Advances in Artificial Intelligence and Machine Learning; Research 3 (4) 1619-1639 Received 27-2-2023; Accepted 4-11-2023; Published 11-11-2023

Performing Stance Classification and Bot Detection on the Indian Farmers' Protest – A Study to Unveil Hidden Perspectives.

Sarabjeet Kaur Kochhar skaur@ip.du.ac.in
Indraprastha College for Women,
University of Delhi.
New Delhi, India.

Chinmay Chahar chinmaychahar@gmail.com
Department of Information Technology
Indira Gandhi Delhi Technical University for Women,
New Delhi, India

Corresponding Author: Sarabjeet Kaur Kochhar

Abstract

The presence of illegal, harmful content, rumors, misinformation, and Twitter bots has consistently brought the social media platforms such as Twitter into the spotlight. Therefore, it is advisable to exercise caution when analyzing tweets. To establish the credibility of any patterns and findings derived from tweets, it is essential to thoroughly investigate the source and authenticity of the tweets in question. This paper advances in this direction by introducing a novel approach involving bot detection and a comparative analysis of human and bot-generated tweets related to the farmers' protest. A framework for knowledge differentiation is deployed to accomplish this goal. The framework unearths the global perspectives of people about Indian farmers' protests, in the form of stances, the results of which serve as nuggets of knowledge derived at the lower level of abstraction. Unexpected results of stance detection motivated the study of bot detection in each tweet of each stance. Knowledge discovered by bot detection and characterization studies was thus built over stance detection and yielded higher-order knowledge nuggets, which identified the widespread presence of bots in tweets both for and against the protest, thus establishing the misuse of social media platforms like Twitter to influence and control the narrative of the social events that significantly impact people's lives. Characterization of issues being tweeted by humans vs. bots in favor of and against farmers' protests was accomplished by conducting a comparative analysis of N-grams in each category. Vocabulary analysis established that texts tweeted by bots mimicked the vocabulary pattern of the tweets by human users. Research inferences such as these can be invaluable for policy makers, enabling them to gain a macro-level understanding of the situations on the ground level and leverage such information for making policy decisions, in order to be prepared to handle similar situations in the future.

Keywords: Farmers protest, Knowledge differentiation, Stance detection, Bot detection, N-grams, Vocabulary Analysis

Citation: Sarabjeet Kaur Kochhar and Chinmay Chahar. Performing Stance Classification and Bot Detection on the Indian Farmers' Protest – A Study to Unveil Hidden Perspectives Advances in Artificial Intelligence and Machine Learning. 2023;3(4):92.

1. Introduction

Agriculture is a crucial sector of the Indian economy, employing around 50% of the workforce and contributing about 15% to the country's GDP [1]. However, Indian farmers are known to be facing several challenges, including limited access to credit, outdated infrastructure, and low prices for their produce. In September 2020, the government introduced three farm bills aimed at transforming the farming sector by facilitating private investments, eliminating intermediaries, and removing intervention in price-setting mechanisms. However, the Indian Farmers' Protest started in November 2020 in response to the government's decision to pass the three agricultural laws. The farmers' unions and various organizations opposed the laws, claiming that they would benefit large corporations and harm small-scale farmers. Narayanan argues that the acts rather than focusing on the welfare of the farmers, facilitated the agritech companies and retailers by removing constraints on buyers and shifted the control of trade in the hands of central government from the state governments [2]. Singh et. al. claim that the Farm Acts were aimed at eviction of smallholders from agriculture and argue that state autonomy should be protected [3].

The protests began in November 2020 and gained momentum as farmers from across the country joined in. The protests were mostly peaceful, with farmers staging sit-ins and blocking highways and railways. The government responded with attempts to break up the protests, including water cannons, tear gas, and barricades. In January 2021, the Supreme Court of India put the laws on hold and formed a committee to negotiate with the farmers. However, the farmers' unions rejected the committee, claiming that it was biased in favour of the government. The protests continued, with the farmers demanding repeal of the laws and a legal guarantee for minimum support prices for their crops. In February 2021, violence erupted during a tractor rally organized by the farmers in the national capital, resulting in clashes with the police and the death of several protesters. The government responded by arresting protest leaders and tightening security measures.

The protests have had several impacts, including drawing attention to farmers' issues, challenging neoliberalism, uniting farmers across India, receiving international attention, and impacting politics in India. They also highlighted the need for policy changes to support small-scale farmers and improve their livelihoods. The farmers protest also sparked a storm over Twitter, which led to the popularization and internationalization of the issue. With Canadian Prime Minister Justin Trudeau, international celebrities like Greta Thunberg, lawyer Meena Harris, media person Mia Khalifa, and Indian political bigwigs like Home Minister Amit Shah and celebrities like Akshay Kumar, Sachin Tendulkar, Virat Kohli, Ajay Devgan, Karan Johar, and Lata Mangeshkar taking their stances on Twitter about the Farmers Protest, the issue naturally grabbed unparalleled attention from Twitter users throughout the world. Figure 1(a) shows a word cloud depicting the locations of tweets. Larger font sizes show that Delhi, Punjab, and Mumbai residents posted the largest number of tweets regarding the farmers' protest in the period of study. In fact, the farmers' issue also had supporters from other metropolitan cities such as Chandigarh, Bengaluru, and Hyderabad. The word cloud also shows that countries like Canada, Pakistan, the United States, and England did not lag behind in expressing their sentiments on the issue on Twitter. Figure 1(b) not only brings forth the names of Indian cities such as Uttar Pradesh, Jaipur, Rajasthan, and Haryana but also shows China and Zhengzhou city.

This resultant war of Tweets has been the center of attention of many research studies. Most of the machine learning research works in the area have however been limited to sentiment analysis of

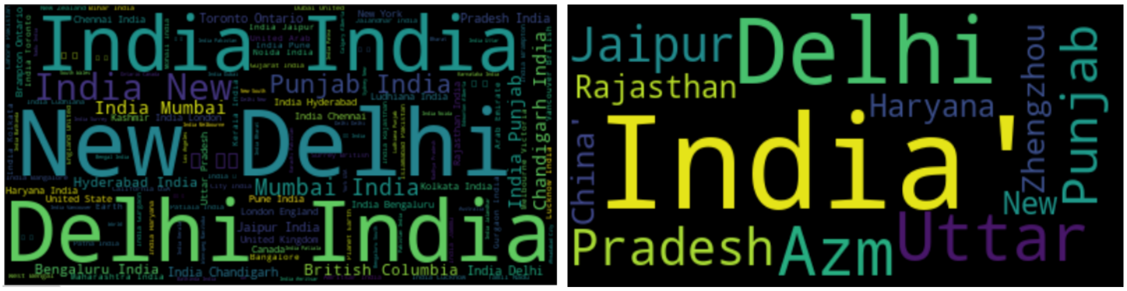

Figure 1: (a), (b): Word Cloud Depicting the Influence of Farmers' Protest in the Cities Worldwide.

tweets. Some are focused on studying the role and impact of influencers and some others restrict themselves to stance detection. It is important to note here that it is pertinent to ascertain the credibility of the source of tweets for any analysis done on them to be treated as credible. Naturally, it becomes crucial to ascertain how many of the Twitter users tweeting about the farmers' protest were real and how many of them were bots. This is an area that has received almost no consideration from the machine learning community so far, especially concerning the farmers' protests, though there is extensive work on the detection of twitter bots in general.

In this paper, we work on the principle of knowledge differentiation [4–6], wherein knowledge is uncovered at two different levels of abstraction, built one over the other. Such a framework is instrumental for the analysis and comprehension of the discovered knowledge from multiple perspectives, leading to better actionability. The framework undertakes the following tasks:

At the lower level of abstraction, we implement and compare six classifiers, namely the SVC, Multinomial Naïve Bayes, Logistic Regression, Decision Tree, Random Forest, and Passive Aggressive classifier to perform stance detection in tweets pertaining to the farmers issue as Anti-farmers, Pro-farmers, and Neutral. We worked with over two lakh tweets, collated from two sources. Around one lakh tweets were scraped directly from Twitter, pre-processed and a sample from them was subjected to human expert annotation for preparing the training set. Manual annotation allowed blending in human subjectivity and led to the discovery that the dataset scraped directly from Twitter was very unbalanced. It had a very insignificant number of tweets against farmers. Around one lakh tweets using the same search phrases, and of the same time period were therefore downloaded from Kaggle, merged into the dataset, and again subjected to annotation.

