TO IDENTIFY WHICH COUNTRY HAS MOST PEOPLE SUPPORT

Dr.S. Krishnaveni,
Assistant Professor Department of B. Com (Business Analytics),
PSGR Krishnammal College for Women Coimbatore, India
krishnavenis@psgrkcw.ac.in

K.S. Jaya Ragavi,
UG Scholar, Department of B. Com (Business Analytics),
PSGR Krishnammal College for Women, Coimbatore, India.
selvajayaragavi@gmail.com

ABSTRACT

The intense and dangerous turmoil provoked by the breakdown in Russia-Ukrainian relations has escalated into a crisis that now afflicts both European and global affairs. In the widely used message of tweet using Twitter which includes the geographical place through specific Twitter API using consumer key, access token, bearer token and access token secret key. Knowing the region of a tweet is beneficial for lots facts analytics questions. This research is looking on the finding out of which country has most people support. An accuracy of 82% changed into done the use of dataset trained on features together with the customers' time zone, the user's language, and the parsed consumer location. The classifier plays nicely on lively Twitter countries which includes the Russia and Ukraine.

Key Words: Twitter, Keys, API, Tweets, Russia and Ukraine.

I.INTRODUCTION

Twitter has come to be a number one supply of records for information-hungry audiences round the world seeking to make feel of the Russian invasion of Ukraine. At the equal time, it's being utilized by the governments of Russia and Ukraine to set the time table for wider media reporting. Official Russian government bills were determined to be amplifying seasoned-Russia disinformation on Twitter. Meanwhile, the Ukrainian authorities has taken to the platform to appeal to its two million fans for assist. Mass conversation commenced as political communication meant to establish and control empires. Twitter has added another element to the mixture, and brought immediacy to strategic political conversation. In asymmetric conflicts, a hit twitter account can be a beneficial weapon against an adversary with many weapons and tanks. Twitter hashtag alone won't be able to instigating considerable exchange, however it could certainly play a position.

II. OBJECTIVE

Analyzing which country has more people support through the twitter dataset which have been extracted using twitter API app through using credential keys (consumer key, access token, bearer token and access token secret)

III. RELATED WORK

Which country has most people support Collaborative Python Related work. In Twitter some terms are used that are used a lot. For example, a hashtag is a keyword that Starts with a # and identifies a topic the tweet is related to. A tweet can also contain a link or A location. A retweet is a copy of a tweet That is posted to your own wall. A retweet can be used to share messages. Twitter and other social networks consist of several building blocks [1]. In Twitter the blocks Sharing, Conversations and Reputation are most

important. Every tweet has some impact for the user. Kawaka et al. [2] have measured the following Characteristics of Twitter. There are few users which have a lot of followers and on the contrary, there are many users which have only a few followers. When a user has more than 100 followers, then they have more than 1000 tweets. However, when a user reaches more than 10000 followers The number of tweets is not evenly distributed anymore. Twitter users will mostly read tweets of users that have the same time zone, same country or city. There are few people who are really Influencing the twitter world. Some users have a reach of more than 10000 people. But even small Users can reach up to 1000 users using retweets.

The Tweets are freely available to any developer [3]. All this Information is available through the Twitter API. We can search the Twitter API for a certain Hashtag, for example #ukraine. There are several approaches to determine the location of a tweet user. One can look at the tweets Text, examine meta-information or look at the relationships of users. Cheng et al. [4] used tweet text to determine the location.

Hecht et al. [5] has the same approach. They are using a Bayed Classifier for words, that is optimized using the extra condition that require that a word is at least Used once by a certain number of users in that class [6]. When this is not the case the probability of A word is zero, in all other cases the Bayes function is used. Their accuracy is around 74% using a uniform dataset.

It is observed that some of the popular and frequently used hashtags refer to place names and people names2.[7] A place name can be related to a location, town, city or a country. Country names are most often used in this category. These country hashtags are added to show that the particular tweet's content is related to the country.

The exploratory nature of this study demanded the employment of content and structural analysis methods to gain deeper understanding of the data. The three main methods used in the study are text classification, social network analysis and content analysis [8]. Text classification was conducted in two phases. The first phase included manual classification of about 500 tweets by three coders. Seven composite categories were identified after the manual classification process [9]. The first coder identified 23 subcategories which were used for coding by the other two Coders.[10]

