

Correlating Crime and Social Media: Using Semantic Sentiment Analysis

Rhea Mahajan¹, Vibhakar Mansotra²

Department of Computer Science and IT, University of Jammu, Jammu, India

Abstract—Crimes occur all over the world and with regularly changing criminal strategies, law enforcement agencies need to manage them adequately and productively. If these agencies have prior data on the crime or an early indication of the eventual felonious activity, it would encourage them to have some strategic preferences so that they can deploy their restricted and elite assets at the spot of a suspected crime or even better explore it to the point of anticipation. So, integration of social media content can act as a catalyst in bridging the gap between these challenges as we are aware of the fact that almost all our population uses social media and their life, thoughts, and, mindset are available digitally through their social media profiles. In this paper, an attempt has been made to predict crime pattern using geo-tagged tweets from five regions of India. We hypothesized that publicly available data from Twitter may include features that can portray a correlation between Tweets and the Crime pattern using Data Mining. We have further applied Semantic Sentiment Analysis using Bi-directional Long Short memory (BiLSTM) and feed forward neural network to the tweets to determine the crime intensity across a region. The performance of our proposed approach is 84.74 for each class of sentiment. The results showed a correlation between crime pattern predicted from Tweets and actual crime incidents reported.

Keywords—Crimes; social media; Twitter; BiLSTM; semantic sentiment analysis

I. INTRODUCTION

With the upsurge of online media, the web has become an energetic and enthusiastic domain wherein billions of people all around the globe associate, offer, post and share their daily activities. Data which is generated by Social Networking Sites is an extremely large data which is growing exponentially at an unprecedented pace. Mountains of raw data is generated daily by individuals on these social networking sites [1]. These sites have changed our lives drastically and their impact on society cannot be overlooked. Facebook, Instagram, and, Twitter are the most popular social net-working sites with 2.5 billion, 1 billion and, .336 billion users respectively all over the world and 241 million, 40 million and 37 million users respectively in India. These numbers vary every day and this rapid growth in the volume of users has provided the predictive ability in extensive fields such as personality prediction [2], stock market trends [3], election results [4], the box office performance of movies, etc. [5]. Social media allows its users to share their apprehensions, ideas and daily activities on the web. This shared content by the individuals when joined together provides a rich resource of naturally occurring data. Status updates from Facebook, tweets from Twitter and pictures from Instagram provide information about the social behavior of its users. Our enchantment to social media has grown in the last

decade to the pinnacles which can only be compared to the billions they have been valued for. Its growth and impact is unparalleled, to say the least. While they have developed into different entities, their usefulness and social impact have always been a subject of debate. The influence can be judged from the fact that the fake news travels or gets viral faster than the real and valuable information. This effect has only increased and sometimes does get morphed into something unpleasant and hostile, where these interactions have gravitated towards the unconstructive side of things which includes bullying, trolling, stalking, social media trials etc. This impact is also tipping the scale towards more and more pessimism.

The present crime prediction models commonly depend on relative static highlights including long haul verifiable data, topographical data, and, segment data. This data changes gradually after some time, which means these conventional models couldn't catch the transient varieties in criminal activities [6]. The primary downside of these models is that they diminish the social setting to verifiable criminal records while disregarding information on the social conduct of the users of available on social networking sites including the victim and the criminal as keeping an on eye the social behavior information of an enormous society is a difficult and challenging task [7].

Twitter is picked over other online social media sites because it is one of the most popular micro-blogging sites for its political potential value and transparency and the way that anybody can get to geo-tagged tweets created in a given region or territory. Moreover, people are very vocal about their views and opinions and do not hesitate to express them through their tweets. So, this research is inspired by the fact that the enormous data available on these sites can be used to bring out a significant amount of information for the administration and law authorities which will eventually be used to predict criminal behavioral patterns.

In this paper, an attempt has been made to predict crime pattern using geo-tagged tweets from five regions of India. We hypothesized that publicly available data from Twitter may include features that can portray a correlation between Tweets and the Crime pattern using Data Mining. We have further applied Semantic Sentiment Analysis using BiLSTM and feed forward neural network to the tweets to determine the crime intensity across a region. BiLSTM is a variant of LSTM and is more powerful than LSTM as it overcomes the problem of gradient explosion that occurs in LSTM. The results showed correlation between crime pattern predicted from Tweets and actual crime incidents reported. Fig. 1 shows framework of the proposed research.