At the higher level of abstraction, we built over the tweets, multi-classified as stances, representing the knowledge derived at the lower level of abstraction. Bot detection was performed for each class of the tweets separately and their results compared and analyzed to draw significant inferences.

Further, the tweets, classified as humans or bots were merged with the Farmers' Protest User Dataset [7], to map the user and tweet characteristics and perform two kinds of analysis namely, analysis of the N-grams and a vocabulary analysis to characterize and compare tweets spread by bots vs. the texts tweeted by human users.

Definitive, useful and interesting conclusions emerged from the study presented in this paper. Results of the stance detection (classification) reported highest number of anti-farmer tweets, followed by neutral and then the pro-farmer tweets. These results were contrary to the expectation, results of other related research frameworks, sentiments reported by national and international dailies and the initial discovery of very few anti-farmer tweets during annotation. So, the results drawn from the lower level of abstraction, encouraged us to perform the study of bot detection in each class of tweets, i.e., at the higher level of abstraction. Bot detection on the classified results confirmed a larger presence of bots in the anti-farmers' category as compared to the pro-farmers' category, thus corroborating the classification results. A very small difference in the percentage of bots in the Anti-farmer and Pro-farmers' category, led to the inference that technology was rampantly used to drive the campaigns in all the categories of users. Study of N-grams presented a comparative analysis of issues being tweeted by humans vs. bots in favor of and against farmers' protest. Vocabulary analysis established that texts tweeted by bots were designed in line with the vocabulary pattern of the tweets of human beings.

2. Related Works

We present a comprehensive discussion of the machine learning works related to the farmers' protest and work done in this paper, organized under the following heads: Sentiment analysis, Role of influencers, Stance detection and Bot Detection.

2.1 Sentiment Analysis

Most of the research endeavors focusing on the farmers' protest tweets are centered primarily around sentiment analysis. Neogi et. al. study the sentiments of people regarding farmers protest by using four classifiers to predict sentiment polarity [8]. Bag of Words and TF-IDF machine learning techniques are compared and conclusion that Bag of Word approach depicts better accuracy than TF-IDF is drawn. Tiwari et. al. use word embeddings that are fed into Bidirectional Encoder Representation from Transformer (BERT) to perform sentiment analysis of farmers protest tweets [9]. The work reports an overall positive and neutral sentiment of tweets (81%) towards the farmers' protest. A list of most frequently talked about topics related to the protest is also presented by deploying LDA modeling. Singh et. al. also perform sentiment analysis of Twitter data collected using popular farmers protest hashtags and vote for Naive Bayes algorithm's efficiency in sentiment analysis over Support Vector Machine and Logistic Regression [10]. Mahajan et. al. also categorize the prevalent sentiment on twitter about farmers protest by using a BiLSTM Model [11]. Predicting democratic protests is yet another work for sentiment identification of the tweets in farmers protest is presented in [12].

2.2 Role of influencers

Some research endeavors study role of influencers such as the politicians and celebrities on the political opinion by analyzing farmer protest tweets. Celebs such as American musician Rihanna started a tweet storm, garnering over half a million reactions, bringing tremendous fanfare and global limelight to the issue, and triggering tweets from other international and national celebrities like climate activist Greta Thunberg, lawyer Meena Harris, media person Mia Khalifa, and Indian celebrities. Mishra et. al [13] collected around 2 lakh tweets to study the impact of tweets of public figures on Rihanna's tweet on the farmers' protests and the hashtags #IndiaAgainstPropaganda and #IndiaTogether. The tweets are classified to pro and anti-stances by training a Word2Vec model on an initial set of high precision keywords. Conclusions confirming hate speech, trolling of celebrity Rihanna on the basis of her gender, race, nationality and religion, celebrity collusion and coordinated tweeting are drawn. Emergence of alternative narratives regarding the farmers' protest from the misinformation spread on the twitter was also established.

The role of influencers on twitter regarding the three events related to government initiatives namely,. Article 370, CAA, Farm Bills, and the COVID-19 pandemic are studied by clustering the user embeddings, obtained using Google's Universal Sentence Encoder (USE) [14]. Retweet and polarity metrics are used to quantify the prejudiced engagement of influencers and draw out their characteristics. The study concludes that while the influencers engage in a partisan manner on the policy-based issues, such as Article 370, CAA, Farm Bills, always aligning with one of the political parties, on COVID-19 crisis, the influencers converged with the thought process of the government.

A study on the role and influence of Canadian Twitter users in garnering support for the farmers protest has been presented by Monteiro [15]. The work also studies accusations of international conspiracy against India, levelled against Disha Ravi, Nikita Jacob and Shantanu Muluk and overrule any commonality in their tweets using similarity measures, leading to the conclusion of no evidence of an international conspiracy.

2.3 Stance Detection

Detecting stances from tweets has received considerable attention of the machine learning community. Suta et. al. combine topic modelling features with classification for stance detection in tweets [16]. The topic modeling features are exploited to generate an explanation of stance labels by studying the most relevant topics within the tweets. Upadhyaya et. al. examine the association of emotions with the temporal perspectives to detect the underlying stance of the tweet [17].

Only a couple of works have, however, used stance detection to analyze farmers' protests. Mahajan et. al. attempt to trace the day and location of Farmers protests by applying Bidirectional Long Short Term Memory (BiLSTM) Model on the tweets related to the protests [11]. In one of the closest works related to the work presented in this paper, Kamble et. al. use transfer learning to train a classification model to predict public stances for farmers protest, on the same dataset as used in this paper [18]. The paper uses the concept of transfer learning to extend an existing ULMFiT (Universal Language Model Fine-tuning) model by Howard and Ruder, for stance prediction of tweets regarding farmers' protests. While we merged in more tweets for balancing the dataset, they use a technique called artificial super sampling to do the same. The work claims F1-score of 0.67 and accuracy of about 0.7, while the Passive- Aggressive classifier deployed in our work achieves F1-Score of 0.78 and accuracy of 0.77. Neha et. al. perform a comparative analysis of tweet clusters using the semantic difference in the clusters to identify narratives in three protests around government policy such as mass mobilization, on-ground activities and call-to-action for people's participation [19].

2.4 Bot detection

Though a large body of research has been dedicated to bot detection on Twitter, no work so far has focused on detecting the bots from the tweets on farmers' protest. Chavosi et. al. worked on using a lag-sensitive hashing technique called De-Bot to find corelated Twitter user accounts by clustering user accounts that are highly synchronous for long periods of time [20]. Kantepe et. al. perform feature extraction by analyzing a user's tweets, profile, and temporal behavior [21]. The suspended Twitter user accounts are then used to classify the bots. Wei et. al. use bidirectional Long Short-term Memory (BiLSTM), with word embeddings to classify human and spambot accounts on Twitter [22]. Feng et. al. deploy an information network to learn heterogeneous influence amongst Twitter users to perform heterogeneity-aware Twitter bot detection [23].

In contrast to the above works, the work presented in this paper focuses on classifying the tweets into one of the following three stances namely: pro-farmers, anti-farmers or neutral. Bot detection is performed on tweets belonging to each category of these stances and statistics compared. Tweets were further mapped with user characteristics to do a comparative analysis of the kind of tweets posted by humans vs. bots. Vocabulary analysis of the merged tweets was done to draw inferences about the difficulty level of the text used by humans vs. bots in their respective tweets.

