

CS5624 Natural Language Processing

Fake News Detection

Group number - 17

Group members:

111901032 Mohammed Abdul Azeem

111901052 Vikas Naik Dharavath

111901015 Bandi Kaushik

Introduction

Fake news is a type of yellow journalism, it encapsulates pieces of news that may be hoaxes and is generally spread through social media and other online media. Fake news was present way before social media was introduced, but it got multiplied when social media got introduced. Fake news is the type of news that is deliberately created to spread rumors and misinformation. Fake news is generally spread through social media platforms.

As you know, social media is used today to read the news. Previously, experts collected and published news, but now anyone can collect and publish news. As a result, the number of fake websites on the Internet is increasing, and fake news articles are more likely to be published daily than before, making it difficult to verify the authenticity of the news. In addition, people are benefiting from Clickbait and the publication of fake news on the Internet.

Major Problem of Fake News

These days when we are scrolling through web pages and clicking on clickbait, users are led to a page that contains false information. Fake news affects people's perceptions. The rise of fake news has become a global problem, and even major tech companies such as Facebook and Google are struggling to resolve it. It can be difficult to determine if a text is true without additional context and human judgment.

Purpose

The purpose of this project is to develop a method for detecting and classifying fake news using natural language processing. The main goal is

- To identify fake news, which is a classic text classification problem.

- We collected the data, preprocessed the text by using the stemming process and also by removing stop words, and transformed the article into a supervised model function.
- To develop a model that classifies certain items as Fake or Real.
- To identify the best classifier for this problem with the help of accuracy, precision, recall, and f1_score.

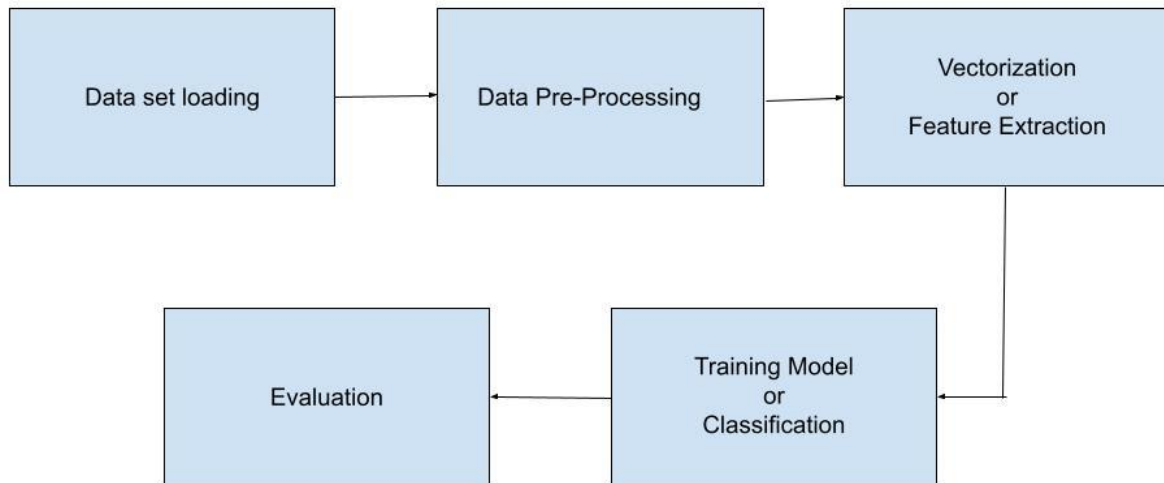
So in this project, we propose a machine learning model to identify whether a news article is fake or not based on its content. We also used NLP to detect fake news. There are several attempts to make use of AI technologies, specifically machine learning techniques and NLP to detect fake news and stop it from spreading. As human beings, we can read a sentence or a paragraph and we can interpret the words with the whole document and understand the context. Given today's volume of news, it is possible to teach a computer how to read and understand the difference between Real and Fake news using NLP. In this project, the main building blocks are Dataset and Machine learning algorithms such as some classifiers. We got our dataset from the website mentioned below.

https://www.uvic.ca/ecs/ece/isot/datasets/fake-news/index.php?utm_medium=redirect&utm_source=/engineering/ece/isot/datasets/fake-news/index.php&utm_campaign=redirect-usage

