Comparison Study of different ML/DL approach on Twitter Sentiment Analysis For General Election

**Class** : XLS EN\_CSE 4410/6410

**Project Participants** : Md. Asaduzzaman Jabin,

**Student ID(s)** : 811562873 (Graduate Student)

**Date** : Aug 29, 2022

# Introduction

In modern life, social media craze controls the lives of people, impact the business, and helps to understand the customer base. Large tech giant – MAANG- Meta, Amazon, Google built a 30 trillion dollar business every month from only keep their user engaged in their ecosystem. So, Emotions and sentiment of the user depict a vital information visually rather than articulately in business model, democracy and social rights, marketing. Also, it plays a crucial role in the different areas- e.g, social media interaction, AdSense, client-based business model, suspecting social mob, identifying national security threats, metaverse, and AR/VR game console design, customer’s psychological study like - human behavior in social media understanding, detection of mental breakdown, social media harassment and synthetic social bullying. So, detection of social media sentiment in real-time with high recognition rate is still a challenging task. Text based sentiment usually performed in NLP using four-stages consisting of- pre-processing, extraction of feature, and classify expression in a dynamic manner, and training and testing.

In this project, I am going to apply various combination of machine learning and deep learning methods- e.g., Naïve Bayes, SVM and Bi-LSTM architecture. The project’s novelty comes up with the combination hyperparameters in different layers of the ML model and mask the sentiment with respective analysis. The existing maximum accuracy achieved over this dataset is <70% only. As a result, I found a scope to work on it to get more accurate prediction of sentiment. We found a dataset using tweets’ API, pre-processed the data, extracted the right features and applied the Naive Bayes, SVM, Bi-LSTM Classifier to obtain public opinions. As a result, we identified outliers, analyzed controversial and swing states, and cross-validated election results against sentiments expressed over social media. The results reveal that the election outcomes coincide with the sentiment expressed on social media in most cases. The pre and post-election sentiment analysis results demonstrate the sentimental drift in outliers. Our sentiment classifier shows an accuracy with a precision of 93%.

# Literature Review

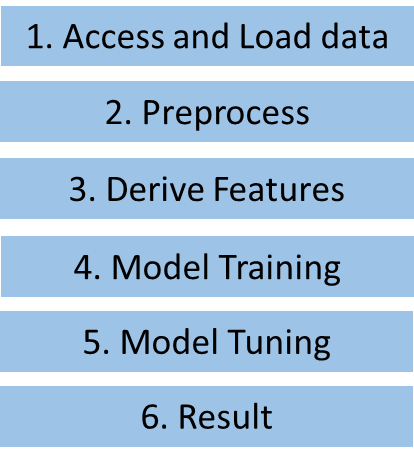
Electronic media such as- Facebook, Twitter, Instagram, LinkedIn and discord are common ways to know the opinion and feedback of mass people. Normally, People share rumors, news, political views, social events, political campaigns, business promotions, global events, promoting social works and developments, and expressing sentiment about elections. In fact, social media plays a vital role first time in US election campaign in 2008. Mr Obama utilize the power of social media- twitter significantly for political campaign. Even In 2016, Mr Trump won a victory over Mrs Clinton which made everyone shocked (CBS 16). Pre-election polls suggested the candidate’s dominance over opponent party. It pulls all the ML enthusiast and business leaders to apply ML on social media data. Even Trump dubbed it a critical tool that played a pivotal role in his victory (CBS, 2016).

Fig 1: The overall workflow of sentiment analysis using ML.

Several works on the sentiment analysis in all around the world have already been carried out in 2012 U.S. election, employing the Naive Bayes Classifier using Unigram features [1], in 2016 U.S. election- lexicon feature extraction [2], Sentistrenght [3] and other elections, such as Indian PM elections [4,5], Iranian elections [6], Singaporean elections [7], and Colombian elections [8]. These works provide insights into social media sentiments as well as their correspondence with the actual election results; further details are provided in Section 3. Similarly, social media analysis of the 2020 U.S. election can also potentially unveil several hidden sentiments about both candidates. The study can become yet more critical since the shadows of rigging are cast on the elections. The sentiment analysis becomes more interesting since votes were cast via postal services. But, it is hard to recognize sentiment sarcasm trivially. In some cases, the negative sentiments are classified as positive due to their writing styles. Twitter is a much better place for sentiment analysis as compared to other social medias [9]. In this sense, some people on social media might not be serious in really expressing their actual feelings. Therefore, their sentiments do not reflect the true picture.