3. Research Framework and Implementation Details

In this section, we present outline of the framework, used in this work. Details of the programming environment and the underlying implementation are also provided in this section. *The general framework for knowledge differentiation and instantiation of the same used in this paper is presented in Figure 2.*

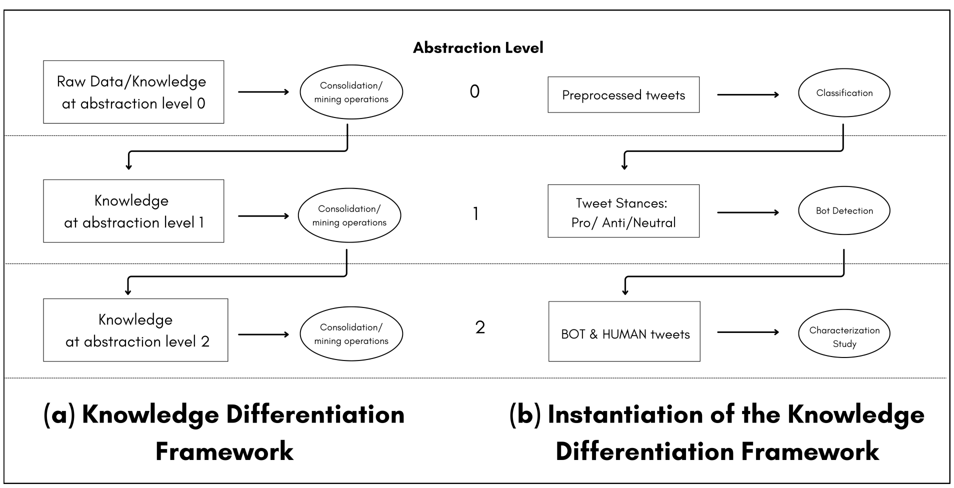

Figure 2: (a) General template of the Knowledge Differentiation Framework and (b) Instantiated Framework used in the Paper.

As per the knowledge differentiation model, data mining and/or consolidation operations are performed over knowledge nuggets at a lower of abstraction to yield higher level knowledge nuggets [4]. Each level of abstraction provides a higher, possibly novel way of understanding the data. Multiple perspectives allow better analysis and comprehension of the discovered knowledge, leading to better actionability.

In the instantiated framework deployed in this paper, pre-processed tweets can be treated as the knowledge at zeroth abstraction level. Classification performed over the tweets yields stances, representing knowledge nuggets at abstraction level one. Bot detection performed separately over tweets belonging to each stance yield identification of tweets as either BOTS or Humans. Characterization study is performed to draw out peculiar traits of bot and human tweets and compare them at the abstraction level two. It can be seen that each level of abstraction provides newer, enriched perspectives into the tweets. It is pertinent to point here that the original framework for knowledge differentiation also mentions the concept of consolidation of knowledge over different periods of time. However, in this paper we have restricted ourselves to mining operations.

The detailed schematic diagram of the framework based on knowledge differentiation used in this paper is presented in Figure 3. Google Colab, the free cloud notebook environment supporting Python 3.6.9 with the Google Compute Engine at the backend was deployed for implementation of the framework. 1.29 GB of RAM was used and the 41.88 GB disk Space was used. Twitter was chosen to source data for understanding the perspectives of people from world over about the farmers' protest since it boasts of a very large subscriber base consisting of people from all ages, beliefs and nationalities. Moreover, the microblogging site provides a platform for unconstrained communications and has already been used as a critical resource by researchers in times of health crisis such as the COVID-19 pandemic [6, 24] and other natural calamities such as earthquakes [25], Floods [26] etc. We also used Farmers' Protest User Dataset from Kaggle [7] to extract tweets regarding farmers' protest, in order to introduce diversity into the dataset and balance the data. The heterogenous data from the two aforementioned data sources was merged. The merged data was subjected to human expert annotation in order to prepare a balanced, training sample. The tweets were then preprocessed by converting them to lower case and all punctuations, emoticons, and special characters were removed using the re library. Next phase of preprocessing included splitting the tweets into tokens using function word_tokenize (), extracting common stems of the words using stemming and lemmatizing the tweets using the NLTK class WordNetLemmatizer (). The stop words that didn't add value to the text, were removed then from the tweets, with the help of the stopword library of NLTK. TfidfTransformer, CountVectorizer and TfidfVectorizer functions were imported from the sklearn.feature_extraction.text for feature extraction and for converting a collection of raw documents to a matrix of TF-IDF features. Pandas was used for importing CSV files and the data analysis. NumPy was used for working with arrays.

Stance Detection was performed over the preprocessed tweets using six different classifiers. For stance detection, function LogisticRegression was imported from sklearn.linear_model to deploy machine learning with Logistic Regression. The function train_test_split was imported from sklearn.model_selection and used for splitting arrays or matrices into random train and test subsets. The functions pyplot was imported from matplotlib for plotting graphs, visualizations and Seaborn was imported to plot various kinds of statistical graphs. The functions log_loss, precision_score, recall_score, f1_score and confusion_matrix was imported from sklearn.metrics for importing various evaluation metrics and computing the confusion matrix to evaluate the accuracy of a classification

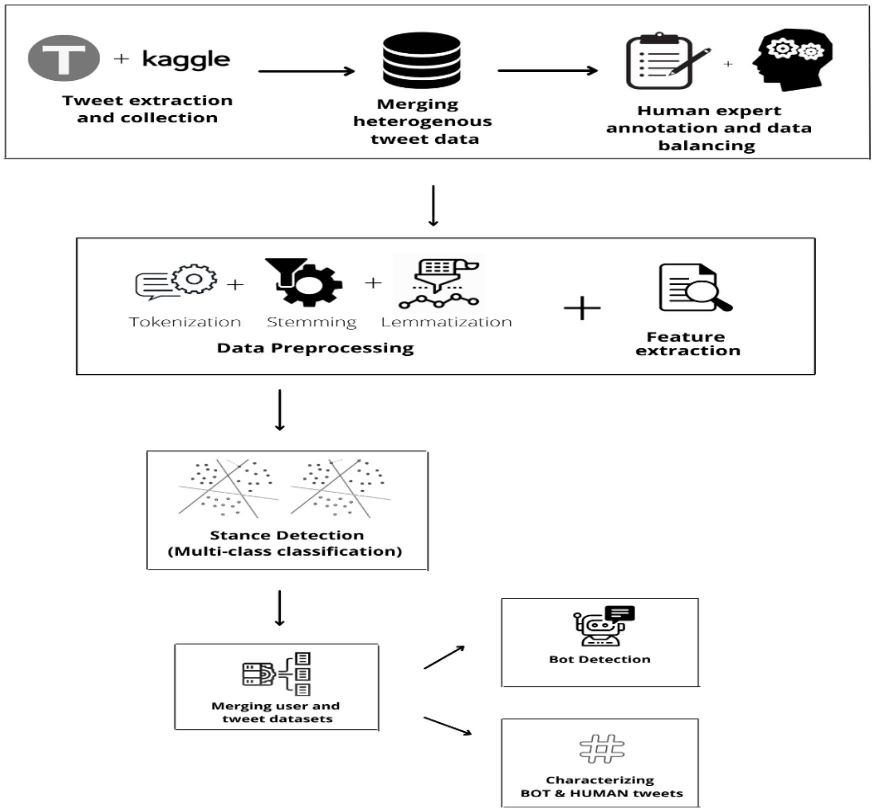

Figure 3: Detailed Schematic Diagram of the Implemented Framework.

respectively. MultinomialNB was imported from from sklearn.naive_bayes, DecisionTreeClassifier from from sklearn.tree, and SVC from sklearn.svm to implement the naive Bayes algorithm, decision tree classifier and the Support Vector Classification respectively. StandardScaler was imported from sklearn.preprocessing for performing Standardization.