Steps Involved

- Data set loading
- Data Pre-Processing (removing Stopwords, Stemming, Dropping duplicates, and removing meaningless char from the text.)
- Vectorization (representing in numeric format)
- Applying Classification and model construction (to predict the output of unlabelled data)
- Evaluation

In this, we use the NLTK toolkit, which includes a number of libraries and many NLP-oriented programs. Even the machine learning algorithms for data clustering, regression, and classification have been imported, e.g. B. Scikit Learn. These libraries are important factors in the program, which was developed in combination with other libraries such as NumPy. In this project, Natural Language Processing (NLP) was used as a computational tool. The PANDAS library was used for natural language processing and analysis.



Dataset Loading:

We first load the two CSV files. Here we have two types of datasets in which one file has all the data about the news that are Real and the other has all the data about the news that are Fake. So we added a column “Label” to both the datasets and added Real and Fake data to that column in the files True.csv and Fake.csv respectively.

	title	text	subject	date	label
343	Trump criticizes treatment of Flynn as unfair	WASHINGTON (Reuters) - President Donald Trump ...	politicsNews	December 4, 2017	Real
7914	President Obama Announces The Major, Unfulfil...	President Obama delivered an address from the ...	News	February 23, 2016	Fake
971	This Really Happened: Fox Interrupted Trump T...	If today on Fox News was any indication at all...	News	June 29, 2017	Fake
8240	U.S. warns Russia against interfering with Wes...	OXFORD, England (Reuters) - U.S. Defense Secre...	politicsNews	September 7, 2016	Real
17950	Former Salvadoran president Calderon dies at 69	SAN SALVADOR (Reuters) - Former Salvadoran Pre...	worldnews	October 9, 2017	Real
14290	WATCH HILLARY LIE About Libya To Supporters: "...	The families of the four dead Americans you le...	politics	Mar 15, 2016	Fake
11760	PRESIDENT TRUMP SHOCKS PRESS CORP...Makes Unanno...	President Trump bolted from the White House We...	politics	Feb 1, 2017	Fake
11411	JUDGE NAPOLITANO: THREE INTEL SOURCES Have Dis...	JUDGE NAPOLITANO was on FOX & Friends this mor...	politics	Mar 14, 2017	Fake
3193	CNN Host STUNNED As Trump's Team REFUSES To C...	Donald Trump s team certainly doesn t make it ...	News	January 2, 2017	Fake
17890	Jared Kushner NEVER Registered To Vote As A "F...	Meanwhile, as President Trump continues to mee...	left-news	Sep 29, 2017	Fake
12184	OOPS! OBAMA Tells Troops No Foreign Terror Att...	In a speech at MacDill Air Force Base in Tamp...	politics	Dec 9, 2016	Fake
468	Trump Just Threw A Colossal Hissy Fit On Twit...	Donald Trump has the maturity of a toddler, an...	News	August 30, 2017	Fake
2616	Trump's Own SCOTUS Nominee Just Threw Him Und...	Judge Neil Gorsuch, Trump s nominee for the Su...	News	February 8, 2017	Fake
3938	FBI's Comey on Trump travel ban: 'citizenship ...	WASHINGTON (Reuters) - FBI Director James Come...	politicsNews	May 3, 2017	Real
12613	WOMAN INTRODUCING HILLARY Refuses To Say "One ...	Watch Hillary laugh when the woman who is intr...	politics	Oct 26, 2016	Fake
15712	U.S. APOLOGIZES FOR HUMAN RIGHTS VIOLATIONS AT...	Obama s apology tour for the greatest country ...	politics	May 11, 2015	Fake
3472	Top House Intel panel Democrat rebukes chairma...	WASHINGTON (Reuters) - The U.S. House Intellig...	politicsNews	June 1, 2017	Real

