Analysis Of The Sentiment Surrounding The Citizenship Amendment Bill In India

Sai Abhishek Rupakula
Luddy School of Informatics, Computing and Engineering
Indiana University
Bloomington
srupakul@iu.edu

Daivik Uggehalli Dayanad
Luddy School of Informatics, Computing and Engineering
Indiana University
Bloomington
daugge@iu.edu

Sahithi Ancha
Luddy School of Informatics, Computing and Engineering
Indiana University
Bloomington
sancha@iu.edu

ABSTRACT

Microblogging involves using short messages to engage and interact with people. Twitter is one such microblogging site that is favored by people across all demographics and nationalities. Factors such as being able to post frequently, and real time sharing have ensured that it is much easier to get your opinion across in a very less amount of time [1]. This paper focuses on performing a sentiment analysis on the data extracted from twitter to label them (about the Citizenship Amendment Bill in India) and then using machine learning models such as Naive Bayes, Support Vector Machine and Random Forest to classify these tweets. The tweets have been obtained over a span of 2 months from 15 December 2019 to 15 February 2020.

Keywords
Twitter · Sentiment Analysis · Textblob · Wordcloud · Machine Learning · Naive Bayes · Support Vector Machine · Random Forest

1. INTRODUCTION

The advent of the internet has drastically changed the face of communication. From being able to connect and talk to people instantly to obtaining most of our information online, rather than mainstream media, the transformation has been rapid. This ease in communicating with others (sometimes anonymously) also means that any inhibitions that people may have held about expressing themselves online have been lost as well. This shift is also evident in how political and social issues play out on social media sites. An increasing number of governments and political figures have turned to social media to convey information or announcements and put forth their own opinions. This also means that people can voice their own opinions in response to what is being said [2]. In situations where the mainstream media is reserved about expressing unbiased views or the truth itself, social media can serve as an indispensable tool in getting one's point across. This narrative that plays out online can in many situations offer valuable insights about the general sentiment surrounding a topic. Whether it be business, politics or a social cause, social media sites are very good places to stay tuned-in with what the consumer or voter expects from the government or business and how the general public perceive a certain issue.

Twitter is one such social networking site where people can follow people or topics they care about. One can use hashtags to participate in discussions revolving around issues they care about. With around 330 million monthly users [3], one can safely say that twitter is among the most popular social media sites today around the world with India happening to be in the top ten countries in terms of the number of users on Twitter at 13.15 million users [4]. India's relatively young demographic also means that a good chunk of the population is tech savvy and are active users of the internet and according to one statistic, 52.3% of social media users are millennials, 28.4% are Gen Z while 15.1% of them are aged 35-44 [5]. Also an increasing number of the people are turning to social media for information instead of the mainstream media [6]. Taking into account all these factors, we decided that carrying out a sentiment analysis on twitter about the Citizenship Amendment Bill would be a good source to find the general public opinion surrounding the issue.

2. BACKGROUND

The Citizenship Amendment Bill (CAB) had originally been drafted by the Indian Government in its previous term in 2016. The Lok Sabha passed the Citizenship Amendment Bill in January 2019 but with the end of its tenure in May, the legislation lapsed. The new CAB was passed by both houses of Parliament in December. An illegal migrant is defined as people who either entered the country without proper documents, or stayed on beyond the permitted time. The Citizenship Amendment Act (CAA) amended the Citizenship Act of 1955 to make it easier for illegal immigrants belonging to Hindu, Sikh, Jain, Buddhist, Parsi and Christian minority communities from Bangladesh, Pakistan and Afghanistan to acquire Indian citizenship if they left their parent countries to escape religious persecution. The bill is about protecting religious minorities who fled to India to avoid persecution by allowing them to become citizens. However, illegal immigration from India's neighbouring countries has always been a problem in the north-eastern states of the country. While the government has made it clear that this move will not affect the existing populace as long as they can prove that they were citizens of the country, a large chunk of the population criticised this move and also led to large scale protests [7]. There were two sides to this protest - people in the north-eastern states of India believed that this move would encourage more immigrants to cross-over from the neighbouring country of Bangladesh and jeopardize their livelihoods while another part of the population believed that this bill will render many muslims already in India, stateless. On the other hand, another segment of the citizens believe that this move was long overdue to prevent the movement of illegal immigrants across the border. Our purpose is to analyse the general sentiment and find which issues are most commonly talked about.