Bot detection was then performed over the tweets belonging to each stance and to compare and contrast salient characteristics of tweets by bot vs. humans, N-gram analysis and vocabulary analysis was carried out. For the mentioned tasks, the following python libraries, packages and modules were used: pandas, numpy, matplotlib.pyplot, seaborn, tweet-preprocessor, nltk, wordnet, modules accuracy_score, precision_score, recall_score, f1_score, roc_auc_score, LogisticRegression, scale, normalize, confusion_matrix from sklearn.metrics, SimpleImputer from sklearn.impute, train_test_split from sklearn.model_selection, DecisionTreeClassifier from sklearn.tree, GaussianNB from sklearn.naive_bayes, N-grams from nltk.util, Stopwords from nltk.corpus, CountVectorizer from sklearn.feature_extraction.text, Counter from collections, WordCloud, STOPWORDS from wordcloud, flesch_reading_ease from textstat. A non exhaustive list of functions employed includes str.lower, str.replace, word_tokenize, PorterStemmer, and WordNetLemmatizer.

4. Stance Detection

Stance Analysis falls under the ambit of opinion mining tasks. The task of stance detection undertakes to determine the standpoint, position or judgement towards a given proposition. This goes further than sentiment detection, wherein the emphasis is determination of mere emotional polarity of a text., i.e., whether a piece of text is positive, negative, or neutral. Stance detection finds important applications in analytical studies monitoring public opinion pertaining to the political and social issues, product reviews, marketing research, and brand management etc., where gauging public opinion and its fallout is important. Peoples' stance serves as feedback that may significantly affect key governmental and organizational policies, product improvement designs, and marketing plans etc.

Classification, a predictive, supervised learning task, is one of the data mining techniques, is used for stance detection. For a very primitive understanding, using classification, the task of stance detection involves assigning predetermined stance labels to the future objects based on a model (called a classifier), learned from the objects that have already been labeled (called the stance training set).

4.1 Data Extraction and Preprocessing

Initially more than one lakh tweets were collected on the topic 'Farmers' Protest' using two methods. Twint, a library that doesn't use Twitter's API, allows one to scrape not only tweets, but a user's attributes such as their followers, following, etc., while evading most API limitations. Twint was used for retrieving tweets for the time period August - November 2021. Some of the key search phrases used to retrieve tweets were "Farmers Protest", "Kissan Morcha" and "Kissan Protest". However, the collected dataset was found to be highly unbalanced. To balance the dataset, 1,000,000 tweets on the same subject and same time period were then retrieved from the Farmers' Protest User Dataset on Kaggle [7] and merged with the tweets scraped directly from Twitter.

It is desirable to pre-process the tweets before performing any data mining task. The key motivation is to avoid the GIGO i.e., Garbage In Garbage Out mode. The tweets, in the form posted on social media are highly unstructured, in mixed case, contain stop words, punctuation symbols, numbers and special characters, such as URLs, mentions, abbreviations, slangs, hashtags and emojis etc., most of which do not contribute to the data analytic task at hand. The following pre-processing tasks were therefore performed on the scraped raw data. Redundant and null columns were removed. Feature selection was also performed next to remove the attributes that seemed to be of little value for the task at hand. Emojis, URLs, mentions, images and videos were removed from the tweet text. Since the scope of current study was limited to English language, text from local languages like Hindi and Punjabi was removed. The text of all tweets was then converted to lowercase for uniformity. The tweets were then tokenized and all unnecessary special characters and stop words were removed. To extract the root form of the remaining words in tweet text, they were subjected to lemmatization.

4.2 Implementation

The quality of training set is an important determiner of the quality of the classification process. Human Expert Annotation is a process deployed to classify some of the existing objects by domain experts, in order to train a model for classification. Human annotation has been extensively used in a variety of advanced natural language processing projects, in the fields of machine learning and artificial learning, that involve complex linguistic and visual training [27]. While the data driven automated models are important tools for extracting knowledge from gigantic datasets, the importance of blending in human subjectivity via domain experts for identification, tagging and extraction of useful linguistic patterns is undeniable [28]. In the current work, Human Expert Annotation was deployed for multi class classification, wherein initially a randomized sample of one fourth tweets out of the 3386 tweets scraped manually using Twint, were manually annotated into one of the three classes namely, anti-farmers, Pro-farmers and neutral. The manual annotation led to the discovery of many tweets that contained the search keywords but were completely irrelevant to the task at hand. Human expert annotation also led to the discovery that the dataset scraped from Twint was very unbalanced. There were very few tweets against farmers (Stance Prediction of Tweets on Farmers Protests in India, no date). As noted in Section 4.1. we therefore downloaded tweets using the same search phrases and same time period from Kaggle [7] and merged into the dataset. A portion of the merged dataset was again subjected to annotation, class distribution of which is shown in Table 1, and was used to train the classifiers.

Table 1: Count of labeled tweets in manually annotated training data

Labels	Tweets
Pro-farmers	215
Neutral	198
Anti-farmers	124

Six algorithms were chosen as candidates for classification of pre-processed unlabeled tweets, namely the SVC, Multinomial Naïve Bayes, Logistic Regression, Decision Tree, Random Forest, and Passive Aggressive classifier, with the help of the respective libraries for the chosen classifiers imported from the Sci-kit learns package, as detailed in Section 3.

The frequency of each feature, as per their weightage in the document was computed using the class TfidfVectorizer (Term Frequency - Inverse Document Frequency). The sklearn metric module was used to derive the evaluation metrics, namely, accuracy, precision, recall, and F1-score. The confusion matrix and heat map depicting values of the evaluation metrics precision, recall, F1-score for the six classifiers chosen for evaluation are shown in Figures 4 (a) – (f).

A comparative analysis of the performance of the chosen classifiers on the test set, evaluated based on the four metrics, namely accuracy, precision, recall and F1-score is tabularized above. Table 2 shows that the accuracy of three classifiers namely, Multinomial NB, Logistic regression and the Passive Aggressive Classifier was higher than the other three classifiers. The table also clearly indicates a tie in F1-scores, which is a weighted harmonic mean of precision and recall, for Logistic regression and Passive Aggressive classifiers. However, a slightly better precision and recall of the passive-aggressive classifier helps to break its tie with the Logistic Regression classifier and it was thus selected for classification of the tweets.

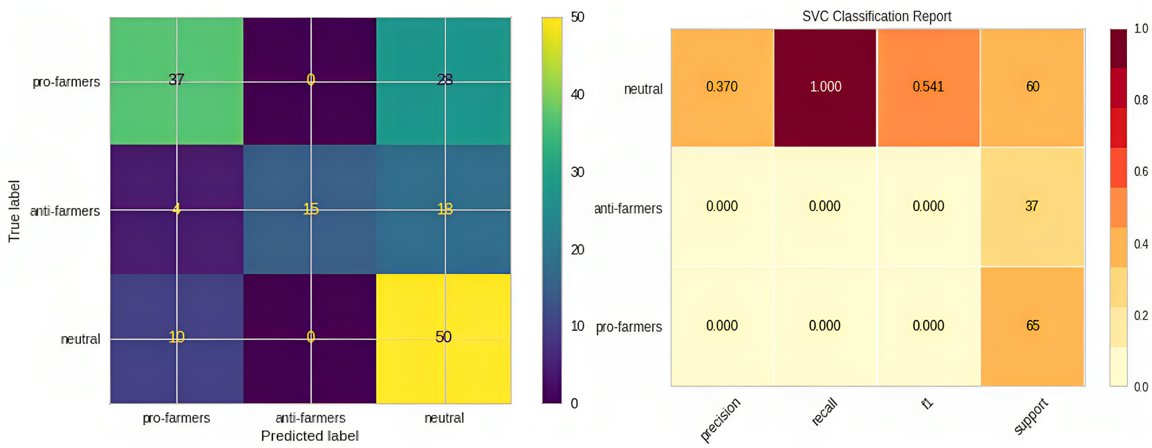

Figure 4: (a): Confusion Matrix and Heat Map for the SVC classifier.