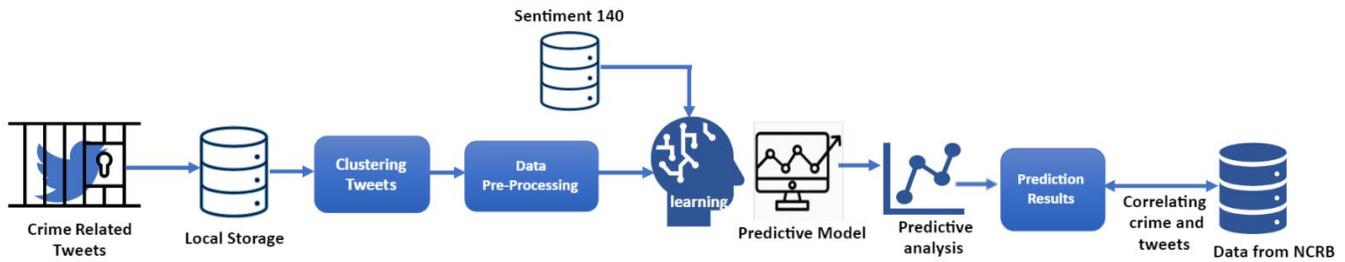


Fig. 1. Framework of the Research.

This paper is organized as follows: After brief introduction in Section I, Section II provides a summary of related works in area of crime Prediction using data from social networking sites. Section III gives the description of the data set and process of data acquisition. Section IV describes the proposed approach, which is followed by Section V, where performance of the classifier on various evaluation metrics is presented. Section VI and Section VII presents correlation analysis and hypothesis testing, respectively. Finally, we have concluded the paper with some future work guidelines in Section VII.

II. RELATED WORKS

Recent studies have attempted to fit in data from Twitter into their predictive models for crime assessment. The purpose of integrating Twitter data for crime prediction is to take into account significant amount of information available on Twitter about the social conduct and mobility of the users. Geber [8] is the first one to introduce social media content to model crime prediction. To address the use of tweet content in determining the crime pattern of a particular location, Geber used latent Dirichlet allocation on tweets that showed an improvement on models using conventional historic data as crime predictors for stalking, criminal damage and gambling. Even though, it is the foremost study to examine tweet text, Gerber's use of LDA is challenging given that it is an unsupervised technique, which meant correlation between word clusters and the crimes are not driven by previous theoretical insights. This resulted in correlations that seemed comparatively worthless. Wang et al. [9] extracted event-based topics from real time tweets to predict hit-and-run incidents in Virginia. Even though their approach was novel, the source of data was limited to a set of manually selected news portals and the massive amount of information backed by the citizens was neglected.

Chen et al. [10] utilized the sentiment in Tweets together with weather data in KDE for predicting the time and location of the theft. However, their study was restricted to spatial information such as weather data for specific time and location Brandt et al. [11] studied the relationship between mobile populations as recorded by Twitter's geotagging facility and the location of different types of crime. They concluded the absence of tweets was predictive of assaults and thefts. Similarly, Malleon et al.[12] have used a number of geographic analysis methods to model crime risk using tweets

for mobile populations. The main drawback of these studies was that tweet text was not taken in consideration, instead focusing purely on geolocation data. It was also concluded that KDE is a location dependent technique cannot be easily generalized. There may be some type of crime that does not occur in the vicinity of previous locations and incidents and the population of an area can change frequently.

In addition to the above studies, sentiment analysis has also been a key instrument in Crime detection and prevention. Zainuddin et al. [13] applied sentiment analysis to crime related tweets through the use of model that was based on Natural Language Processing techniques and SentiWordNet, the model had the capability to detect the subjectivity of crime and then predicted crime through hate tweets. Machine learning algorithms has also been used to solve the task of sentiment analysis of Tweets [14][15]. Pang et al. [16] performed a comparative study involving algorithms such as Naïve Bayes, Support Vector Machine and maximum entropy to determine sentiment polarity for movies reviews. These studies were effective but ignored the semantics to capture the meaning of the tweets.