3. LITERATURE REVIEW

Many researchers have dealt with the topic of performing a sentiment analysis on twitter data over the years. Most of the work done was based on improving the current classification techniques and proposing new and more effective models. Vishal and Sonaware [8], perform a sentiment analysis on twitter data and then use lexicon based techniques and machine learning algorithms such as Naive Bayes, Max Entropy, and Support Vector Machine to perform opinion mining on the tweets obtained. In addition to this, evaluation metrics such as accuracy, precision and recall are also used to assess the performance of each of these models. Another example is by Agarwal and Xie [9], where three models are used to classify the twitter data/tweets as positive, negative or neutral. The three types of models used to accomplish this are: a unigram model (uses over 10,000 features), a feature based model (using 100 features) and a tree kernel based model. On comparing the three, the researchers find that the latter two models outperform the unigram model significantly. The third example of the same is Wang and Wei [10] which proposes a graph-model with three algorithms such as Loopy Belief Propagation, Relaxation Labelling, and Iterative Classification to perform sentiment analysis at the hashtag level. The results from this model are then compared with results for the same from a two-stage Support Vector Machine and the effectiveness of the proposed model is assessed. However, the focus of this paper is on analysing the sentiment surrounding a certain issue and not on improving the models themselves. For this purpose, we stick to commonly used and popular methods to carry out our analysis. We perform a Sentiment Analysis using Textblob to label the data because manually labelling a dataset of 7200 is time consuming and not feasible. Three models, Naive Bayes, SVM and Random Forest are trained to classify this dataset and then compared to see which model performs the best.

4. DATA AND METHODS

The entire process from start to finish can be summarised into Figure 1.

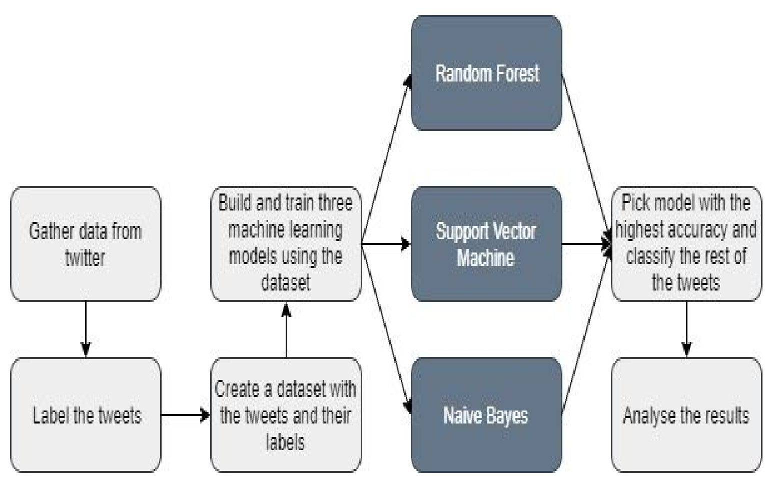

Figure 1: A brief summary of the process we employ from start to finish.

- 7200 tweets are extracted from twitter
- 3600 tweets are labelled
- A dataset is created
- Three machine learning models are built and trained
- Accuracy and f-1 scores for the three models are compared
- Model with highest accuracy and f1 score is picked
- The rest of the 3600 tweets are classified using the picked model
- The final results are analysed

4.1 Data Extraction

In this project the first step was to identify all the tweets through which the argument or discussion was going on. As we all know in twitter, groups can be identified only by the hashtag they use. Also this project focuses on a sensitive issue so we had to do a lot of research and identify the essential hashtags. According to our research the important hashtags, the time frame and the language the tweets were collected in are given in Table 1.