fake.csv

	title	text	subject	date	label
0	As U.S. budget fight looms, Republicans flip t...	WASHINGTON (Reuters) - The head of a conservat...	politicsNews	December 31, 2017	Real
1	U.S. military to accept transgender recruits o...	WASHINGTON (Reuters) - Transgender people will...	politicsNews	December 29, 2017	Real
2	Senior U.S. Republican senator: 'Let Mr. Muell...	WASHINGTON (Reuters) - The special counsel inv...	politicsNews	December 31, 2017	Real
3	FBI Russia probe helped by Australian diplomat...	WASHINGTON (Reuters) - Trump campaign adviser ...	politicsNews	December 30, 2017	Real
4	Trump wants Postal Service to charge 'much mor...	SEATTLE/WASHINGTON (Reuters) - President Donal...	politicsNews	December 29, 2017	Real
5	White House, Congress prepare for talks on spe...	WEST PALM BEACH, Fla./WASHINGTON (Reuters) - T...	politicsNews	December 29, 2017	Real
6	Trump says Russia probe will be fair, but time...	WEST PALM BEACH, Fla (Reuters) - President Don...	politicsNews	December 29, 2017	Real
7	Factbox: Trump on Twitter (Dec 29) - Approval ...	The following statements were posted to the ve...	politicsNews	December 29, 2017	Real
8	Trump on Twitter (Dec 28) - Global Warming	The following statements were posted to the ve...	politicsNews	December 29, 2017	Real
9	Alabama official to certify Senator-elect Jone...	WASHINGTON (Reuters) - Alabama Secretary of St...	politicsNews	December 28, 2017	Real
10	Jones certified U.S. Senate winner despite Moo...	(Reuters) - Alabama officials on Thursday cert...	politicsNews	December 28, 2017	Real
11	New York governor questions the constitutional...	NEW YORK/WASHINGTON (Reuters) - The new U.S. t...	politicsNews	December 28, 2017	Real
12	Factbox: Trump on Twitter (Dec 28) - Vanity Fa...	The following statements were posted to the ve...	politicsNews	December 28, 2017	Real
13	Trump on Twitter (Dec 27) - Trump, Iraq, Syria	The following statements were posted to the ve...	politicsNews	December 28, 2017	Real
14	Man says he delivered manure to Mnuchin to pro...	(In Dec. 25 story, in second paragraph, corre...	politicsNews	December 25, 2017	Real

true.csv

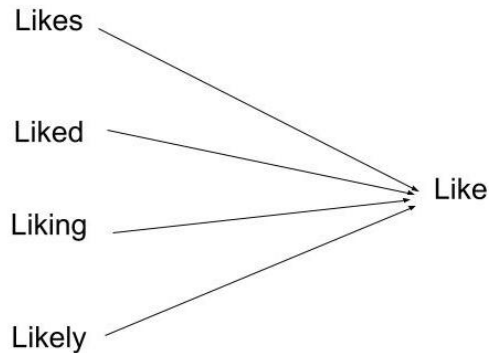
And finally merged both datasets and scattered the data and used the result of that merged data as our dataset. This final dataset consists of 44898 rows and 5 columns. The first column is "Title" which is the headline of the news, the second column is "Text", the third column is "Subject", the fourth column is "Date" and the last column is "Label" with two classes Fake and Real.

Then we declared two variables x and y. Where we store the column text and label in the variables x and y respectively.

Data Pre-Processing:

Due to the nature of the dataset used, it contains a lot of noise as it is a natural language. To make the data apt for the algorithms to work on, after storing the data in a variable, we then preprocess the data which is stored in variable x. This step performs stemming, removing stopwords, duplicate drops, and even removing nonsensical characters in the text, such as exclamation marks and commas. Stemming is the process of reducing a word or token to its root/basic words. This is used to reduce the forms of words in data. For example, if you use words such as actor, actress, or acting, the root word is "act", so all prefixes and suffixes are removed, resulting in "act."

STEMMING



Stopwords are basically words that add value to other words or define a relation between words. They can largely contain adjectives, adverbs, prepositions, conjunctions, and determiners. Since our data set consists of multiple elements, it is imminent to remove these stop words before feeding the data as input to the classifiers. These words include, for example, a, an, another, nor, but, or, to, still, in, etc. To do operations, I created a function called preprocessing. Therefore, give this function the variable `x` as input. This function in the first step removes everything except the alphabet as they can be small or large. Remove all number punctuation, etc. Then convert everything to lowercase and split the data into a list. We used a portal stemmer library which is imported from the NLTK toolkit, then stem each word and remove the stopword at the same time. Then connect all the words and return the processed data.