In this paper, we have tried to overcome the drawback of above studies by collecting real time tweets for a period of 21 days across five regions of India to capture dynamic movement of the user. Further, we have used combination of BiLSTM and feed forward neural network to find sentiment polarity of the Tweets. The strength of BiLSTM is that it provides extra training by traversing the text twice from left to right and right to left, there by extracting the semantics of the words in context of the information preceding and succeeding it and therefore can capture long term contextual dependencies and global features from the sequential text.

So, keeping in view the various trends of research carried out using social media in particular Twitter, it needs no mention that social media mining is an important area of research and by the application of various data mining techniques can generate very impressive and interesting patterns as well as outcomes which can be analysed, interpreted and can be used for the benefit of the society especially in crime Prediction and detection and in the scenario of evolving protest and riots. Table I lists some of the important works done in area of crime Prediction using tweets.

TABLE I. LISTS SOME OF THE IMPORTANT WORKS DONE IN AREA OF CRIME PREDICTION USING TWEETS

Author	Application	Technique used	Dataset Used	Evaluation results
Geber(2014)[8]	Twitter-based model for crime trend prediction to determine crime rates in the prospective time frame.	Text analysis- filtering including stop word reduction and low-frequency term reduction Predictive analysis-linear support vector classifier	Historic tweets were collected from Chicago city for a period of three years combined with other datasets such as unemployment rates and weather conditions.	Results revealed correlation between features extracted from content as content-based features and the crime trends.
Wang et al. (2012) [9]	Twitter based criminal incident prediction on Hit and Run cases.	Text analysis-Semantic Role Labelling (SRL) and Dirichlet allocation Predictive analysis-linear modelling	Real-time Tweets using Twitter API	F1 score-80%of verbal SRL and 72%of nominal SRL
Chen et al.(2015) [10]	Twitter based model for time and location prediction in which specific type of crime will occur.	Text analysis-Sentiment Analysis by the lexicon-based method Predictive analysis-linear modelling via logistic regression Comparative analysis- hot spot mapping with kernel density estimation(KDE)	GPS tagged tweets from Chicago city of US; combined with weather data and historic crime data from Chicago	Performance measure -Area Under Surveillance Curve(AUC) Predicted AUC-0.67 Actual AUC-0.66 Error-1.5%
Aghababaei et al. (2016) [17]	Twitter based criminal incident prediction on 25 types of crime.	Text analysis-statistical language processing and spatial modelling Predictive analysis-logistic regression Comparative analysis- hot spot mapping with kernel density estimation(KDE)	Geo-tagged tweets from Chicago city of US and historic criminal data.	Of the 25 crime types, 19 showed improvements in Area Under Surveillance Curve (AUC) when adding twitter topics to the KDE-only model.
Almehmadi (2017) [18]	Twitter-based model to predict crime by analysing language usage in Tweets as a valid measure.	Text analysis-WEKA and Ranker algorithm Predictive analysis- SVM classifier was used to classify the data to the proposed class: offensive or non-offensive language	GPS tagged tweets were collected from Houston and New York for three months.	With a binary SVM classifier, 96.19% correct classification accuracy was achieved. Results show accuracy by class for cross-validation with ROC 77%.
Ristea (2018) [19]	Twitter based opinion mining and spatial crime distribution for hockey events in Vancouver.	Spatial clustering, opinion mining and regression analysis was used in order to find meaningful explanatory variables for crime occurrences.	Crime data for Vancouver was obtained from Vancouver Open Data Catalogue. Geo-referenced tweets were obtained using the Twitter Streaming Application for 2014-2016 i.e. for two hockey seasons.	Results showed the influence of social media text analysis in describing the geography of crime along with the importance of additional criminogenic factors
Siriraya et al. (2019) [20]	Twitter based crime investigation tool that provides contextual information about crime incidents by visualizing spatial and time-based characteristics of a crime.	Various tweet vectorization strategies (pre-trained word vectors from the GloVe model4 , Doc2Vec etc.) and classification models (Logistic Regression, SVM etc.) were used to investigate the performance in classifying negative tweets.	Geo-tagged tweets were collected for a period of one year from San Francisco.	The results showed that using the GloVe model to represent the tweet words and the linear kernel SVM to perform binary classification resulted in the best performance (a stratified 5 fold-cross validation showed an F-score of 0.80 as opposed to 0.70 for the SVM-Doc2Vec model)

III. DATASET DESCRIPTION

We began our research by identification of five regions of India; determined by Nation Crime Records Bureau as per prevailing crime rate. They are Uttar Pradesh, Madhya Pradesh, Maharashtra, Bihar and Delhi-NCR. Then, we collected crime related Tweets from Twitter and crime data from various national and local online news portals and NCRB¹ from 2 December 2019 to 22 December 2019. Crime against women, Crime against children, Murder, Suicide, Cyber Crime and violence due to riots and protests were six categories of crime for which data has been collected.