HASHTAGS	**TIME FRAME**	**LANGUAGE**
#CAA	15th december 2019 to 15th february 2020	ENGLISH
#NRC	15th december 2019 to 15th february 2020	ENGLISH
#CAA_NRC	15th december 2019 to 15th february 2020	ENGLISH
#CAA_NRC_PROTESTS	15th december 2019 to 15th february 2020	ENGLISH
#CITIZENSHIPAMENDMENTBILL	15th december 2019 to 15th february 2020	ENGLISH
#INDIAAGAINSTCAB	15th december 2019 to 15th february 2020	ENGLISH

Table 1: Hashtags used to extract the data

As mentioned the bill was introduced on 15th of december so we started collecting the tweets from 15th december 2019 till 15 february because after 15th of february the entire nation was under covid-19 [11] scrutiny so there was not much chaos involved with regards to the CAB bill. Also for our understanding we pulled the tweets where the language is only english because there were many tweets which had other indian languages such as hindi, telugu etc. We need to have more data for proceeding with this research because we had to clearly state the point of view of people involved.

We first used pythons inbuilt library Tweepy to collect all the data required for processing [12]. To use Tweepy we first need to create a developer account and request for API credentials. Twitter has changed its policies and now asks everyone to fill out a questionnaire to request for the access. The administrator reviews the answers and then grants the credentials. We used all the 6 above hashtags and pulled the data. The amount of data we pulled was very less (1000 > data) . The reason being Tweepy allows you to access the data which was pulled in the last seven days. Also this research is being conducted in the month of April so there were not many posts available for us to analyse. Sentiment analysis with very little data would result in some incorrect outcomes.

After a lot of research we decided to use a package called GetOldTweets (got) developed by jefferson henrique [13]. There was an updated version of this package GetOldTweets3 (got3) which we used for this project to pull all the old tweets. This module basically bypasses the limitations of twitter's API. There are tools which give you older tweets but they are all paid versions ,got3 is free to use and a very powerful tool. Basically when we scroll down the twitter we start getting more and more tweets by making calls to a JSON provider. The got3 model automates the calls to JSON provider and bypasses the restrictions. With this model we can get older tweets, get tweets by a specific username, getting username and a query by making use of the dates, getting the top tweets by the user. There are also few functionalities which we can use while making use of the model mainly specifying the number of tweets we need, mentioning the start and end dates, language of the tweets we need and location from which the tweet was tweeted. Using this package we pulled 1200 tweets for each and every hashtag so our data set was close to 7200 tweets. We then divided these 7200 tweets into 2 parts. The initial 3600 tweets were labelled to train the machine learning models.The second set of 3600 were used to analyze the sentiment from our trained machine learning model.

4.2 Data Preprocessing

For us to perform sentiment analysis we first need to check the tweets if they are in a proper format so that later on there wouldn't be any problem while performing machine learning and analysis. We felt that the raw tweets obtained were redundant and highly susceptible. There are few things we did for preprocessing mainly

- Remove all the URLS which were present in the data
- Removing the word “RT” which was before most of the tweets

- Remove all the mentions (@username)
- Removed all the numerical characters and punctuations in the sentences
- Removing all the stop words

In addition to the above step we also removed all the emoticons from the tweets. Having emoticons is not bad but the reason for removing the emoticons is discussed in the limitations section. All the irrelevant tweets were also removed in this process. Before deploying machine learning models for text classification, the data must be cleaned or pre-processed in order to achieve better results. The number of words before pre-processing the 3600 was 83,507 and after preprocessing was 56,260. Significant characters were eliminated before doing further analysis.