Sentiment analysis involves classifying opinions in text into three main categories, i.e., “positive” or “negative” or “neutral” [10]. Sentiment information can be extracted using various ways, including speaker recognition [11], physical activity recognition [12], philological signals [13], human facial features [14], and textual information expressed over social media. Sentiment analysis is employed in numerous fields for opinion mining, such as focusing on multi-level single and multi-word aspects to manifest several domains in Twitter datasets [15], in recommendation systems [16], being employed for business intelligence [29], for finding public opinion about a particular rule before presentation [17], in comments analysis [18], News and print media sentiment Analysis [19], commercial movie reviews analysis [20], publishing or advertising [21-22].

Text

Description automatically generated

Fig 2: The human level sentiment analysis policy.

Moreover, social media might not represent complete sentiment in elections, since all voters are not present. While social media might not represent everyone completely, it provides a sample space of people’s opinions. Also, some people might not want to reveal their views due to privacy issues, so even if they are on social media, they might not express their true opinion [18]. Nevertheless, despite all these limitations, social media sentiment analysis provides the nearest approximation of public sentiment. To detect sarcasm, we used Countvectorizer() along with tokenizer(), further details are provided in Section 3.2.

Diagram

Description automatically generated

Fig 3: The overall model architecture for sentiment analysis.

Chart, treemap chart

Description automatically generatedChart, waterfall chart, treemap chart

Description automatically generatedIn this project, pre- and post-election sentiments for both candidates in each state for Indian PM election are considered [23]. Outliers are well known in fundamental data mining tasks to find extreme values laying outside the trends followed by other data samples. For closely contested states, finding the outliers is significant for data analysis. Moreover, we analyzed public sentiment rather than election results state wise.

Chart, waterfall chart

Description automatically generated

Fig 4: The Confusion matrices for Naïve Bayes, SVM, Bi-LSTM Respectively for sentiment analysis.

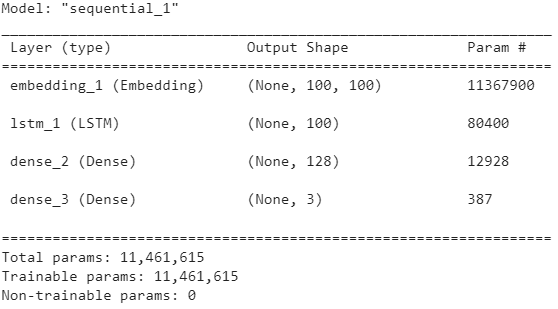
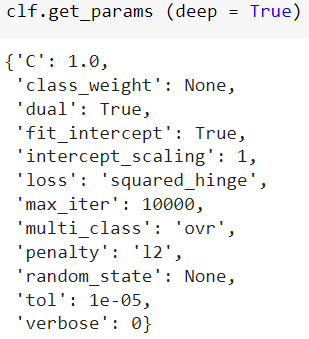
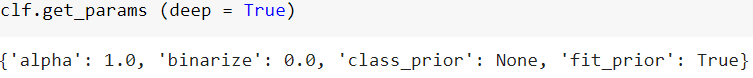
# Methodology

In this method, we use Pytorch in python as programming language. Initially, we evaluate the feasibility of our machine learning algorithms – Naïve Bayes, SVM and Bi-LSTM using classification learner app toolbox in MatLab find the quick evaluation of data. Here, Fig. 1 depicts the overall workflow of the machine learning process to detect the sentiment of the people. The data set consists of few characteristics-

* Kaggle Twitter Data Set- Prime Minister Election [23]
* Size- Number of tweets: 162K
* Sentiment Class/ Data Labeling: Negative (-1), Neutral (0), and Positive (+1)
* Class Distribution: Negative (23%), Neutral (33%), and Positive (44%)
* Attributes: 2 , Tweet- clean\_text (string), category (numeric)