To extract the data from Twitter, we need to create an account on Twitter. Then, Twitter requires its users to register an application. This application authenticates our account and provides the user a access token and consumer key which then can be used to connect with twitter and download tweets. Crime related and Geo-tagged real-time tweets were collected from above mentioned Indian regions using geo-tag filter of Twitter Streaming API.

We ran the data collection process which resulted in over 30,000 tweets from 512 users in our database shown in Fig. 2. This data contains information such as user ID, the screen name, number of followers, date, the tweet itself, device used to post the tweet source, the user-defined location, coordinates, agender, retweets and user mentions.

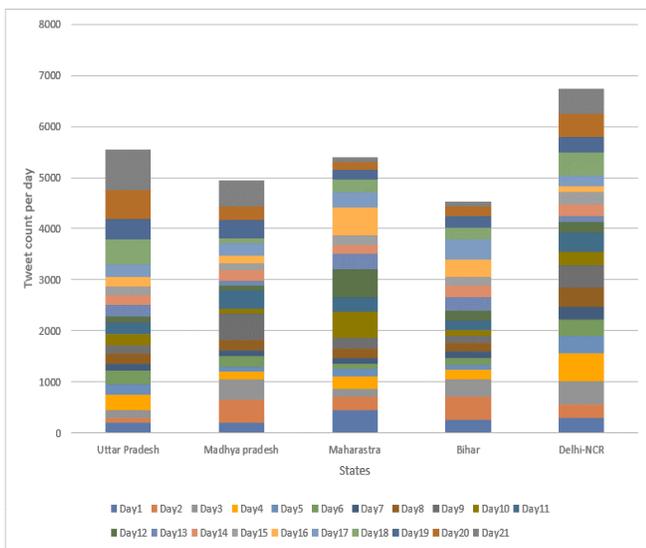


Fig. 2. Distribution of Tweets Day-wise from Five Regions across India.

An English language filter was applied and 29 different keywords were used while streaming real-time Tweets. Tweets were collected using a keyword search strategy [21]. Keywords used to identify a specific crime type were rape, dowry, abduction, kidnapping, child labor, depression, anxiety protest, etc. are listed in Table II. The Tweets were extracted in JSON format imported to a pandas Data frame in Python and were finally downloaded in CSV file format. We extracted the tweets using the geo-tag filter option of Twitter's streaming API and bounding box. Tweets were then clustered on the basis

of similarity i.e. crime type and location using K-means clustering and Jaccard Distance metric to make them organized as shown in Fig. 3.

TABLE II. KEY WORDS

S. No.	Crime Type	Key Words
1.	Crime against women	dowry,rape,assault,abduction,metoo
2.	Crime against children	kidnapping, child labor, minor
3.	Murder	kill, gun, shot, arms, murder
4.	Suicide	Depression, suicide, anxiety mentalhealth
5.	Cybercrime	fraud, stalking, trolling, bullying
6.	Violence due to protest and riots	antiCAA, anti-NRC, hateIndia, protest ,just violence, riots

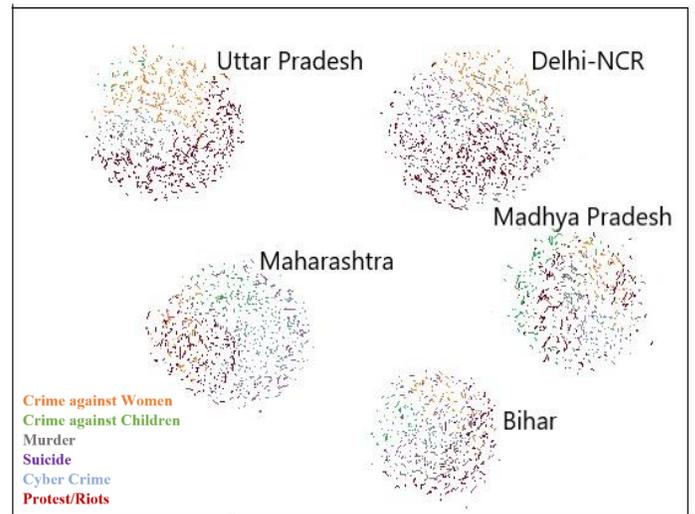


Fig. 3. Tweets Clustered on the basis of Crime Type and Location.