We also split the initial 3600 tweets dataset into a 70% training set and 30% testing set. The text is converted to a matrix of token counts using the CountVectorizer. It is then used to transform a count matrix to a normalized tf-idf representation. Finally, we train Naive Bayes and Linear SVM , Random Forest Classification using the scikit-learn library [14].

4.3 Labelling the Data

For the machine learning model to classify our tweets into different classes it would require some labelled data but we manually downloaded all the tweets and processed them which don't have any labels. Mainly we wanted to divide our tweets into 3 different classes positive,negative and neutral. The positive class would have all tweets whose opinions were positive about the Bill, negative would have all the negative reviews about the bill and neutral tweets would have opinions in which feelings were not expressed.

We had 2 options for labelling our data since we don't have any labels. The first option was to manually label all the tweets. Our dataset size was almost about 3600 tweets and manually labelling giving all the classes to them would have been a very tedious process given the time frame. So we came up with an idea of using TextBlob [15] for giving the sentiment score for all the tweets and then classifying them. TextBlob is nothing but a python library for processing textual data. It divides all the words into different natural language processing tasks such as text analysis, sentiment analysis etc. It is one of the simplest NLP packages available and a very powerful tool to use.

The sentiment functionality in Textblob takes in all the tweets, processes them and gives out 2 values one is polarity and the other one is subjectivity. For our analysis we would only consider the polarity score. Usually the score ranges from -1 to +1 value. We have taken all the values and if the value is greater than 0 then we have labelled the tweet as “Positive”, if the value is less than 0 then we have labelled it as “Negative” and if it is 0 then we have labelled it “Neutral” In this was by using the polarity score we have labelled the entire dataset. We maximized our effort in trying to make sure that the dataset is not biased.

4.4 Machine Learning models

4.4.1 Naive Bayes Classifier

Once we have the features, we can train a Multinomial Naive Bayes Classifier to predict our label of Positive, Negative or Neutral sentiment of the tweet. This classifier is based on Bayes theorem: P(A|B) = [P(B|A)*P(A)]/[P(B)] where, A and B are events. P(A|B) gives us the conditional probability of an event A occurring given event B. This is the posterior probability. P(B|A) is the conditional probability of an event B occurring given event A. This is likelihood. P(A) is probability of event A occurring independently and is called the prior probability. P(B) is the probability of event B occurring independently and is called the evidence [16]. In simple words, the multinomial theorem involves multiple occurrences of words and checking if a given word appears in the positive, negative or neutral words category. Pipeline class of scikit-learn acts like a compound classifier [17]. We have achieved an accuracy of nearly 68%.

Label	**Precision**	**Recall**	**f1-score**	**Support**
Neutral	1.00	0.15	0.27	222
Positive	0.64	0.95	0.76	513
Negative	0.75	0.60	0.67	345

Table 2: Classification report for the Naive Bayes model.

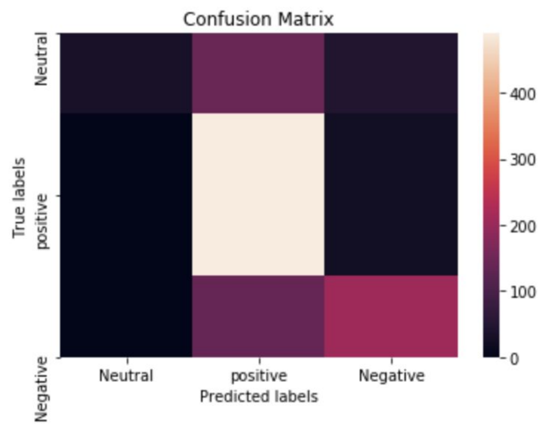

Figure 2: Confusion matrix for the Naive Bayes model.