To understand how sentiment is extracted in human level please refer to the Fig. 2. Basically, we need pick a few influential words from the whole sentence and assign a tag for each part of speech, The clean\_text attributes contains the public tweets (string) and category (numerical -1<x<1) and category defines the mapped sentiment (Negative (-1), Positive (+1), Neutral (0)). Then, each adjective words in the text will finally determine the sentiment from the text. However, Fig. 3 displays the ML and DL model architecture for sentiment analysis. Please be noted that, difference between ML and DL is that, DL does not any feature extraction process, it already embedded in the first layer (Embedded) of LSTM. But, for ML, we have the flexibility to choose our feature extractor (Countvectorizer – Acc- 75%). On the other hand, TFIDF and Hash verctorizer shows lower accuracy for machine learning models. One important thing we need to mention that, we need to do data cleaning and dropping the missing/ duplicate samples before extracting the features. Then , we need to apply lowercase for the symmetry and do eliminate the stop words (unwanted words – articles, preposition, conjunctions ,etc) from the data by using regex command in python. To know more, please refer to as Supplementary material slides (Page 5-8). After feature scaling, we split the data in 40%: 40%: 20% ration respectively to train, validation, and test set for both ML and DL approach. Let’s summarize the overall approach-

* Data splitting: 40%, 40%, 20% respectively train, validation, and test.
* Improved Accuracy: Bi LSTM 93% > SVM (89%) > Naïve Bayes (75%)
* For, Bi-LSTM, Dense (128) provides an accuracy of 93%, 256 gives 75%, 64 depicts 72%
* For SVM, we use Countvectorizer with accuracy of 89% (alpha = 1.0, L2 regularize, Tolerance 1e-5, ), TfidfVectorizer(73%), HashingVectorizer (63%)
* In Naïve Bayes, alpha =1.0, 10 fold puts an accuracy 75% , for alpha =2.0, 5 fold gives 69%



(a)

(b)

(c)

Fig 5: (a) Tuned parameters for Naïve Bayes, (b) SVM, and (c) Bi-LSTM models.

# Result Analysis

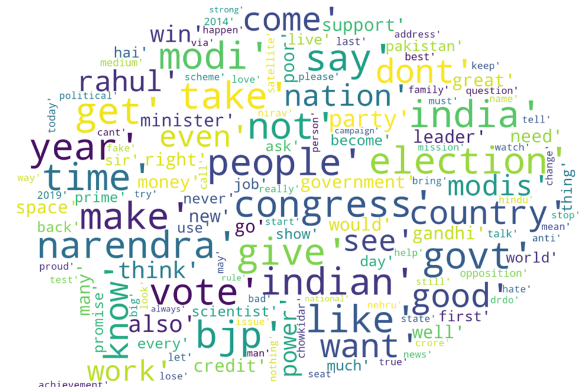
After testing the model, we display the confusion matrix for each algorithm. Here we, found that positive and neutral tweets depict more confusion to the model training are depicted in Fig. 4.

We considered 10-fold cross validation both for Naïve Bayes and SVM to avoid over fitting problem and used validation set to tune the model hyperparameters. The tuned hyperparameters for the model are resulted in Fig. 5.

Overall, we compile the code on Core-i9, 16 GB, RTX 3050Ti architecture. All the processing time are resulted in the same machine architecture for whole data sets of 162K tweets. All the performance comparison is done on testing set (32K tweets 20% of the whole dataset) is depicted in Fig. 6. The most used buzzwords are demonstrated as word cloud is showed in Fig. 7 (left)

A picture containing text

Description automatically generated

Fig 6: Accuracy comparison of Naïve Bayes, SVM, and Bi-LSTM algorithms for testing set.

