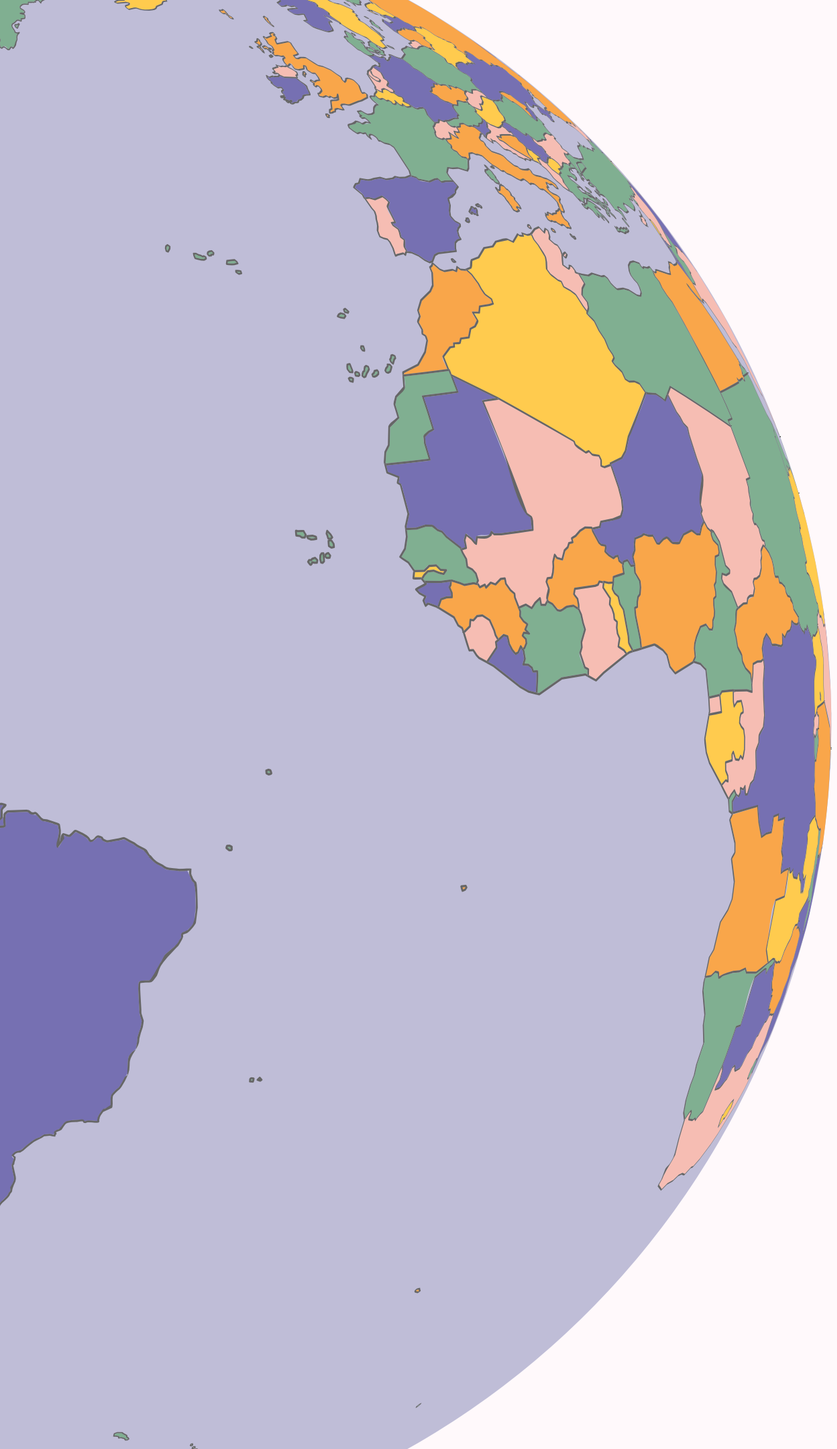
(Problem statement: To predict whether a news is fake or not)

Fake news detection

Gouravi Dorkhande - C2201988l917

Akshada Malpure - C2201988l934

Reena Prasad - C220l988l946



Objective

(what is addressed by our project:)

Fake news has become a huge issue in our digitally-connected world, the fake news spreads like wildfire and decieve people with wrong information. Our objective through this project is to provide with the insight that whether a news spread on WhatsApp is fake or authentic.