4.4.2 Linear Support Vector Machine

Support Vector Machine is a supervised machine learning classifier, which uses a hyperplane to separate the different labels (positive, negative and neutral). The best hyperplane that distinguishes these labels is defined by the hyperplane with the maximum margin. This algorithm has a technique called the Kernel Trick. The SVM kernel is a function that takes low dimensional input space and transforms it to a higher dimensional space. It converts non-separable problems to separable problems. It is mostly useful in non-linear separation problems, which involves more than 2 classes [18]. The hinge loss function is soft margin/lazy, only updates the model parameters if an example violates the margin constraint, which makes training very efficient. L2 regularization adds an L2 penalty equal to the square of the magnitude of the coefficients and all coefficients are shrunk by the same margin [19]. By doing so, we achieved nearly 73% accuracy.

Label	Precision	Recall	f1-score	Suport
Neutral	0.96	0.39	0.56	222
Positive	0.69	0.94	0.79	513
Negative	0.78	0.65	0.71	345

Table 3: Classification report for the Linear Support Vector Machine model.

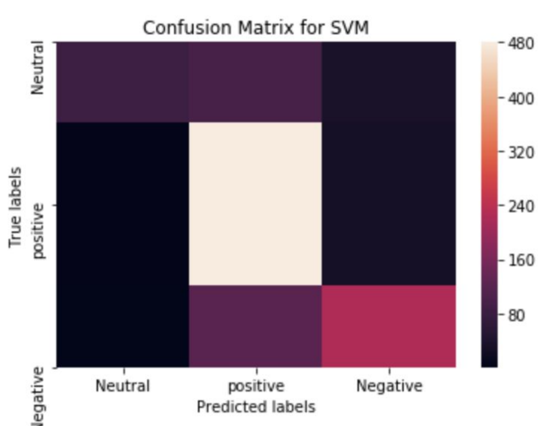

Figure 3: Confusion matrix for the Linear Support Vector Machine model.

4.4.3 Random Forest Classification

This algorithm has been widely used to train models to perform sentiment analysis tasks for its higher accuracy on large datasets. It is an ensemble tree based machine learning algorithm, which aggregates votes from different decision trees from a randomly selected subset of training set, to classify the final label during testing. As Random Forest uses the concept of Bootstrapping, where each tree works on the subset of the total training data, thereby each tree is trained on the different value of training data. It is very robust in terms of noise. Random Forest seems to be more accurate because of Bagging, so that output of all decision classifiers will be taken as an average. [20] Our random forest classifier predicts 3 labels positive, negative and neutral better than Naive Bayes and Linear SVM. By utilizing this model we attained an accuracy of 81.4%.

Label	Precision	Recall	f1-score	Suport
Neutral	0.98	0.53	0.69	222
Positive	0.75	0.97	0.85	513
Negative	0.88	0.76	0.82	345

Table 4: Classification report for the Random Forest model.

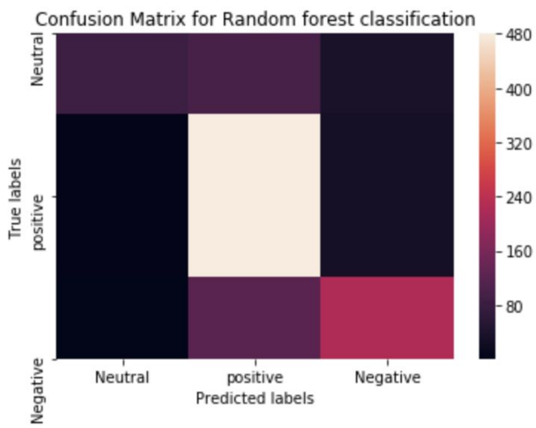

Figure 4: Confusion matrix for Random Forest classification.

MODEL	ACCURACY
Naive Bayes	68%
SVM	73%
Random Forest	81.4%

5. RESULTS

We have seen above from the classification report about the key metrics in the classification problem. We have recall, precision, f1 score and support for each and every class. The Precision gives us an idea on how many of them are correctly classified in that class. Recall basically tells us how many values of that particular class are found over the total number of values in that particular class. F-1 score is nothing but a harmonic mean between precision and recall and support is the number of occurrences of a given class. We have seen that F-1 scores for random forest classification model are much higher compared to the other models. Also the accuracy of our random forest model is greater than the other two models. So based on the conclusions from the accuracies and scores we decided to move forward with the Random Forest Classification Model for our Sentiment analysis task.