Once the tweets were collected, NLTK² package with pip package manager in Python was used for processing text in tweets. The steps include removal of extra places, URL, stop words, tokenization which refers to dividing the text into a sequence of words and lemmatization i.e. reducing different types of words with similar meaning with their root. Tweets were then embedded into vector form using word2vec vectors using Google News vectors for obtaining vector representations of words with Skip-gram architecture.

IV. SEMANTIC SENTIMENT ANALYSIS

We have used BiLSTM and feed forward neural network as shown in Fig. 5 to determine the sentiment polarity of the tweets. Conventional RNNs can only process the data in one direction and none of the attention is given to process future information. To overcome this limitation, the concept of Bidirectional RNN came into existence. Bi-directional RNN has the ability to traverse the data in both directions with different hidden units acting as forward layers and backward layers. Bidirectional LSTM (Bi-LSTM) was introduced by Graves et al. [22] combining Bidirectional RNN with LSTM

¹National crime records bureau <https://ncrb.gov.in/en>

² <https://www.nltk.org/book/ch01.html>

cell. The output of forward states is not used as an input for backward states and vice-versa in BiLSTM thus, overcoming the problem of gradient explosion.

Sentiment140³ data set from Kaggle has been used to train our Classifier. It contains 1.6 million tweets extracted using the Twitter API. The tweets have been annotated as negative, positive and neutral with respective sentiment scores and they can be used to detect sentiment of the brand, product, or topic on Twitter .The input to the BiLSTM is set of word vectors $W=\{w_1, w_2, \dots, w_n\}$. At each step from $i \dots n$, a forward Long Short Memory (LSTM) takes the word embedding of word w_i and previous state as inputs, and generates the current hidden state. A backward LSTM reads the text from w_n to w_i and generates another state sequence. The hidden state h_{si} for word w_i is the concatenation of h_{si} vector forward and h_{si} vector backward thereby capturing the semantics of the word in context of the information preceding and succeeding it . The output of BiLSTM is fed into the feedforward neural network. Finally, the probability of a tweet t_i belonging to a sentiment class S is obtained using Softmax function

$$p(t_i|\hat{\theta}) = \frac{\exp(\beta_i^T \hat{\theta})}{\sum_{j=1}^s \exp(\beta_j^T \hat{\theta})}$$

where β_i (weight vectors) are parameters in SoftMax layer. The activation function for neural network is ReLU. In order to prevent the over-fitting in the training process and co-adaptations of units, dropout of 0.5 is applied.

HYPERPARAMETERS

Epochs	50
Learning rate	10^{-3}
Optimizer	Adams
Max length	148
Dropout	0.5
Batch size	64
Nodenum	128
Vector size	300

The output from this sentiment analyser in the form of heat map and corresponding sentiment score is shown in Fig. 4. In the heat map, intensity of blue colour shows the accumulated sentiment of Tweets on a particular day. Tweets that were categorized as Negative (dark blue) were identified as contributing to the crime intensity of that place.

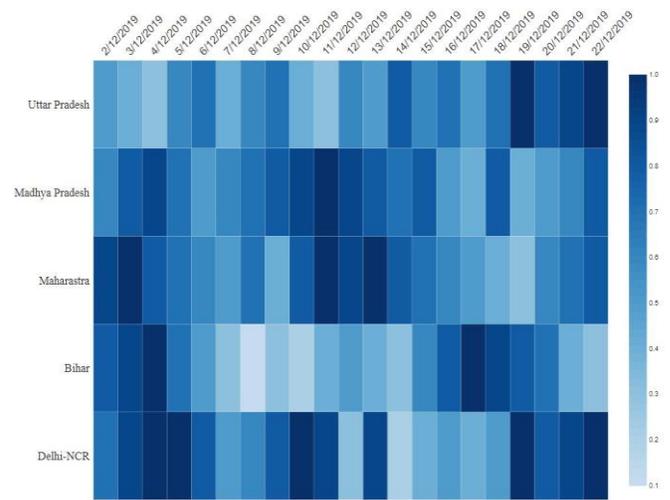


Fig. 4. Heat Map and Corresponding Sentiment Score during Observed Time.