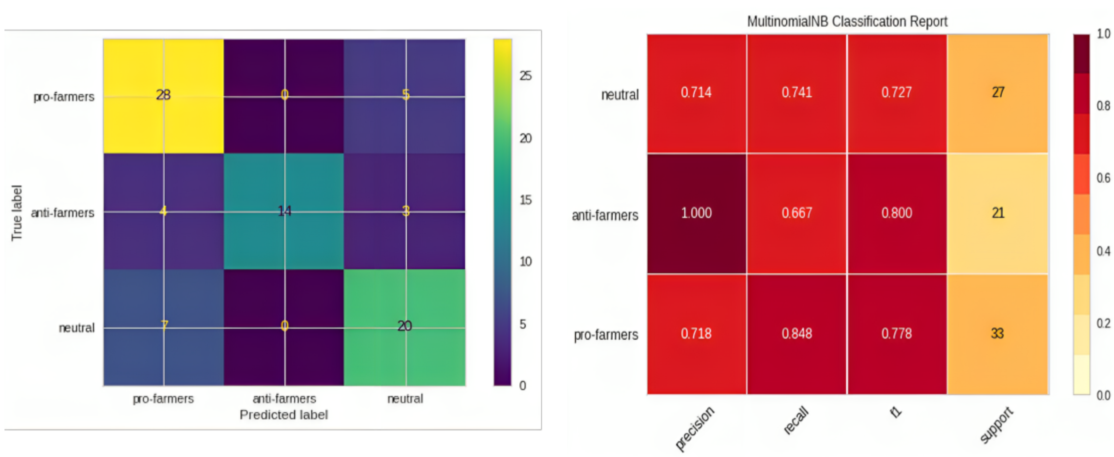

Figure 4: (b): Confusion Matrix and Heat Map for the Multinomial NB classifier.

4.3 Results and Inferences

Results of the classification using passive-aggressive classifier on the pre-processed tweets are as shown in Table 3. Contrary to the expectation, number of anti-farmer tweets (39.87%) led the count, followed by neutral (38.65%) and anti-farmer tweets (21.71%). Interestingly, as reported in Section 4.2, while annotation, the number of anti-farmer tweets were so few that the dataset had to be balanced. Also, this result seemed to defy the very strong support in favor of farmers, both nationally and internationally, as reported by major newspaper dailies in India and abroad. The result was also contrary to the results presented by Neogi et. al. [8], who reported that the majority of tweets were neutral, followed by tweets expressing positive sentiment and the negative sentiments coming in last. We therefore decided to further study each class of tweets separately to find the clues underlining such a result. Section 5 elucidates the details of the further investigations conducted on the classified tweets.

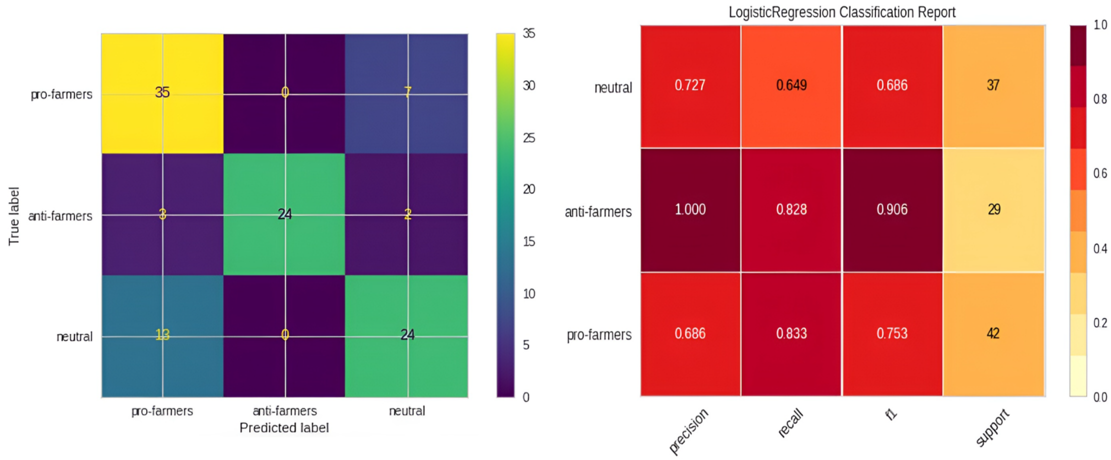

Figure 4: (c): Confusion Matrix and Heat Map for the Logistic Regression classifier.

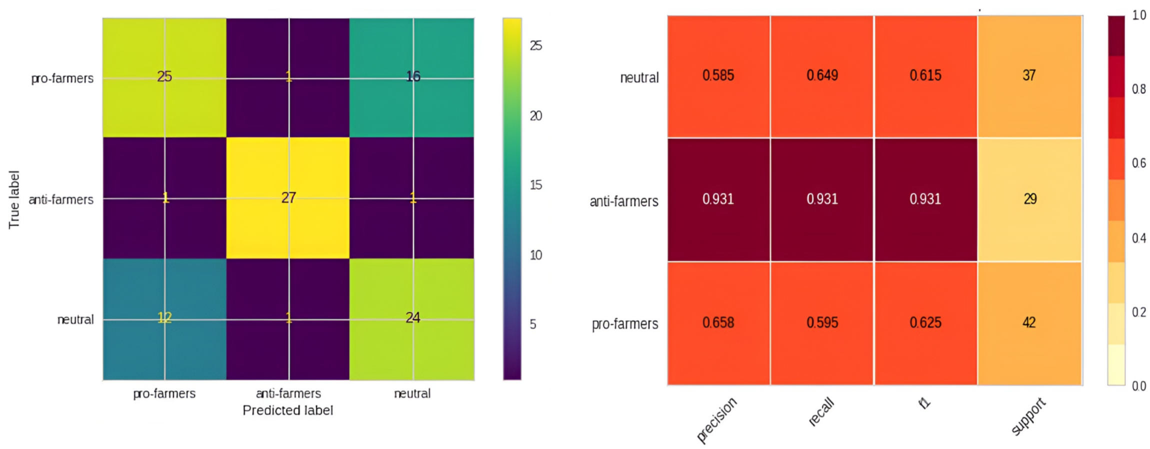

Figure 4: (d): Confusion Matrix and Heat Map for the Decision Tree classifier.

5. Bot Detection

Results obtained from the stance detection of tweets set the stage for bot detection carried out on the classified farmers' tweets, as detailed in this section. The other motivation for doing bot detection were the following observations:

Most of the anti-farmer and neutral tweets were tweeted at the same time as the pro farmer tweets.

The tweet posting timings appeared to be round of clock

A large number of tweets were posted from same places.

These observations prompted us to check for the presence of bots, details of which are presented in following subsections.

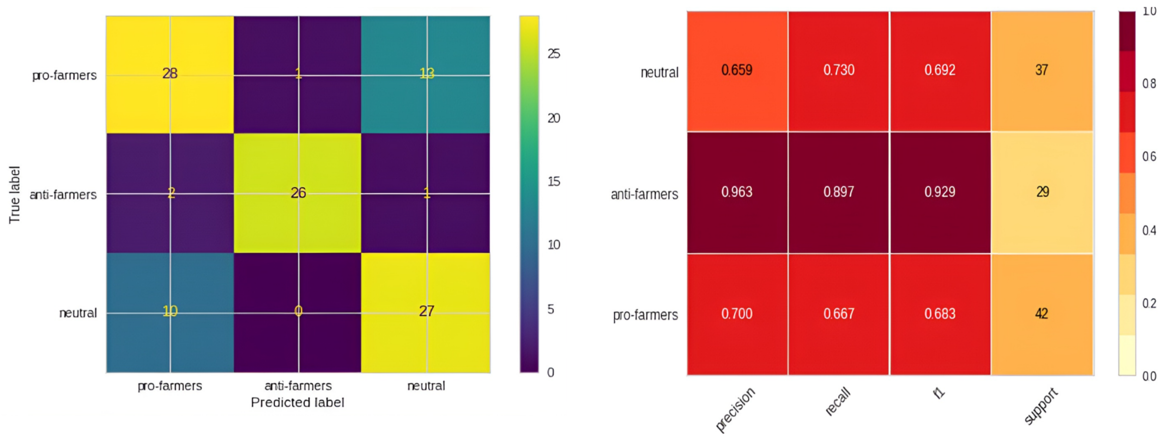

Figure 4: (e): Confusion Matrix and Heat Map for the Random Forest classifier.