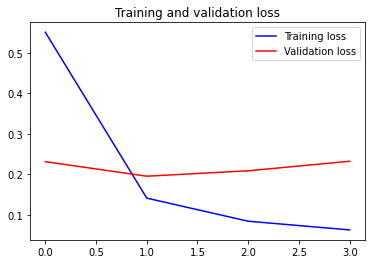
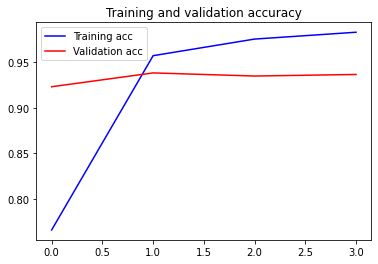


Fig 7: (Left) The word cloud for most used buzzwords in tweet sentiments, (Right) The loss and accuracy trends for our highest accurate model (Bi-LSTM 93% testing accuracy)

But from Fig. 7 (right), shows the loss and accuracy trends of the whole training and validation process for Bi-LSTM (most accurate 93% in testing). Here we can notice that the loss (blue) during training goes almost to zero and validation remains flat in the whole process (red). However, accuracy goes up to 98% during training and but validation loss remains almost flat. Actually, validation set helps use recorrect the hyperparameter again after training process which ensures the generalization of the data in future scenario.

# FINAL REPORT - Conclusions

In this project, we have done the comparison study of different ML and DL algorithms for sentiment analysis of an election campaign data. We use simple procedure and model architecture to classify 3 sentiments. From testing results, we can say SVM is overall best approach (89%, 600 seconds) when we consider both learning time and accuracy as crucial parameters, but in terms of accuracy, LSTM will be the best model (93% accuracy). Our model will generalize for all future data, since we consider holdout validation and 10-fold cross validation took in account during training process. Moreover, from the confusion matrix we can say that positive and neutral sentiment are more complex to classify than negative sentiment.

# Appendix

* Our Code: Google drive:
* <https://colab.research.google.com/drive/1Ghc0IgW7HIeUPfO5UTYX7yXxLXzk7rS9?usp=sharing>
* Reference data set:

<https://www.kaggle.com/code/subhadeepdebnath/twitter-sentiment-analysis-lstm-complete-guide/data>

Code Reference 1: <https://github.com/soham2707/Twitter-Sentiment-Analysis->

Code Reference 2: <https://www.kaggle.com/code/moathmohamed/twitter-sentiment-lstm-98>

* Presentation: video uploaded in google drive

**Reference:**

[1] Wang, H.; Can, D.; Kazemzadeh, A.; Bar, F.; Narayanan, S. A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In Proceedings of the ACL 2012 System Demonstrations, Jeju, Kprea, 8–14 July 2012; pp. 115–120.

[2] Nausheen, F.; Begum, S.H. Sentiment analysis to predict election results using Python. In Proceedings of the 2018 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 19–20 January 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1259–1262.

[3] Caetano, J.A.; Lima, H.S.; Santos, M.F.; Marques-Neto, H.T. Using sentiment analysis to define twitter political users’ classes and their homophily during the 2016 American presidential election. J. Internet Serv. Appl. 2018, 9, 18.

[4] Jose, R.; Chooralil, V.S. Prediction of election result by enhanced sentiment analysis on Twitter data using Word Sense Disambiguation. In Proceedings of the 2015 International Conference on Control Communication Computing India (ICCC), Trivandrum, India, 19–21 November 2015; pp. 638–641.

[5] Singhal, K.; Agrawal, B.; Mittal, N. Modeling Indian general elections: Sentiment analysis of political Twitter data. In Information Systems Design and Intelligent Applications; Springer: Berlin/Heidelberg, Germany, 2015; pp. 469–477.

[6] Salari, S.; Sedighpour, N.; Vaezinia, V.; Momtazi, S. Estimation of 2017 Iran’s Presidential Election Using Sentiment Analysis on Social Media. In Proceedings of the 2018 4th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), Tehran, Iran, 25–27 December 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 77–82.

[7] Choy, M.; Cheong, M.L.; Laik, M.N.; Shung, K.P. A sentiment analysis of Singapore Presidential Election 2011 using Twitter data with census correction. arXiv 2011, arXiv:1108.5520.