V. EVALUATION METRICES

We have evaluated our classifier on various metrics. Precision, Recall, and F-score have been used for assessing the performance of the proposed model by finding the Confusion Matrix which contains information about actual and predicted classifications done by a classification system. The performance of classifier shown in Table III has been calculated by taking the average of the three metrics for each class of sentiment.

- Precision = True Positive / (True Positive + False Positive)
- Recall = True Positive / (True Positive + False Negative)
- F1-Measure = $2 * (\text{Precision} * \text{Recall} / (\text{Precision} + \text{Recall}))$

TABLE III. PERFORMANCE OF THE CLASSIFIER

Positive Sentiment			
Precision	Recall	F-Measure	Performance
91.54	92.82	92.18	92.18
Neutral Sentiment			
Precision	Recall	F-measure	Performance
80.70	82.10	81.39	81.39
Negative Sentiment			
Precision	Recall	F-measure	Performance
78.30	83.30	80.72	80.77

³ <https://www.kaggle.com/kazanova/sentiment140>

Algorithm

Input: Sentiment140(T_{train}), Real-time Crime Related Geo-tagged tweets

Output: Probability of Tweet belonging to sentiment class s

Step 1 : Install dependencies tweepy, tensorflow, keras

Step 2: Import packages os, json pickle, numpy, myplot

Step 3: Authentication with twitter using access keys and tokens

Step 4: Extract Tweets using Twitter Streaming API using geo-filter and keyword search strategy

Step 5: Cluster the Tweets on basis of similarity using Jaccard distance.

Step 5 : Obtain the set of word vectors $t=\{w_1, w_2, \dots, w_n\}$ using word2vec from Google News

Step 6: Process the tweets using NLTK package and prepare the data for model fitting

Step 7: Initialize BiLSTM model hyperparameters

Step 8: For each sentence $t \in T_{train}$

- Generate expression sequence and output eigenvector $h_s=\{h_{s1}, h_{s2}, \dots, h_{sn}\}$ through BiLSTM
- The output of BiLSTM is fed to feed forward neural network
- Apply Back propagation algorithm to adjust model parameters and word vectors;
- Apply activation function Softmax to calculate the output probability of Tweet belonging to sentiment class S .

$$p(t_i|\hat{\theta}) = \frac{\exp(\beta_i^T \hat{\theta})}{\sum_{j=1}^s \exp(\beta_j^T \hat{\theta})}$$

Step 9: For each $t \in T_{test}$

Classify the sentiment polarity of the real time tweets using trained model.