After we process the data, we split the two variables into training and testing sets i.e. into `x_train`, `x_text`, `y_train`, and `y_text`. Here we used 80% of the data for training the machine learning model and the remaining data to test the model that is trained.

Vectorization or Feature extraction :

After splitting the data we perform vectorization where we convert the textual data into their respective feature vectors. Feature vectors are basically numerical data. In this project, we use a vectorizer called TFIDF vectorizer. Here TF stands for Term Frequency and IDF stands for Inverse Document frequency. So Term Frequency basically counts the number of times a particular word is repeated in a document or in a paragraph or in text. So the repetition tells the model that it is a very important word and it assigns a particular

numerical value to that word which is the main purpose of TF. Inverse Document Frequency, where sometimes a word is repeated several times and doesn't have meaning in it. In simple words, we can say words that occur many times in a document but also occur many times in other documents and may be irrelevant. So IDF is a measure of how significant a term is in the entire Corpus.

$$w_{i,j} = tf_{i,j} \times \log \left(\frac{N}{df_i} \right)$$

tf_{ij} = number of occurrences of i in j
 df_i = number of documents containing i
 N = total number of documents

Therefore, TF and IDF create feature vectors with the given word so that the machine can understand it. So here we are performing TF-IDF vectorization on 'x_train' and 'x_test' variables.

In this process, we fit and transform the data in the two variables x_train and x_test into feature vectors based on the function of TF-IDF vectorization and store them in tfidf_train and tfidf_test. Now we are feeding our data to the machine learning models.

Applying Classification and model construction:

After applying TF-IDF vectorization on x_train and x_test, converting them into the feature vectors and we store them in tfidf_train and tfidf_test, we then train the model using different classification models. Some of the classification models we used are Passive Aggressive Classifier, Multinomial Naive Bayes Classifier, K Nearest Neighbor Classifier, and Logistic Regression Classifier.

1. Multinomial Naive Bayes Classifier:

It is a Naive Bayes classifier for multinomial models. It is suitable for classification with discrete features. It guesses the tag of a text using the Bayes

theorem and calculates each tag's likelihood for a given sample and outputs the tag with the greatest chance.

2. Passive Aggressive Classifier :

Passive Aggressive Classifier belongs to the online learning algorithm in Machine Learning. It remains Passive for a correct classification outcome and turns aggressive in the event of miscalculations, updating, and adjusting. Unlike most algorithms, it doesn't converge.

3. KNN Classifier:

KNN Classifier determines the class of a data point by the majority voting principle. It calculates the distance from all points in the proximity of the unknown data and filters out the ones with the shortest distances to it. Based on the distance it classifies whether the data is fake or real.

4. Logistic Regression :

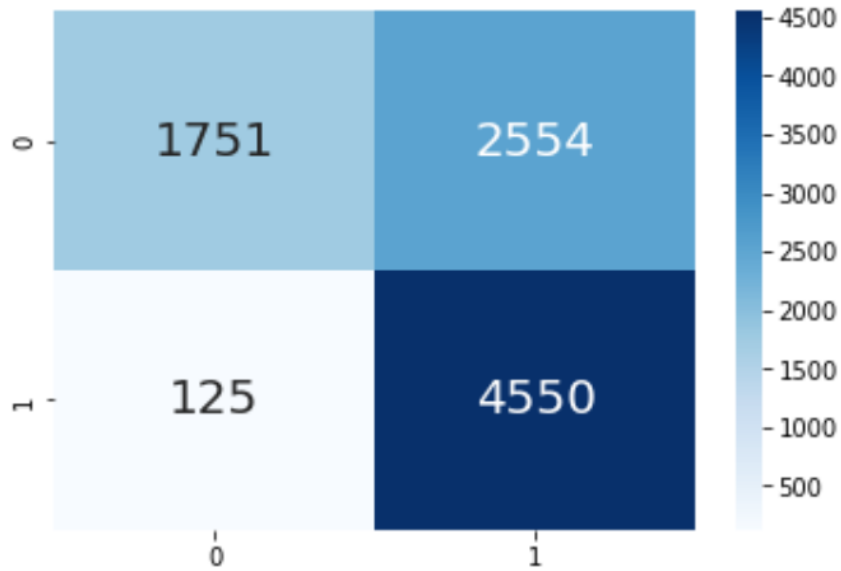
Logistic Regression is a classification algorithm used to assign observations to a discrete set of classes. It is also used to predict the probability of a target variable.