[8] Cerón-Guzmán, J.A.; León-Guzmán, E. A sentiment analysis system of Spanish tweets and its application in Colombia 2014 presidential election. In Proceedings of the IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (Socialcom), Sustainable Computing and Communications (Sustaincom)(BDCloud-Socialcom Sustaincom), Atlanta, GA, USA, 8–10 October 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 250–257.

[9]. Salloum, S.A.; Al-Emran, M.; Monem, A.A.; Shaalan, K. A survey of text mining in social media: Facebook and twitter perspectives. Adv. Sci. Technol. Eng. Syst. J. 2017, 2, 127–133.

[10] Turney, P.D. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. arXiv 2002, arXiv:cs/0212032.

[11] Cumani, S.; Laface, P.; Kulsoom, F. Speaker recognition by means of acoustic and phonetically informed GMMs. In Proceedings of the Sixteenth Annual Conference of the International Speech Communication Association, Dresden, Germany, 6–10 September 2015; pp. 200–204.

[12] Usman Sarwar, M.; Rehman Javed, A.; Kulsoom, F.; Khan, S.; Tariq, U.; Kashif Bashir, A. Parciv: Recognizing physical activities having complex interclass variations using semantic data of smartphone. Softw. Pract. Exp. 2021, 51, 532–549.

[13] Narejo, S.; Pasero, E.; Kulsoom, F. EEG based eye state classification using deep belief network and stacked autoencoder. Int. J. Electr. Comput. Eng. (IJECE) 2016, 6, 3131–3141. [CrossRef]

[14] Kalsum, T.; Mehmood, Z.; Kulsoom, F.; Chaudhry, H.N.; Khan, A.R.; Rashid, M.; Saba, T. Localization and classification of human facial emotions using local intensity order pattern and shape-based texture features. J. Intell. Fuzzy Syst. 2021, 40, 9311–9331.

[15] Janjua, S.H.; Siddiqui, G.F.; Sindhu, M.A.; Rashid, U. Multi-level aspect based sentiment classification of Twitter data: Using hybrid approach in deep learning. PeerJ Comput. Sci. 2021, 7, e433.

[16] Pradhan, R.; Khandelwal, V.; Chaturvedi, A.; Sharma, D.K. Recommendation System using Lexicon Based Sentimental Analysis with collaborative filtering. In Proceedings of the 2020 International Conference on Power Electronics IoT Applications in Renewable Energy and its Control (PARC), Mathura, India, 28–29 February 2020; pp. 129–132.

[17] Guo, X.; Li, J. A novel twitter sentiment analysis model with baseline correlation for financial market prediction with improved efficiency. In Proceedings of the 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS), Granada, Spain, 22–25 October 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 472–477.

[18] Cardie, C.; Farina, C.; Bruce, T. Using natural language processing to improve erulemaking: Project highlight. In Proceedings of the 2006 International Conference on Digital Government Research, San Diego, CA, USA, 21–24 May 2006; pp. 177–178.

[19] Mukwazvure, A.; Supreethi, K. A hybrid approach to sentiment analysis of news comments. In Proceedings of the 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions), Noida, India, 2–4 September 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1–6.

[20] Štrimaitis, R.; Stefanoviˇc, P.; Ramanauskaite, S.; Slotkien ˙ e, A. Financial Context News Sentiment Analysis for the Lithuanian ˙ Language. Appl. Sci. 2021, 11, 4443. [CrossRef]

[21] Singh, V.K.; Piryani, R.; Uddin, A.; Waila, P. Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-levelsentiment classification. In Proceedings of the 2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), Kottayam, India, 22–23 March 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 712–717.

[22] Jin, X.; Li, Y.; Mah, T.; Tong, J. Sensitive Webpage Classification for Content Advertising. In Proceedings of the 1st International Workshop on Data Mining and Audience Intelligence for Advertising, ADKDD ’07, San Jose, CA, USA, 12 August 2007; Association for Computing Machinery: New York, NY, USA, 2007; pp. 28–33.

[23] https://www.kaggle.com/code/subhadeepdebnath/twitter-sentiment-analysis-lstm-complete-guide/data