Modules used

The modules implemented in our project are:

numpy: for construction of array

panda: for formation of dataframe

nltk.corpus: for extracting english stopwords

nltk.stemporter: for stemming of words

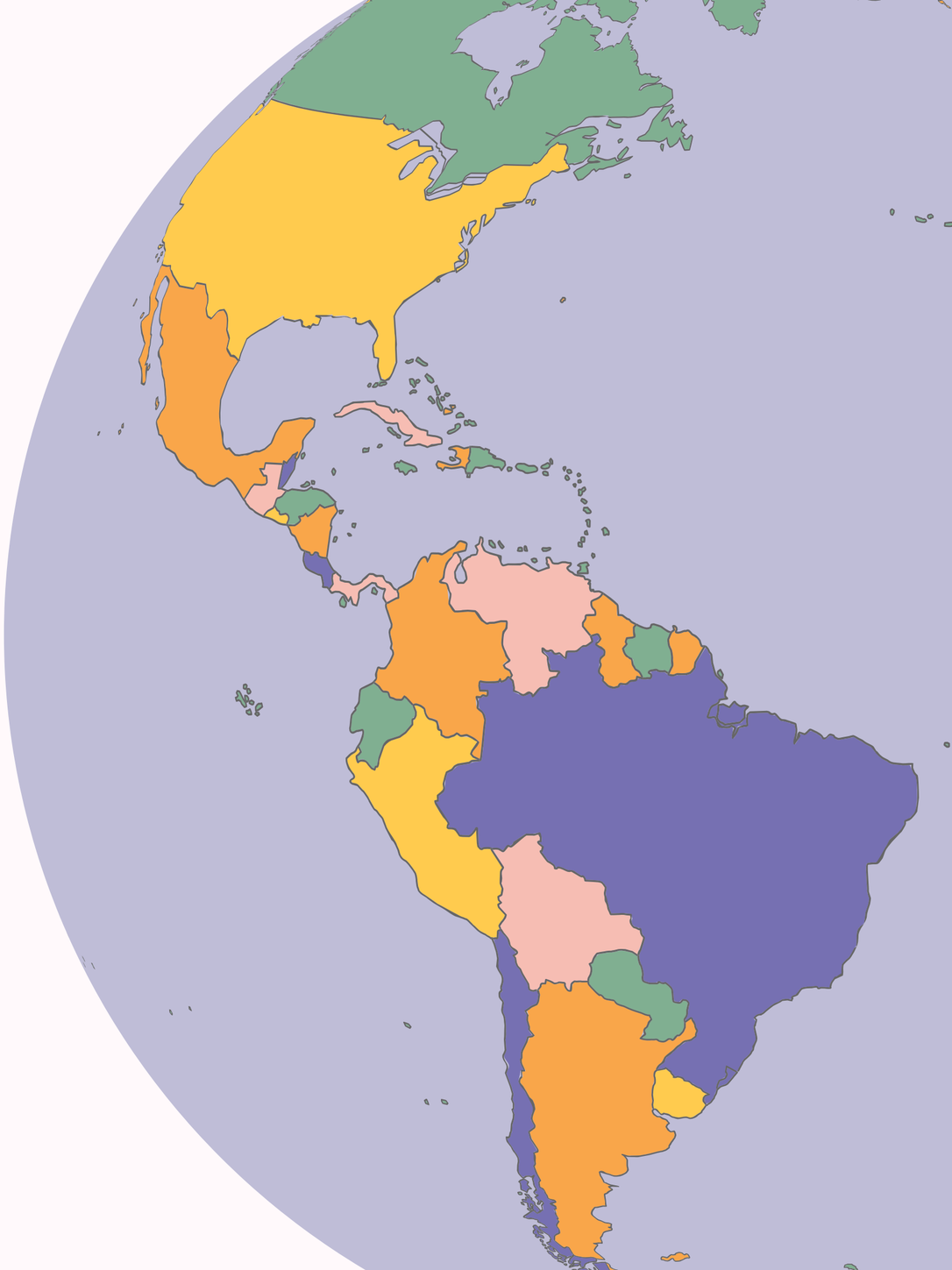
sklearn.feature_extraction.text: TfidfVectorizer is used to get tfidf values

sklearn.linear_model: train_test.split is used to form the training and testing data

sklearn.metrics: accuracy_score is used to get the accuracy score of the model

pickle: used for serialization and deserialization on data structures

pipeline: to pipeline our model



Step 1: preprocessing

Prepare the raw data and make it suitable for machine learning model. Preprocessing includes methods like tokenizing, stemming, spliting