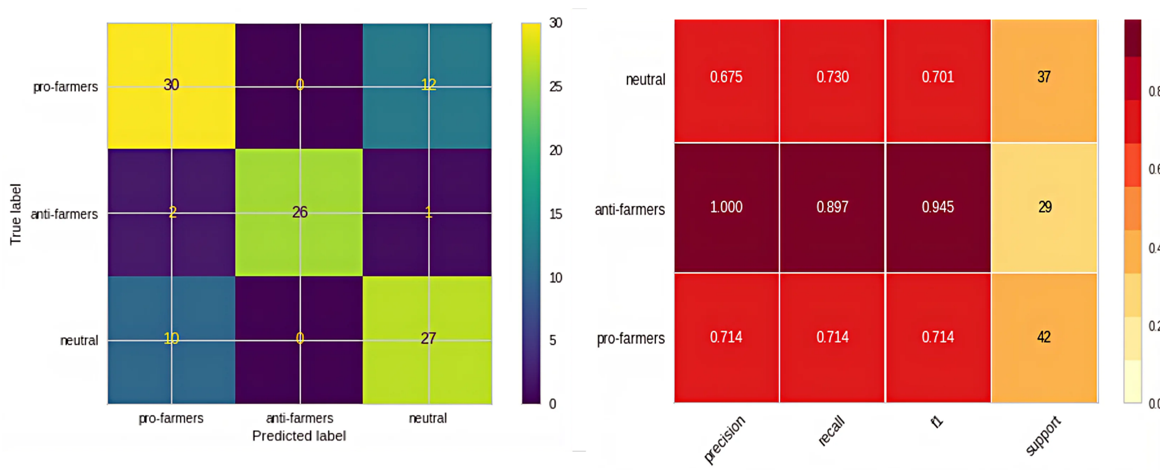

Figure 4: (f): Confusion Matrix and Heat Map for the Passive-Aggressive classifier.

5.1 Data and Implementation

For the task of bot detection, we experimented with two datasets for training the classifier. Twitter bot accounts dataset [29], with over 37438 rows and 20 columns, out of which 12425 rows were labelled as bot and the rest as human accounts, was used first, to train three classifiers, namely, Logistic regression, Gaussian Naive Bayes, and Decision Tree. The following features were extracted from the dataset for this study: 'tweetId', 'userId', 'username', 'verified', 'location', 'followersCount', 'friendsCount', 'statusesCount', 'favouritesCount', 'geo_enabled' = 'TRUE'. To reduce the dependency of the trained classifier models on the composition of the training and test sets, we experimented with two holdouts of the test set specifically, 20% and 40%. The performance of the trained classifiers is shown in Table 4.

The second dataset, known as the dataset for supervised bot detection on Twitter (1.0) [30], was used for training the same classifier models, as above, namely Logistic regression, Gaussian Naive

Table 2: Comparative Analysis of the performance of the SVC, Multinomial Naïve Bayes, Logistic Regression, Decision Tree, Random Forest, and the Passive Aggressive classifiers

Classifier	**Accuracy**	**Precision**	**Recall**	**F1 Score**
SVC	0.63	0: 0.73	0: 0.57	0: 0.64
		1: 1.00	1: 0.41	1: 0.58
		2: 0.52	2: 0.83	2: 0.64
		Avg: .60	Avg:.60	Avg:.62
Multinomial NB	**0.77**	0: 0.72	0: 0.85	0: 0.78
		1: 1.00	1: 0.67	1: 0.80
		2: 0.71	2: 0.74	2: 0.73
		Avg:.81	Avg:.63	Avg:.77
Logistic Regression	**0.77**	0: 0.69	0: 0.83	0: 0.75
		1: 1.00	1: 0.83	1: 0.91
		2: 0.73	2: 0.65	2: 0.69
		Avg:.80	Avg:.77	Avg:.78
Decision Tree	0.70	0: 0.66	0: 0.60	0: 0.62
		1: 0.93	1: 0.93	1: 0.93
		2: 0.59	2: 0.65	2: 0.62
		Avg:.72	Avg:.72	Avg:.72
Random Forest	0.75	0: 0.70	0: 0.67	0: 0.68
		1: 0.96	1: 0.90	1: 0.93
		2: 0.66	2: 0.73	2: 0.69
		Avg:.77	Avg:.76	Avg:.76
Passive Aggressive	**0.77**	0: 0.71	0: 0.71	0: 0.71
		1: 1.00	1: 0.90	1: 0.95
		2: 0.72	2: 0.73	2: 0.70
		Avg:.81	Avg:.78	Avg:.78

Table 3: Results of classification using passive-aggressive classifier.

Labels	**Tweets**
Pro-farmers	21.71%
Neutral	38.65%
Anti-farmers	39.87%

Bayes, and Decision Tree. This open access dataset has 8386 tuples out of which 3474 have been tagged as the genuine users while 4912 are classified as social spambots. The dataset has 69 features categorized as content, account information or account usage features, out of which following features were extracted from the dataset for this study: ’tweetId’, ’userId’, ’username’, ‘verified’, ’location’, ’followersCount’, ’friendsCount’, ’statusesCount’, ’favouritesCount’, ’geo_enabled’ = ‘TRUE’. As in the previous dataset, to make the trained classifiers independent of the composition of the training and test sets, we experimented with two holdouts of the test set. The performance evaluation metrics of the classifiers, with test set percentages as 20% and 40% are shown in Table 5.

Table 4: Comparative Analysis of the performance of the Logistic Regression, Gaussian Naive and the Decision Tree classifiers on test set holdouts of 20% and 40% on Dataset 1.

Metrics	Logistic Regression 20%	Logistic Regression 40%	Gaussian Naive Bayes 20%	Gaussian Naive Bayes 40%	Decision Tree 20%	Decision Tree 40%
Accuracy	68.53	67.63	52.71	57	81.809	77.75
Precision	55.77	57.58	40.28	43.10	71.33	67.00
Recall	12.85	13.38	96.00	88.00	73.08	66.34
F1 Score	0.20	0.21	0.56	0.57	0.72	0.66

Table 5: Comparative Analysis of the performance of the Logistic Regression, Gaussian Naive and the Decision Tree classifiers on test set holdouts of 20% and 40% on the Second Dataset.

Metrics	Logistic Regression 20%	Logistic Regression 40%	Gaussian Naive Bayes 20%	Gaussian Naive Bayes 40%	Decision Tree 20%	Decision Tree 40%
Accuracy	61.8	62.82	75.16	75.32	84.8	84.26
Precision	85.2	75.78	64.08	63.36	80.4	78.13
Recall	1.6	1.30	81.24	81.10	79.5	80.10
F1 Score	0.032	0.025	0.71	0.711	0.79	0.79

5.2 Results and Inferences

Tables 4 and 5 how better results from training the classifiers on the second dataset. Out of the three classifiers (Table 5), decision tree gave the best accuracy, precision and F1-score and was hence selected for classification of the tweet dataset (Section 4.1), to identify the percentage of bots.

Table 6: Results of Bot Detection using Decision Tree classifier.

Class	**Bots**
Pro-farmers	11.82%
Anti-farmers	12.75%

The decision tree classifier identified 12.75% tweets in the Anti-farmers' category as bots and 11.82% of tweets in Pro-farmers' class as bots (Table 6). A leading presence of bots to post anti-farmer tweets may justify a greater number of anti-farmer tweets, reported in the classification results of Section 4.3. However, it is interesting to note a very small difference in the percentage of bots in the Anti-farmer and Pro-farmers' category. This leads to the inference that technology was rampantly used to drive the campaigns in all the categories of users. Another important conclusion is that the study proves the use of free social media platforms like Twitter to spread misinformation and control the narrative of the social events that impact people's lives.