First, we load every classifier with help of inbuilt libraries and train the classifier model with the `tfidf_train` and `y_train` data. After training the classifier model we test the performance of the model using the `tfidf_test` data.

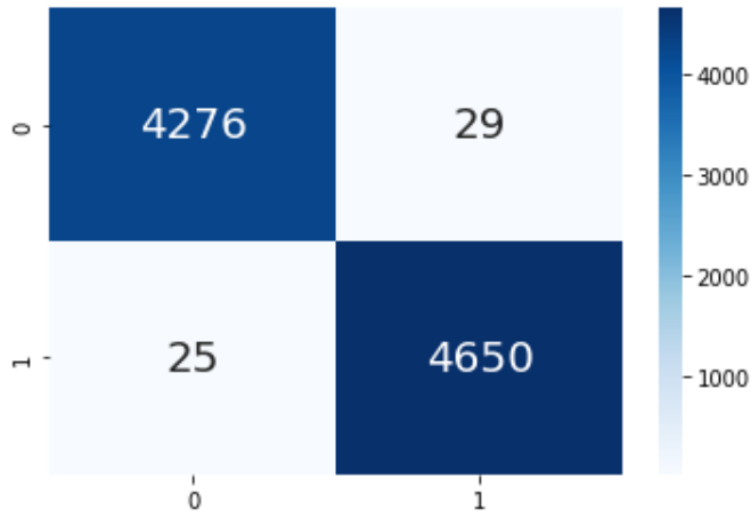
Evaluation :

After we train the classifier model we test the model with `tfidf_test`. After testing the data we then predict the output with the given `tfidf_test`. And also find the Accuracy, Precision, Recall, and `f1_score` of all the classifier models by comparing them with the `t_test`, so that we can know how well they were able to classify the given dataset is compared. For this purpose, we used the confusion matrix. A confusion matrix shows the number of incorrect classifications and correct classifications made by the model. In the confusion matrix, the top left section labels the data that has been correctly classified as Fake, which is True Negatives, the top right section labels the data that has been incorrectly classified as Real news, referred to as False Positives, The bottom left section labels the articles that have been incorrectly classified as Fake news, referred to as False Negative, the bottom right section labels the articles that have been correctly classified as true news, referred to as True Positive.

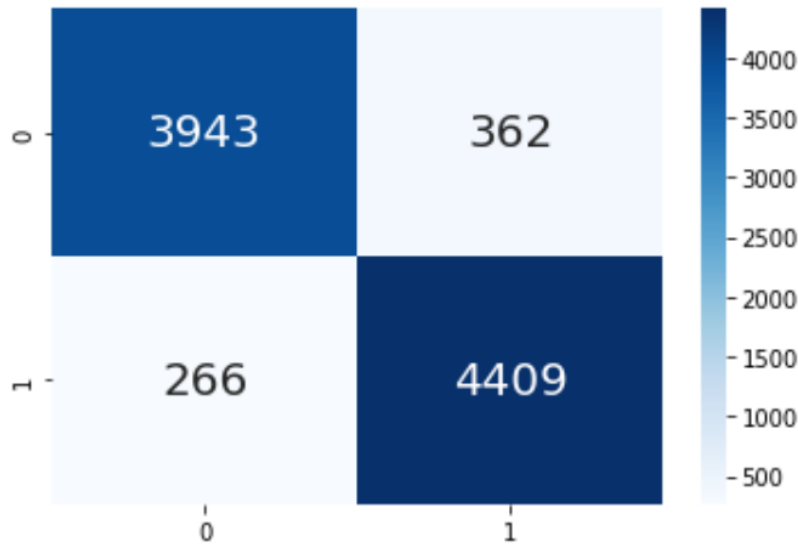
Accuracy for KNN Classifier: 70.16703786191538 %
Precision for KNN Classifier: 64.04842342342343 %
Recall for KNN Classifier: 97.32620320855615 %
f1_score for KNN Classifier: 77.25613379743612 %
<matplotlib.axes._subplots.AxesSubplot at 0x7fe3e3f24950>