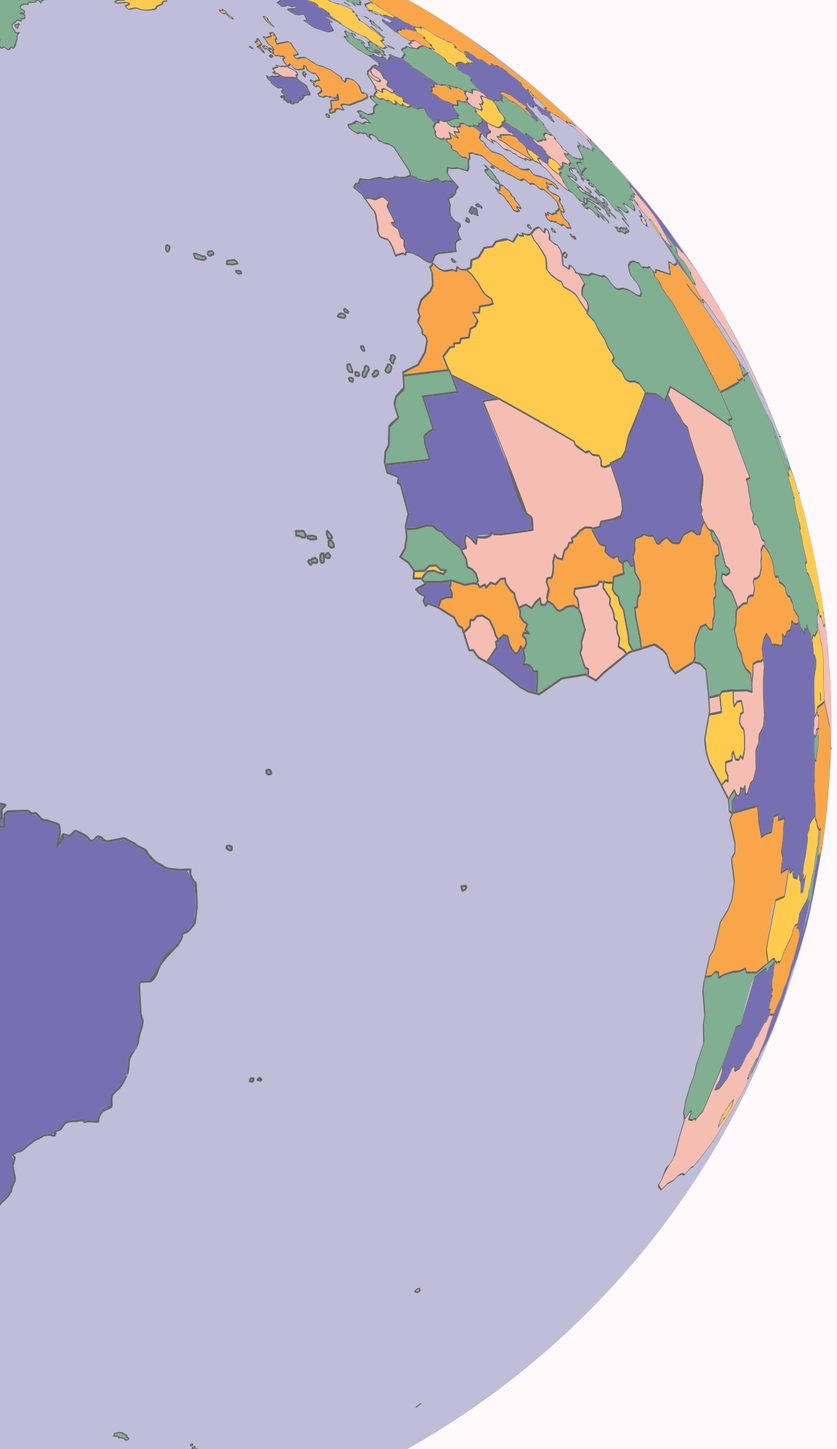
panda: to convert csv to 2D dataframe

Then we tried to find missing values and replace with null strings.

Since the content of news is very large performing on it is very tedious so we combined authour and title columns and performed further on it, also the accuracy was high.