Next, using this Random Forest model, we classify the remaining 3600 tweets into the following three categories - 'positive', 'negative' and 'neutral'. After classification, we generate word clouds for the two sentiments (positive and negative). However, we found that the word clouds for both the sentiments were quite similar and hence merged the two to generate one single word cloud as shown in Figure 5. A word cloud [21] is a python library which is used to depict the most frequently used words in the tweets. From Figure 5 we see that 'caa', 'nrc', 'India', 'protest' etc are the most frequently used words.

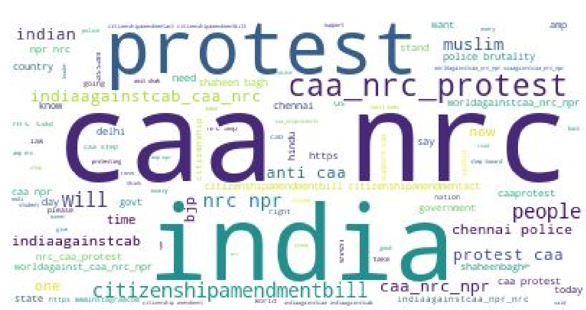

Figure 5: Word Cloud of 100 words

Out of the 100 most frequently used words, we pick the top 25 words and then plot a bar graph to depict the number of times each word is used. These results are depicted in Figure 6.

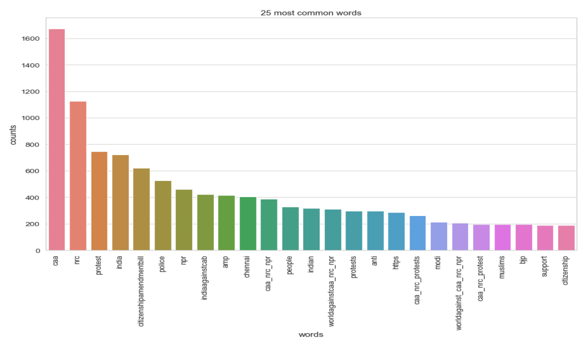

Figure 6: 25 most common words

It is clear that 'caa' is used most frequently. Finally the Sentiments obtained after analysing the tweets using a Random Forest classifier, are put together in a pie chart and the results are depicted in Figure 7. As we can observe, the percentage of positive and negative tweets is 33.2% and 15.4% respectively and the percentage of neutral tweets is 51.3%.

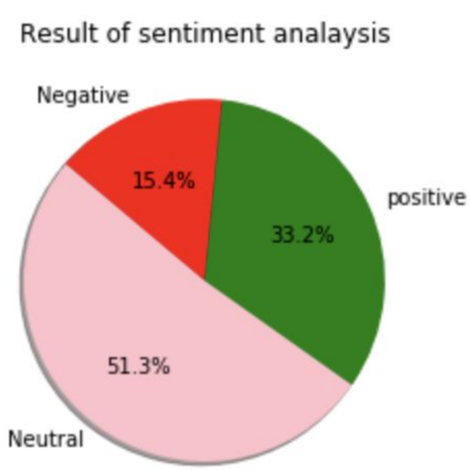

Figure 7: Results of sentiment analysis.

6. LIMITATIONS

Sentiment Analysis is quite a challenging task. While preparing the model we faced quite a few limitations. Firstly our model fails to recognize the subjective portion of the text. It fails to catch the subjective part in the text. Also few of the words in the text are domain dependent but our model does not follow any domain related analysis so that is the reason why our model performs poorly when there are similar words in positive and negative tweets. Also since the tweets contain a lot of sarcasm, negations, jokes,exaggerations and emoticons that are incoherent, the model cannot comprehend whether they are positive tweets or negative tweets.