5.3 Characterizing BOT and Human tweets

It is important to understanding information propagation on social media in order to identify and control the spread of misinformation [28]. We therefore decided to study the tweets, classified as humans or bots, as detailed in Section 5.2, to characterize them for better understandability. For this purpose, the second dataset used in this paper, the dataset for supervised bot detection on Twitter (1.0) [30] was merged with Farmers' Protest User Dataset (Farmers Protest Tweets Dataset (CSV), no date), sourced from Kaggle, to map the user and tweet characteristics, and to help in corroborating the results of bot detection, drawing inferences and their visualization. The merged dataset contained a total of 45824 rows, 9 features by merging 37438 rows of the dataset for supervised bot detection on Twitter (1.0) [30] with 8386 rows of Farmers' Protest User Dataset (Farmers Protest Tweets Dataset (CSV), no date). Nine common features were selected for merging the two datasets, out of which seven were: 'id', 'followers_count', 'friends_count', 'favourites_count', 'statuses_count', 'verified', and 'geo_enabled'. Some features had to be transformed to match their corresponding features in the other dataset:

1. 'account_age' column had to be converted to days by the following transformation:

   feature type.sdf2['account_age_days']= df2['account_age'].apply(lambda row : row//24).

2. The feature 'geo_enabled' had to be binarized, zero symbolizing a NULL location of the tweet and one otherwise. The following transformation function was used for the same:

   data['geo_enabled'] = np.where(data['location'].notnull(), 1, 0).

3. For the age factor of user test dataset, a new feature 'account_age_days' was derived from the created time-stamp

   We undertook two kinds of analysis namely, analysis of the N-grams and a vocabulary analysis to characterize and compare tweets spread by bots vs, the texts tweeted by human users.

5.4 a N-Gram Analysis

Figure 5 shows a comparison of the N-grams drawn from PRO, ANTI and NEUTRAL stances of BOTS and HUMANS respectively and led to the discovery of the following inferences:

Some tweets regarding Indian farmers protest debate in United Kingdom parliament, held on 8 March 4:30pm, invoked by Tanmanjeet Singh Dhesi, a labour party MP in the British House of Commons, were classified as bots. Interestingly, Dhesi himself condemned the two rupees a tweet by Twitter troll factory and the fake accounts for launching fake propaganda, while addressing the House of Commons [31].

Another set of tweets under the hastags #FarmersProtest and #23March_कसािनों_के_साथ, sought to seek law on MSP and withdrawal of black laws were recognized as pro farmers and were classified as tweeted by Bonafede human twitter users by our system. The tweets raised the issues such as "Do you get paid for your work? Why must the farmer labour for nothing?" and demanded "Not meagre handouts or credit leading to debt and suicides". Some more tweets classified in this category were "@PMOIndia @narendramodi These daughters of #BhagatSingh are not for running! They don't

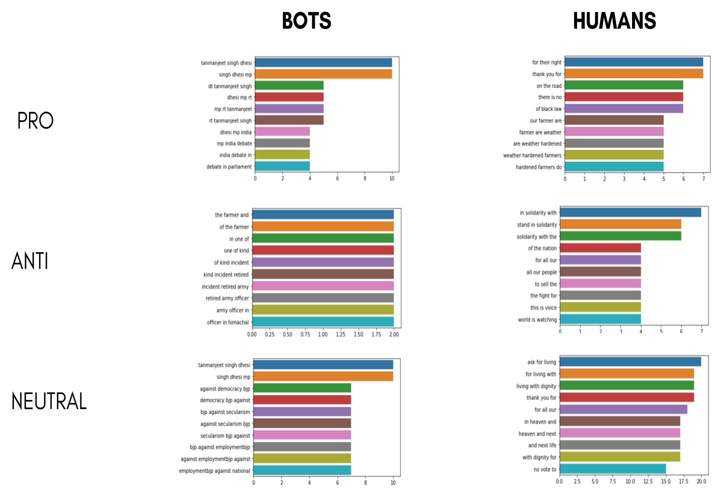

Figure 5: N-grams for stances PRO, ANTI and NEUTRAL tweeted by Bots vs. Humans

run from sending their sons to face the enemy at the border; they won't run from the #FarmersProtest. They seek law on MSP and withdraw of black laws #BoycottBJP_4Farmers", "300+ farmers died protesting for their rights & demand to repeal of Black laws passed during "pandemic" The deplorable state of India's Democracy under Modi BJP black Era!", "#FarmersRoarInBengal #FarmersProtest Farmers will write them off no matter how far they have to go in search of that victory. Farmers have roll up their sleeves for the protest and soon make the leeway of black laws and get out of woods. United we stand 💪", and "@narendramodi These sons of #BhagatSingh are not for running! They don't run facing the enemy at the border; they won't run from the #FarmersProtest. They seek law on MSP and withdraw of black laws #BoycottBJP_4Farmers"

Some tweets popularizing an incident in which a retired Army officer in Himachal Pradesh allegedly disinherited his only son to punish him for his participation in the farmers' protests against the Centre's three new agriculture legislations were flagged off as anti-farmer tweets, spread by bots by our system.

Tweets such as "Disobedience is the true foundation of liberty. The obedient must be slaves. Henry David Thoreau #FarmersProtest100Days #FarmersProtest", "@Kisanektamorcha Fact - they are making fools of themselves! The world is watching your pathetic acts. Grow a pair and stand up for what's right." , and "Yup - their lies are all being exposed. The world is watching! #23March #FarmersProtest" were recognized as anti-farmer, tweeted by humans

The tweets flagged off as neutral, tweeted by bots included texts such as "📌People's power is like a tornado !", "📌BJP would get swept away very soon ! #FarmersProtest #Elections2021", "Secularism is a threat to Indian traditions 😊"

The tweets flagged off as neutral, tweeted by humans included texts such as "@timesofindia For #Modi and friends; joys of corporate five star life. For poor, corporates mean no medical, education, transport, cooking gas, petrol. #FarmersProtest ask for living with dignity for our people. #ModiGovt4AmbaniAdani https://t.co/ppptArGT12", and "@timesofindia #ModiBJP promise

Darama; great things in heaven and next life. They Offer the nation; divide and hate. For this life; No economy, no employment, no basic needs of medical and education, nothing delivered on manifesto. #FarmersProtest ask to living today".

5.5 b Vocabulary Analysis

The Python library "Textstat" was used to calculate the difficulty level of text in the tweet using the Flesch Reading Ease (FRES) score [32], as shown in Figure 6. Figure 7 shows the difficulty line graph of the vocabulary pattern adopted by the Bots vs. the vocabulary pattern followed by the Humans. It can be seen from the figure that bots largely mimic the vocabulary pattern adopted by the humans while tweeting, except for the tweets with high difficulty level.

Figure 6: Flesch Reading Ease Score for vocabulary pattern used in the tweets tweeted by Bots vs. Humans

6. Conclusions and Future Directions

A framework to uncovers the stances taken by people on the protests by Indian farmers. Six classifiers, namely the SVC, Multinomial Naïve Bayes, Logistic Regression, Decision Tree, Random Forest, and Passive Aggressive classifier were implemented and compared to accomplish this task. Bot detection was performed for each stance to compare and contrast the usage of technology by people of different stances. The results were compared and analyzed to show that bots were used by people holding stances both against and pro farmers. Study of N-grams successfully uncovered issues being tweeted by humans vs. bots in favor of and against farmers' protest and vocabulary analysis established that the difficulty level of the vocabulary used for tweet by bots was designed to be similar to the vocabulary pattern of the tweets of human beings.

Limitations of the current work include the study of tweets over a single time period. Extension of the work to perform study over multiple time periods can give insights into the changing characteristics of tweets.