PorterStemmer():Stemming done to convert all words to root words.

Tf-Idf performed: to assign weights to each word



Step 2: Classifier

We have split the dataset into train and test data in 80:20 ratio

Then we have used logistic regression because our class labels are binary so we used this model as it gives best accuracy

We first calculated accuracy of training data and verified it for the test data.

The accuracy was very close, hence we saved this model and sav file

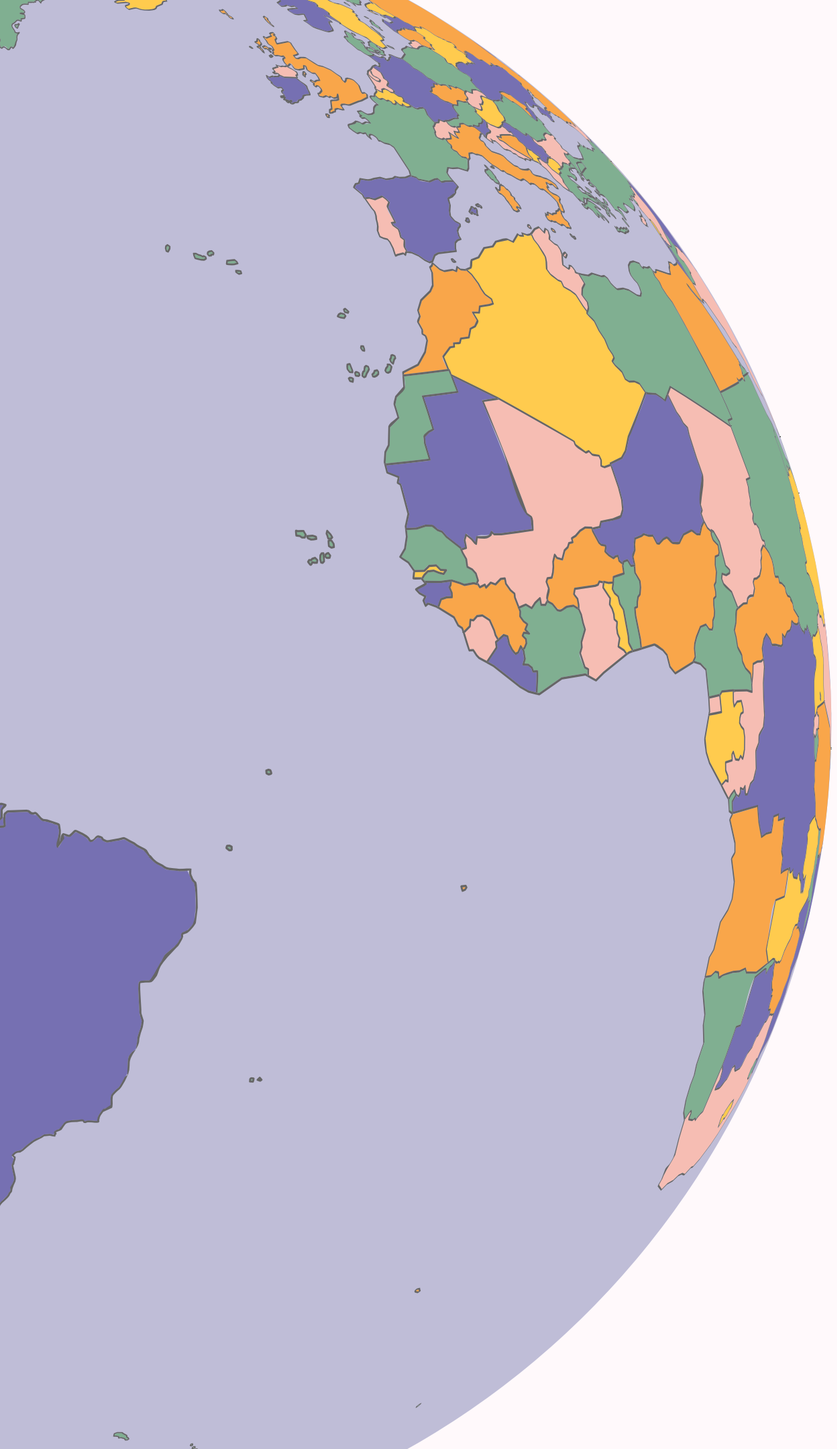
Step 3:Prediction

Pickle module is used

pickle.load is used to deserialize the model