In addition to this, we did not manually label the tweets but used a tool called text blob to do the labelling part so we also need to consider the limitations of the textblob while giving the polarity score in our analysis as well. There are a few cases where bots have been deployed to spread fake news but in our analysis any bot detection tools or packages were not used. Although proper care has been taken to remove tweets which are not related to the context, there are a few cases where tweets which do not carry any meaning were there in the data. As the issue of CAA took place in India there were a lot of other language tweets excluded for the analysis. Performing analysis considering other languages would have made our analysis more accurate.

7. DISCUSSION

When the data from all the hashtags is combined, some of the most popularly used words were ‘caa’, ‘nrc’, ‘india’, ‘protest’, ‘caa_nrc_protes’, ‘indiaagainstcab_caa_nrc’, ‘muslim’, ‘police brutality’ and ‘police’ among others. When we take a look at Figure 6, it is evident that the tweets were more negative rather than positive. The tweets classified as positive talked about the NRC being implemented as soon as possible and why it’s beneficial for the country and specifically about it not affecting Indian muslims. On the contrary, as we already mentioned in the limitations, some tweets that were sarcastic and critical about this issue were classified as positive. Some also talked about how there were several myths surrounding the issue and the need to gain a deeper understanding before protesting it. There was also some talk about pro_nrc and caa rallies and how all of the negativity and lies were being spread by leftist organizations trying to disrupt peace and stability, along with talking about the need to support the religious minorities who escaped to India as refugees and give them a better life.

The negative tweets were largely about alleged police brutality, the Indian government’s fascist agenda and divisive politics and also about the need to stand together and fight for justice. Another important point that came up in many of the tweets (both positive and negative) was the exclusion of lakhs of Indian citizens themselves from the NRC list, who were detained and later found to be Indian citizens in the state of Assam. An overwhelming number of tweets were primarily about protests that happened all over the country bringing many educational institutions and even daily lives to a stand-still in many parts of India. Some words such as ‘chennai police’ and ‘shaheen bagh’ indicate how news of specific incidents of alleged police brutality have taken the country by storm with many expressing angst and outrage over the same. A large chunk attribute these changes to politicians trying to consolidate the voters list so they can manipulate the 2024 general nationwide elections.

To further improve the results of our classification models, we can implement artificial neural networks. These models are also much more effective when it comes to classifying objective as well as subjective tweets and their rate of accuracy also tends to be higher. Another thing we can do is to improve the functionality of tools such as Textblob along with using software to detect and remove tweets by bots. Oftentimes, bots are used to spread political propaganda or false information. Instead of using Textblob to classify the data, we could manually do the same to get a more accurate classification of the dataset. However, this might be time consuming and not always feasible. One more thing to keep in mind is that our analysis is limited to data from twitter. To build a more comprehensive understanding, it’s necessary to understand the ground reality itself and incorporate this data into our study. In this case, to handle larger amounts of data that our local machines might not be able to handle, we can incorporate the usage of cloud technology.

8. CONCLUSION

Implementing NRC means that every resident in the country will have to prove that they are citizens by providing documentary evidence. This is being done under the premise that all the illegal immigrants can be singled out and sent back to their own countries since being an illegal immigrant has deemed them ineligible to apply for citizenship. However, the exceptions being made for the religious minorities has led many to believe that this is an open attack on muslims and made it harder for muslims from the neighbouring countries to immigrate to India.

The opinions surrounding this topic are diverse and strong and as with any story, there are two sides to this as well. Being rendered stateless means that the individual will have no right to vote, own property or have any social security whatsoever while also being an easy target for discrimination [22]. The state of Assam is a live example of what will happen if the current act, without any modifications is carried out. On the other hand, this same state has faced the problem of illegal immigrants coming in from Bangladesh over the years and who have in some cases even posed a security threat. However, a majority of the population detained in detention camps in those states are muslims and it's hard to believe that there's no discrimination happening on the lines or religion [23]. Knowing what has happened, and yet implementing the same is not the way to go when the repercussions are large. What is needed is a reform of policies along with consideration and compromise involving parties on both the sides.