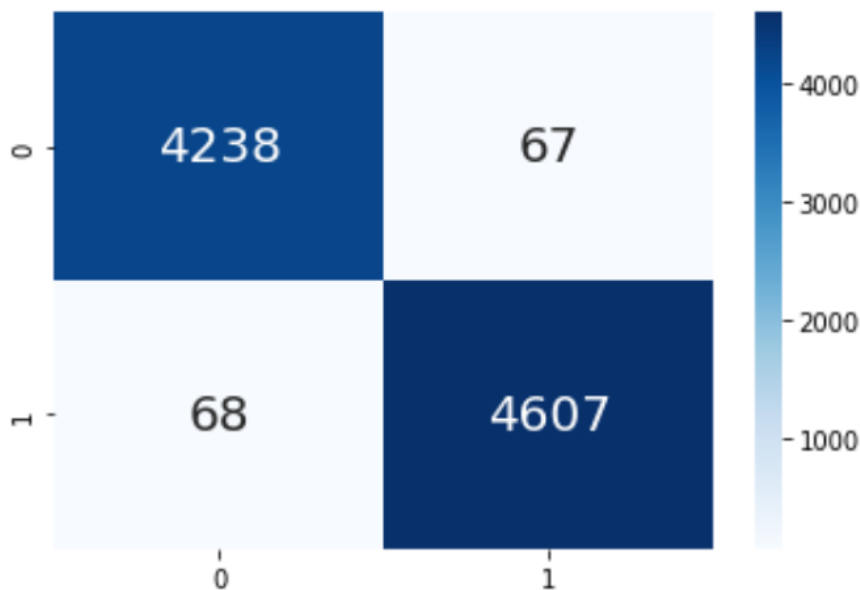
Accuracy for Passive Agressive Classifier: 99.39866369710467 %
Precision for Passive Agressive Classifier Classifier: 99.38020944646291 %
Recall for Passive Agressive Classifier Classifier: 99.46524064171123 %
f1_score for Passive Agressive Classifier Classifier: 99.42270686337396 %
<matplotlib.axes._subplots.AxesSubplot at 0x7fe3e3d03c50>



Accuracy for Multinomial Naive Bayes Classifier: 93.00668151447663 %
Precision for Multinomial Naive Bayes Classifier: 92.41249214001257 %
Recall for Multinomial Naive Bayes Classifier: 94.31016042780749 %
f1_score for Multinomial Naive Bayes Classifier: 93.35168325217023 %
<matplotlib.axes._subplots.AxesSubplot at 0x7fe3e2c4e4d0>



Accuracy for Logistic Regression Classifier: 98.4966592427617 %
Precision for Logistic Regression Classifier: 98.56653829696191 %
Recall for Logistic Regression Classifier: 98.54545454545455 %
f1_score for Logistic Regression Classifier: 98.55599529361429 %
<matplotlib.axes._subplots.AxesSubplot at 0x7fe3e344ca10>



As can be seen from the given confusion matrix for the respective classifiers, the Number of misclassified data is small, which makes it good to implement practically on a large data set. After implementing the classification on the models, the accuracy, precision, recall, and f1_score are estimated. It can be seen that all the classifiers have accuracy above 90% except the KNN classifier.

By looking at the accuracy score we can conclude that Passive Aggressive Classifier is best for this project because of its high accuracy and f1_score compared to other models.

CONCLUSION :

We used TF-IDF Vectorization and have implemented the Fake News Detection model using the Passive-Aggressive Classifier, Multinomial Naive Bayes Classifier, K Nearest Neighbor Classifier, and Logistic Regression Classifier. Passive Aggressive Classifier has high accuracy as compared to other classifier models and selected this model to detect fake news detection. So overall, the performance for our dataset was better with Passive Aggressive Classifier. Also, the confusion matrix has been plotted and Accuracy, Precision, Recall, and f1_score values have been measured for the performance of the analysis purpose.