The work can also be extended to include tweets from multiple languages. Specifically, in context of farmers protest, many tweets were posted in Indian languages such as Hindi and Punjabi. Analysis of emojis, retweets, mentions etc. can also add valuable insights to the existing analysis. Classification using an ensemble model can also be explored for improving the accuracy of the work.

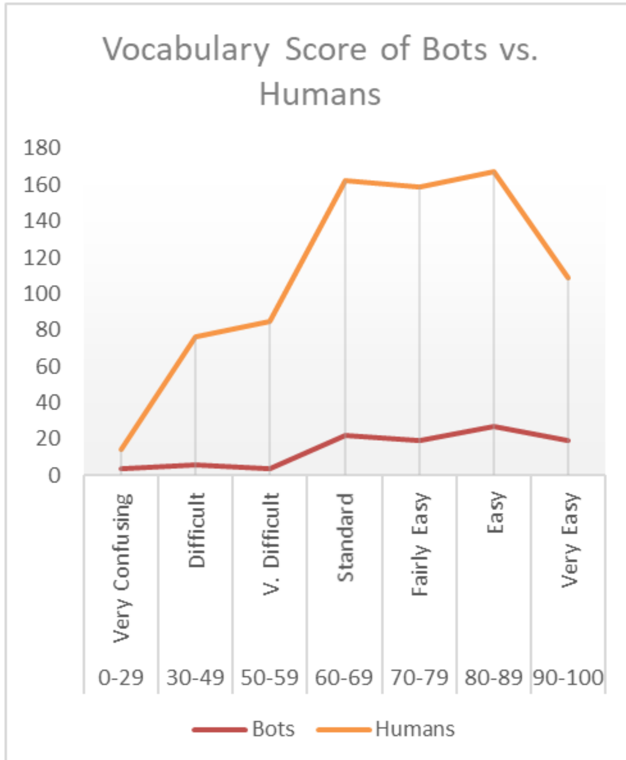

Figure 7: Graph depicting the Vocabulary Pattern Adopted by the Bots vs. the Vocabulary Pattern followed by the Humans.

References

[1] Mehrotra S, Parida J, Sinha S, Gandhi A. Explaining Employment Trends in the Indian Economy: 1993-94 to 2011-12. Econ Pol Wkly. 2014;49:49-57.

[2] Narayanan S. Understanding Farmer Protests in India. Acad Stand Pover. 2021;1:133-140.

[3] Singh L, Shergill BS. Separating Wheat From the Chaff: Farm Acts, Farmers' Protest and Outcomes. Millennial Asia. 2021;12:390-410.

[4] Bhatnagar V, Kochhar S. Modeling Support Changes in Streaming Item Sets. Int J Syst Sci. 2006;37:879-891.

[5] Kochhar SK, Gupta A. Deploying Change Modeling to Study the Evolution of COVID-19 Related Menstrual Health Issues. Adv Artif Intell Mach Learn. 2023;3:1460-1481.

[6] Kochhar SK, Kaur R. Breaking the Taboos: Deploying Knowledge Differentiation to Study COVID-19 Ramifications on Women's Menstrual Health. Int. J. Comput. Digit. 2023;13:1-1.

[7] https://www.kaggle.com/datasets/prathamsharma123/farmers-protest-tweets-dataset-csv.

[8] Neogi AS, Garg KA, Mishra RK, Dwivedi YK. Sentiment Analysis and Classification of Indian Farmers' Protest Using Twitter Data. Int J Inf Manag Data Insights. 2021;1:100019.

[9] Tiwari D, Nagpal B. KEAHT: A Knowledge-Enriched Attention-Based Hybrid Transformer Model for Social Sentiment Analysis. New Gener Comput. 2022;40:1165-1202.

[10] Singh A, Kalra N, Singh A, Sharma S. Sentiment Analysis of Twitter Data During Farmers' Protest in India Through Machine Learning. In: International Conference on Computer Science and Software Engineering (CSASE), IEEE. 2022:121-126.

[11] Mahajan R. Predicting Democratic Protests Paradigm from Twitter: Using Deep learning BiLSTM Model. 2022. TechRxiv Preprint: https://doi.org/10.36227/techrxiv.19127435.v2

[12] Sahithi C, Sreeja Y, Akhil S, Taruni K, Sobin CC. Twitter Sentiment Analysis Using Machine Learning Techniques: A Case Study Based on Farmers Protest. In: Machine Learning in Information and Communication Technology. Proceedings of the ICICT SMIT. 2022:187-196.

[13] Mishra D, Akbar SZ, Arya A, Dash S, Grover R, Pal J. Rihanna Versus Bollywood: Twitter Influencers and the Indian Farmers' Protest. 2021. ArXiv preprint: https://arxiv.org/abs/2102.04031

[14] Dash S, Mishra D, Shekhawat G, Pal J. Divided We Rule: Influencer Polarization on Twitter During Political Crises in India. Proceedings of the International AAAI Conference on Web and Social Media. 2022;16:135-1466.

[15] Monteiro S. Farmer Protests in India and the Mobilization of the Online Diaspora on Twitter. SSRN Journal. 2021.

[16] Gómez-Suta M, Echeverry-Correa J, Soto-Mejía JA. Stance Detection in Tweets: A Topic Modeling Approach Supporting Explainability. Expert Syst Appl. 2023;214:119046.

[17] Upadhyaya A, Fisichella M, Nejdl W. Towards Sentiment and Temporal Aided Stance Detection of Climate Change Tweets. Inf Process Manag. 2023;60:103325.

[18] Kamble A, Badgujar P, Kadam A, Shah D, Kadam AJ. Stance prediction of tweets on farmers protests in India. IJRASET. 2022;10:4288-4295. Int J Res Appl Sci Eng Technol. 2022;10:4288-4295.

[19] Neha K, Agrawal V, Buduru AB, Kumaraguru P. The Pursuit of Being Heard: An Unsupervised Approach to Narrative Detection in Online Protest. International Conference on Advances in Social Networks Analysis and Mining (ASONAM). 2022:256-260.

[20] Chavoshi N, Hamooni H, Mueen A. Debot: Twitter Bot Detection via Warped Correlation. In: 16th International Conference on Data Mining (ICDM). IEEE. 2016;18:28-65.

[21] Kantepe M, Ganiz MC. Preprocessing Framework for Twitter Bot Detection. International Conference on Computer Science and Engineering. 2017:630-634.

[22] Wei F, Nguyen UT. Twitter Bot Detection Using Bidirectional Long Short-Term Memory Neural Networks and Word Embeddings. 2020. ArXiv Preprint: https://arxiv.org/pdf/2002.01336.pdf

[23] Feng S, Tan Z, Li R, Luo M. Heterogeneity-Aware Twitter Bot Detection With Relational Graph Transformers. Proceedings of the AAAI Conference on Artificial Intelligence. 2022;36:3977-3985.

[24] Kochhar SK, Sharma A, Jain D, Rani G, Gaur V. A Blended Approach to Analyze Indian Stock Market During COVID-19. Int J Comput Digit Syst. 2023;14: 9-30.

[25] Crooks A, Croitoru A, Stefanidis A, Radzikowski J. Earthquake: Twitter as a Distributed Sensor System. Trans GIS. 2013;17:124-147.

[26] Varshney A, Kapoor Y, Chawla V, Gaur V. A Novel Framework for Assessing the Criticality of Retrieved Information. Int J Comput Digit Syst. 2022;11:1229-1244.

[27] https://research.aimultiple.com/human-annotated-data.

[28] Soni A, Jain S, Karki M, Gaur V, Kochhar SK. Topic Modelling, Classification and Characterization of Critical Information. Int J Comput Digit Syst. 2023;14:125-137.

[29] https://www.kaggle.com/dsv/1623389

[30] https://zenodo.org/records/5574403

[31] https://timesofindia.indiatimes.com/city/ludhiana/dhesi-hits-out-at-bjym-leader-for-defamatory-propaganda/articleshow/90855272.cms

[32] https://simple.wikipedia.org/w/index.php?title=Flesch_Reading_Ease&oldid=8795738.