IV.METHODOLODY

FLOW CHART

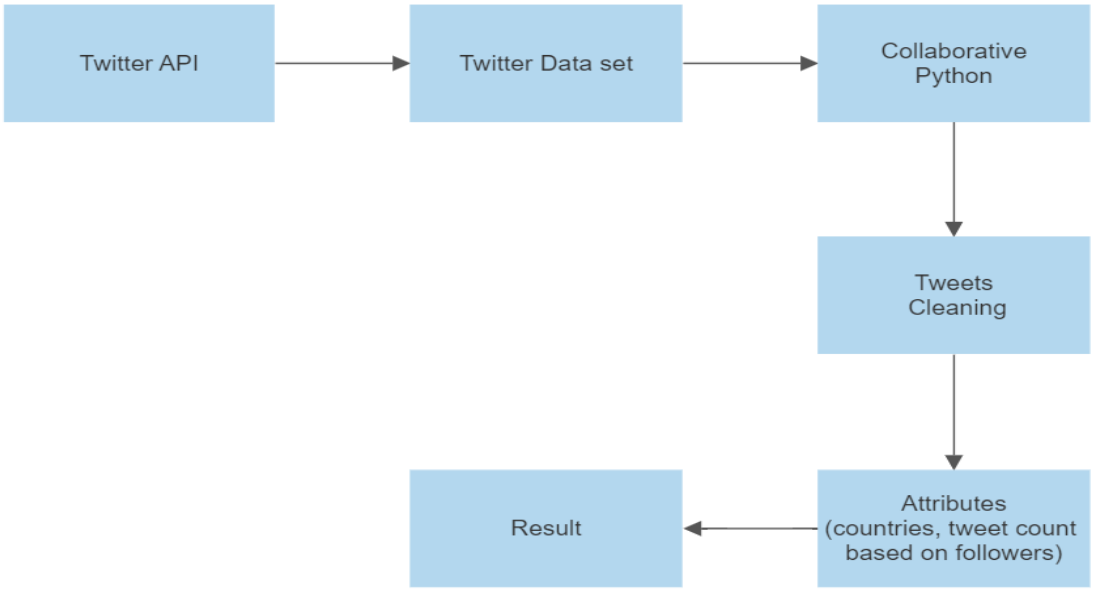

FIG 4.1 Flowchart

A. TWITTER API

The Twitter API allows users to study and write Twitter statistics. Thus, user may use it to compose tweets, study profiles, and access information and a high quantity of tweets on specific topics in precise locations. API stands for Application Programming Interface. When a user wants to get entry to our APIs, they are required to check in an application. By default, programs can most effectively get entry to public records on Twitter.

To be able to get right of entry to Twitter data programmatically user want to test and create an app on twitter developer's internet web site for authentication and thereafter we can get admission to facts with the resource of the usage of Twitter API. To sign in the twitter app, we need to create a brand-new app https://apps.Twitter.Com/. On registering the app, were able to get preserve of consumer key and consumer secret key. Next, From the configuration net web page of the app, we are able to get access token and access token secret, as a way to be used to get entry to twitter on behalf of our utility. We need to keep the ones authentication tokens non-public as they can bemisused. Best exercise is to create a separate config document and preserve the ones tokens.

B. ACCESSING DATASET

Twitter offers REST APIs to connect to their provider. We will use one python library to get entry to the twitter REST API's known as Tweepy. It offers wrapper techniques to without difficulty get right of entry to twitter REST API. To set up Tweepy we are able to use underneath command. Pip set up tweepy. In order to authorize our app to get right of entry to Twitter on our behalf, we need to use the OAuth interface.

C. PREPARING DATASET

Be analysing Twitter accounts or records it's important to analyse the structure of tweets and retweets besides the pre-system the records to cast off non-beneficial terms known as prevent-words. Pre-

processing is inside the simple term approach to take within the records and put together the information for finest output thinking about our requirement.

D. CLEANING TWEETS

Tweets are short messages, confined to a hundred and forty characters in duration. Due to the individual of this microblogging carrier (quick and short messages), people use acronyms, make spelling mistakes, use emoticons and other characters that unique meanings. Following is a quick terminology associated with tweets. Remove punctuations., Tokenization – Converting a sentence into list of terms. Remove forestall phrases. Lemmatization/stemming – Transforming any form of a word to its root phrase.

E. PROCESS

From the dataset attributes (countries, tweet count based on followers) were chosen and coded using collaborative python. It is used to identify which country has most people support based on total followers count for each country and visualized with bar chart.

V. RESULT

```
country_count = tweets.groupby(['Country'])[
    'followers'].sum().to_frame('country_support').reset_index()

fig = px.bar(country_count, x='Country', y='country_support', color='Country',
             title="Country with most people support based on hashtags", labels={ 'country_support': 'Count of followers for each Country'})
fig.update_layout(xaxis={'categoryorder': 'total descending'})
fig.show()
```

FIG 5.1 Code for finding which country has most people support

The above FIG 5.1 shows the code for reach of tweet count based on peoples support country wise.

