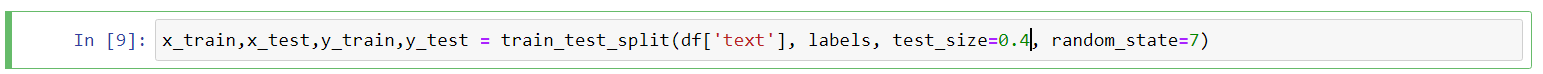
* Getting information about the dataset



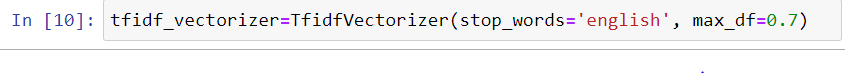
* Getting the labels from the datasets



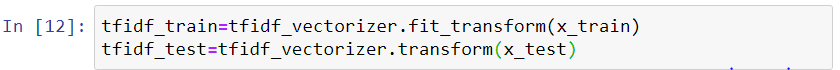
* Split the dataset into training and testing sets.in which we are using trainmg data is 60% of the total data and 40% of the total data as testing data



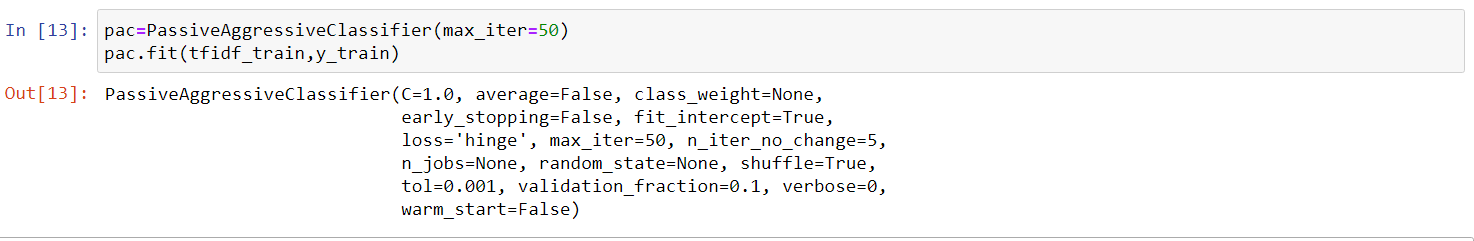
* Now we will initialize a [TfidfVectorizer](https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html) with stop words from the English language and a maximum document frequency of 0.7 (terms with a higher document frequency will be discarded). Stop words are the most common words in a language that are to be filtered out before processing the natural language data. And a TfidfVectorizer turns a collection of raw documents into a matrix of TF-IDF features.



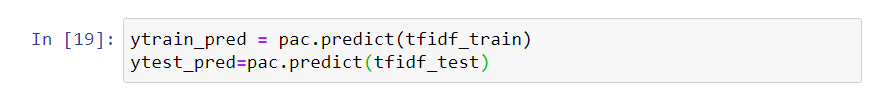
* Now, fit and transform the vectorizer on the train set, and transform the vectorizer on the test set.



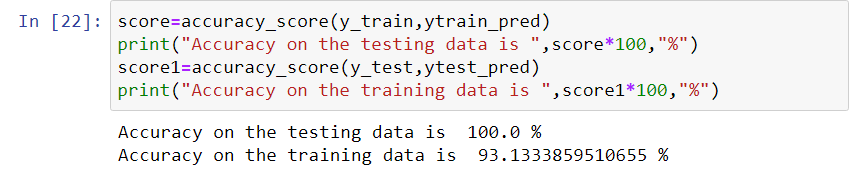
* Next, we’ll initialize a PassiveAggressiveClassifier. This is. We’ll fit this on tfidf\_train and y\_train.



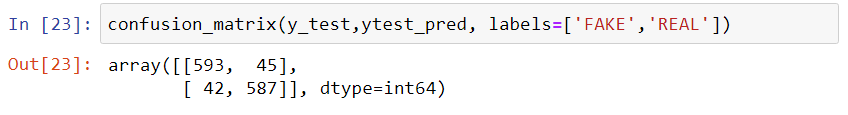
* Now we will perform prediction over the training as well as testing data



* Now we are calculating the accuracy of testing as well as training data



* Now we are going to plot confusion matrix of the testing datato check the the status of the prediction



So with this model, we have 593 true positive and 587 false positive means we are getting the right prediction and 42 false positives and 45 false negatives means we are not getting the right predicition.

Applications:

There are many field in which this model can be used :

* **Social Media Services**
* **Email Spam and Malware Filtering**
* **Online Customer Support**
* **Search Engine Result Refining**
* **Online Fraud Detection**

**Conclusions :**

In the 21st century, the majority of the tasks are done online. Newspapers who were earlier preferred as hard-copies are now being substituted by applications like Facebook, Twitter, and news articles to be read online. The growing problem of fake news only makes things more complicated and tries to change or hamper the opinion and attitude of people towards use of digital technology. When a person is deceived by the real news two possible things happen. People start believing that their perceptions about a particular topic are true as assumed. Another problem is that even if there is any news article available which contradicts a supposedly fake one, people believe in the words which just support their thinking without taking in the measure the facts involved.

This project paper has given an algorithm showing its implementation to detect a fake news and a real one. Not only this single algorithm but we can use many other machine learning algorithm like **CountVectorizer and LogisticRegression model** to implement the same model. All they differ is in accuracy.

In this model I have taken 60% of the data as training data for the machine and 40% data for testing of the model and got 93.13% accuracy.

References :

* 1. <https://www.kaggle.com/c/fake-news/data> : from here I gathered the dataset for my project.
  2. <https://scikitlearn.org/stable/modules/generated/sklearn.linear_model.PassiveAggressiveClassifier.html>
  3. <https://github.com/JanviGupta12/Fake_News> : This is the official repository where my jupyter notebook is uploaded regarding this project.