**Final Project Blueprint**

(Bots in the Net: Applying Machine Learning to Identify Social Media Trolls)

Team-B

Internship batch(1 Jan’21- 30 Jan’21),

Data Science Application Developer, Technocolabs.

Submitted by-

Purvit Vashishtha

Team B Members:

Sagar Gyanchandani

M.V.S Sai Hiranmayee

Ajay D. Chandrawanshi

Nitesh Chandra

Aditya Raval

Aastha Doshi

Deepanshu Rai

Abhishek DasRoy

Ashwin Subramanian

Ali Ahmad Mansoor

Harshit Shukla

Problem Statement

As suggested by the title of the Project i.e. Bots in the Net: Applying Machine Learning to Identify Social Media Trolls (Russian Troll Tweets), we have to implement machine learning or deep learning technique in order to classify whether tweets were posted by trolls using features of the dataset.

Dataset

We are provided with nine datasets of Russian Troll Tweets. There are about three million (3 Million) Russian Troll Tweets in our dataset along with attributes:

| Columns/ Attributes | Details |
| --- | --- |
| **external\_author\_id** | An author account ID from Twitter |
| author | The handle sending the tweet |
| content | The text of the tweet |
| region | A region classification |
| language | The language of the tweet |
| publish\_date | The date and time the tweet was sent |
| harvested\_date | The date and time the tweet was collected. |
| following | The number of accounts the handle was following at the time of the tweet |
| followers | The number of followers the handle had at the time of the tweet |
| updates | The number of “update actions” on the account that authored the tweet, including tweets, retweets and likes |
| post\_type | Indicates if the tweet was a retweet or a quote-tweet |
| account\_type | Specific account theme, as coded by Linvill and Warren |
| retweet | A binary indicator of whether or not the tweet is a retweet |
| account\_category | General account theme, as coded by Linvill and Warren |
| new\_june\_2018 | A binary indicator of whether the handle was newly listed in June 2018 |

**STEPS**

1. Combining the Dataset: There are 9 datasets of Russian Troll Tweets. So, after combining all the datasets we will make a single csv file which contains around 3 Million tweets along with other features.
2. EDA (Exploratory Data Analysis): We will check datatypes and shape of our attributes. We will also use describe() method to derive statistical information of the dataset and df.isnull().sum() to count null values of each column. Also in the EDA part we will use tableau and other methods of matplotlib and seaborn to perform and gain insights as well.
3. Data Cleaning and Preprocessing: We will drop columns which do not contribute in predicting trolls. We will fill the missing values by mean, median or mode whichever is appropriate.
4. Feature Engineering: First of we will separate dependent and independent features of the dataset and then assign it to variables. We will use LabelEncoder() from scikit-learn library to encode the class which we have to predict.We will try to make new columns (features) from the existing columns which can be usefull for fitting the model. Then we will split the dataset into training and testing set using train\_test\_split().
5. Model Selection: We will use a few of machine learning and deep learning techniques for classification problem. Those techniques are: Random Forest, LightGBM and LSTM. Model which would perform best with more accuracy will be used for classification.
6. Prediction: We will first of all create a predicting variable using predict method. After that we will check accuracy of the model which is used. Model with higher accuracy will be used.
7. Deployment: After prediction part we will proceed towards deployment part of the project. We will use heroku for app development for this project by specifying required files like python file containing interface of app, requirements.txt file containing required libraries and few more files.

**CONCLUSION**

We would be able to identify social media trolls, Russian Troll Tweets, using certain machine learning or deep learning algorithm. We can stop the spread of disinformation to millions of people who are using social media.

**Future Work**

Improve dataset of positive labels by collecting known troll tweets from a wider range of sources.