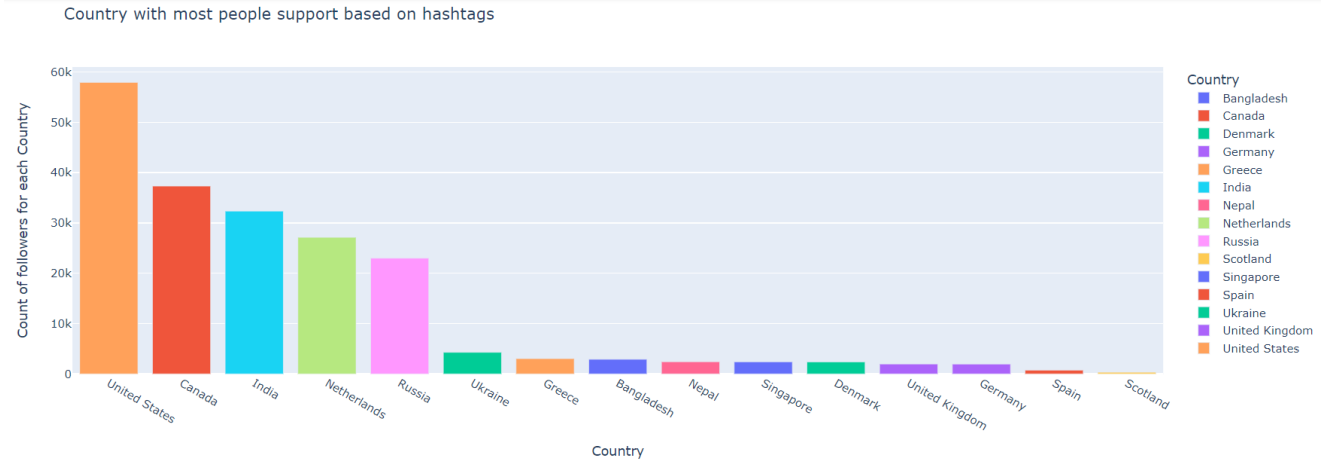

FIG 5.2 country with most people support

In the above FIG 5.2 it shows which country has most people support based on follower's count. X axis represent the countries name and Y axis represent the count of followers for each country. The Bar chart clearly shows, that which country has supported most in the order of ascending to descending. United States supported most and Scotland has supported least. Using Meta-data of a tweet the region can be recognized.

VI. Conclusion

This analysis facilitates to find which united states peoples supported Russia and Ukraine war with an accuracy of 82% on the most important dataset Several functions were tested and the aggregate of Time quarter, man or woman place and geoparser vicinity resulted in the maximum accuracy. The feature patron Language and tweet language did decorate for a few international locations the accuracy. The location of a message may be decided Using a Twitter API. This technique may be implemented to clear out on Countries and to clean up a Twitter movement. In 25% of the tweets capabilities time area and location are empty. In the one's instances, a function the use of the tweet textual content is probably beneficial. The tweeted text can comprise locations, which may be extracted the use of Entity Recognition. The user of the retweeted tweet would possibly relate to the user vicinity of who the tweet retweeted. Also, a hashtag can contain beneficial facts. Using data associated with the hashtag can improve the classifier.

REFERENCE

[1] J. H. Kietzmann, K. Hermkens, I. P. McCarthy, and B. S. Silvestre (2011). "Social media? GetSerious!" Understanding the functional building blocks of social media. Business horizons,54(3):241–251.

[2] H. Kwak, C. Lee, H. Park, and S. Moon2010. "What is twitter, a social network or a news media?". In Proceedings of the 19th international conference on World wide web, pages 591–600. ACM.

[3] Twitter developers, tweet object. https://dev.twitter.com/overview/api/tweets, Nov.2014.

[4] Z. Cheng, J. Caverlee, and K. Lee (2010). "You are where you tweet: a content-based approach to geolocating twitter users". In Proceedings of the 19th ACM international conference on Information and knowledge management, pages 759–768.

[5] B. Hecht, L. Hong, B. Suh, and E. H. Chi (2011). "Tweets from Justin bibber's heart: the dynamics of the location field in user profiles". In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 237–246.

[6] Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). "Twitter Power: Tweets as Electronic Word of Mouth". Journal of the American Society for Information Science and Technology, 60(11), 2169–2188. Doi:10.1002/asi Java, A., Song, X., Finin, T., & Tseng, B. (2007).

[7] Dilip Ratha and Eung Ju Kim Date: March 4, 2022 "Russia-Ukraine Conflict: Implications for Remittance flows to Ukraine and Central Asia" https://reliefweb.int/sites/reliefweb.int/files/resources/EN_40.pdf

[8] Abdur Rasool1, *, Ran Tao[1], Kamyab Marjan 1, Tayyab Naveed 2. 1 School of Computer Science & Technology, Donghua University, Shanghai, China. 2 College of Textile Engineering, Donghua University, Shanghai 201620, China.

[9] Liu, B. (2012). "Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies", 5(1), 1–167. doi:10.2200/S00416ED1V01Y201204HLT016, Aravind Sesagiri Ramkumar, Nanyang Technological University

[10] Thangasamy IA, Leveridge M, Davies BJ, et al. (2014) International urology journal club via Twitter: 12-month experience. European Urology 66: 112–17.