9. REFERENCES

[1]"What Is a Microblog?" *Sprout Social*

[2]Montgomery, Bob. "Experts: Social Media Plays Bigger Role in Politics Today." GoUpstate. GoUpstate, January 24, 2020.

[3]Clement, J. "Twitter: Number of Active Users 2010-2019." Statista, August 14, 2019.

[4]Clement, J. "Countries with Most Twitter Users 2020." Statista, April 24, 2020.

[5]Talkwalker.com. "Social Media Statistics in India." Talkwalker. Accessed April 30, 2020.

[6]Pti. "Paradigm Shift: 80 per Cent Indians Consume News from Social Media Rather than a Newspaper." The Economic Times. Economic Times, July 23, 2017.

[7]"Citizenship (Amendment) Act, 2019." Wikipedia, Wikimedia Foundation, 27 Apr. 2020, en.wikipedia.org/wiki/Citizenship_(Amendment)_Act,_2019#Reception.

[8]Kharde, Vishal, and Prof Sonawane. "Sentiment analysis of twitter data: a survey of techniques." *arXiv preprint arXiv:1601.06971* (2016).

[9]Agarwal, Apoorv, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca J. Passonneau. "Sentiment analysis of twitter data." In *Proceedings of the Workshop on Language in Social Media (LSM 2011)*, pp. 30-38. 2011.

[10]Wang, Xiaolong, Furu Wei, Xiaohua Liu, Ming Zhou, and Ming Zhang. "Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach." In *Proceedings of the 20th ACM international conference on Information and knowledge management*, pp. 1031-1040. 2011.

[11]"Coronavirus Disease 2019 (COVID-19)." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, www.cdc.gov/coronavirus/2019-ncov/index.html.

[12]"Tweepy." *Tweepy*, www.tweepy.org/.

[13]Henrique, Jefferson. "GetOldTweets-Python." *GitHub*, 21 Nov. 2018.

[14] "Learn." scikit. Accessed May 2, 2020. https://scikit-learn.org/stable/.

[15] "Simplified Text Processing." *TextBlob*, textblob.readthedocs.io/en/dev/.

[16]Ray, Sunil, and Business Analytics and Intelligence. "Learn Naive Bayes Algorithm: Naive Bayes Classifier Examples." Analytics Vidhya, April 1, 2020. https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/.

[17] Susan Li, "Multi-Class Text Classification Model Comparison and Selection", Sep 25, 2018,https://towardsdatascience.com/multi-class-text-classification-model-comparison-and-selection-5eb066197568

[18]"SVM: Support Vector Machine Algorithm in Machine Learning." Analytics Vidhya, April 15, 2020. https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/.

[19]Stephanie. "Regularization: Simple Definition, L1 & L2 Penalties." Statistics How To, October 14, 2018. https://www.statisticshowto.com/regularization/.

[20] L. Breiman, Random forests, Machine Learning, vol. 45. Issue 1, pp. 5-32, 2001.

[21]"(Tutorial) Generate Word Clouds in Python." DataCamp Community. Accessed May 2, 2020. https://www.datacamp.com/community/tutorials/wordcloud-python.

[22]Shankar, Soumya. "India's Citizenship Law, in Tandem With National Registry, Could Make BJP's Discriminatory Targeting of Muslims Easier." The Intercept, January 30, 2020. https://theintercept.com/2020/01/30/india-citizenship-act-caa-nrc-assam/.

[23]Ahmed, K. Anis. "Why Is India Making Its Own People Stateless?" The New York Times. The New York Times, September 11, 2019. https://www.nytimes.com/2019/09/11/opinion/india-assam-stateless-bengalis-muslim.html.