and then we have performed prediction on the news article provided by user

This answer we have returned to the app



Step 4: Bot Connect

flask: to connect it with whatsapp

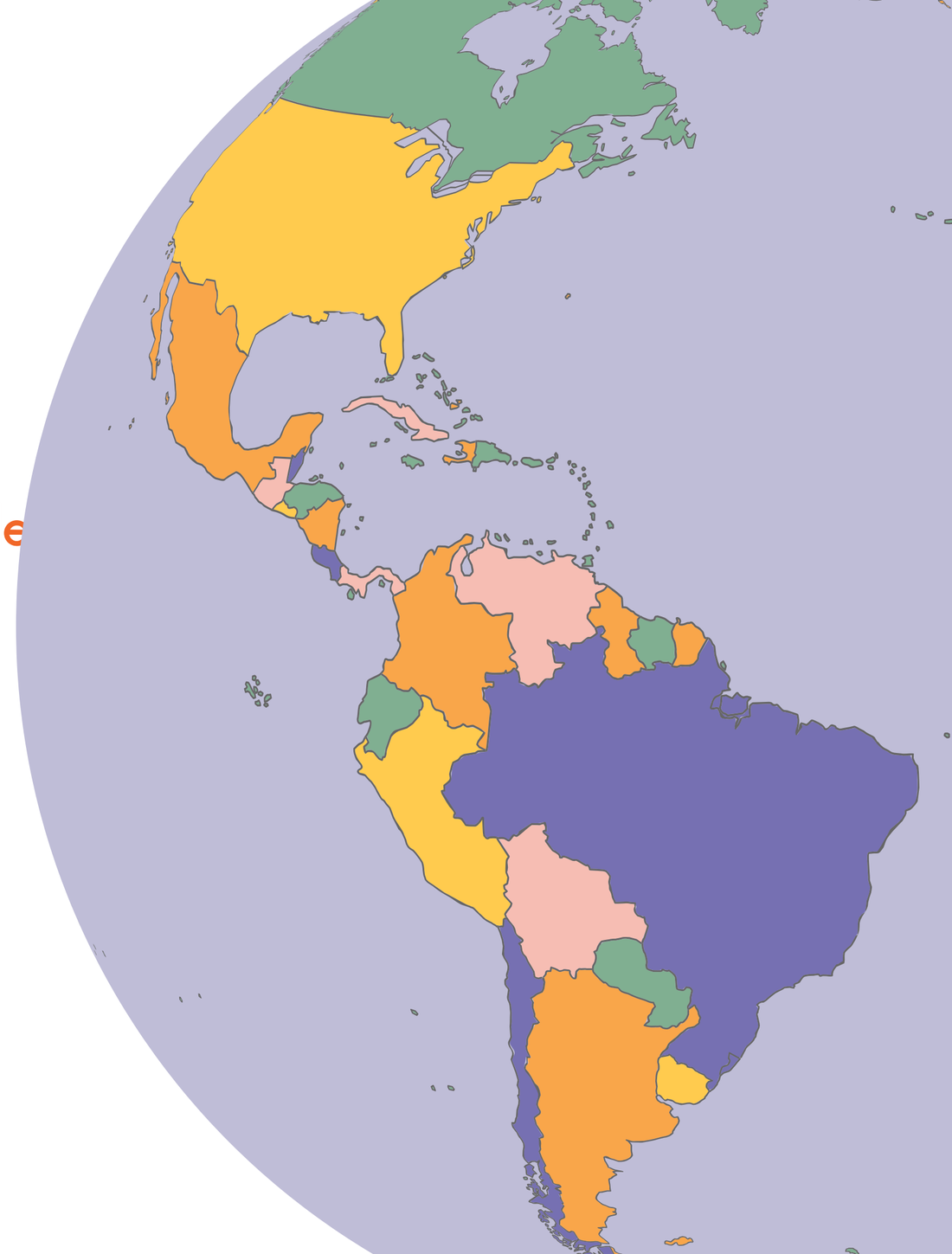
twilio: whatsapp sandbox is used to create whatsapp bot

ngrock:

Uniqueness of project

We haven't kept the project Console based, instead we have used Twilio WhatsApp API through which we are accepting the news article headlines and feeding this data to the model in order to detect the class label for the news.

Inorder to get connected to Twilio WhatsApp API we have used Ngrok which sends requests to it.



Thank You!