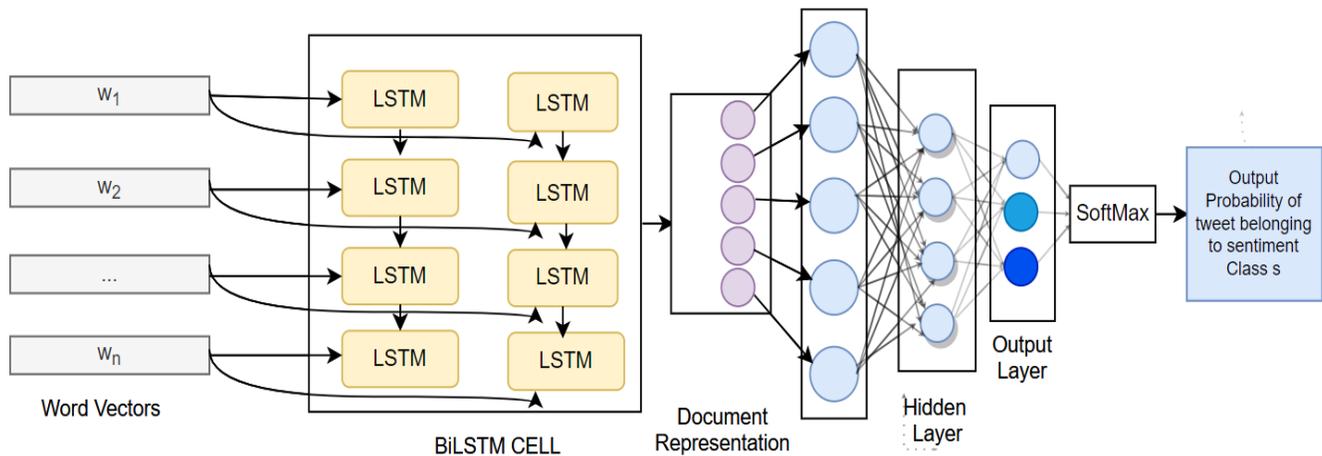


Fig. 5. Working of Sentiment Analyser.

VI. CORRELATING CRIME AND TWEETS

We have used Pearson’s correlation coefficient (r) as a statistical measure of the strength of a linear relationship between predicted crime pattern (Fig. 7) from tweets and actual crime reported by news portals and media (Fig. 6). The correlation(r) between crime predicted and crime reported is shown in Table IV.

Pearson’s correlation coefficient (r)

$$r = \frac{N\sum xy - (\sum x \sum y)}{\sqrt{[(N\sum x^2 - (\sum x)^2)(N\sum y^2 - (\sum y)^2)]}}$$

$$r^2 = \frac{(N\sum xy - (\sum x \sum y))^2}{\sqrt{[(N\sum x^2 - (\sum x)^2)(N\sum y^2 - (\sum y)^2)]^2}}$$

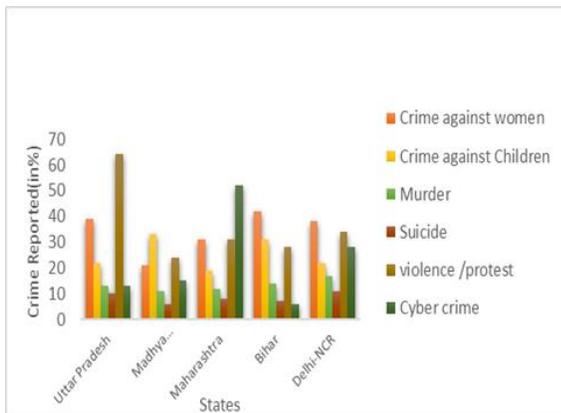


Fig. 6. Crime Incidents Reported from 2 Dec 2019 to 22 Dec 2019 as Per NCRB and News Portals.

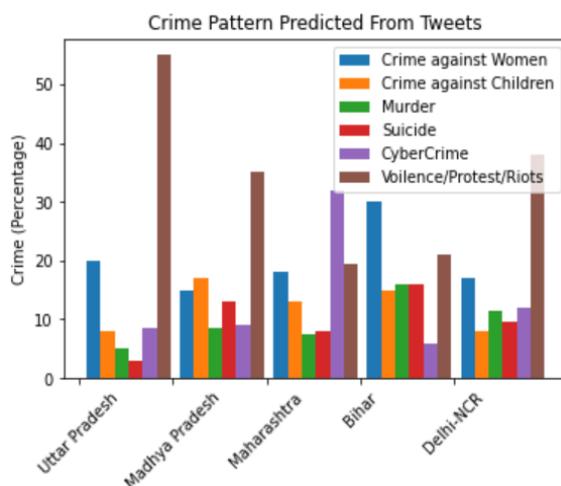


Fig. 7. Crime Pattern Predicted from Tweets from 2 Dec 2019 to 22 Dec 2019.

TABLE IV. HYPOTHESIS TESTING

	r	r ²	t-test stat	p-value
Crime against Women	0.7927	0.6284	2.2522	0.1079
Crime against Children	0.7978	0.6365	2.2918	0.1057
Murder	0.3722	0.1385	0.6945	0.5372
Suicide	-0.2218	0.0492	-0.3939	0.7209
Cyber Crime	.9499.	0.9023	5.2639	0.1333
Violence protest	0.8068	0.6509	2.3652	0.0989

VII. HYPOTHESIS TESTING

Null hypothesis Ho: Publicly available data from Twitter do include features that can portray a correlation between Crime pattern predicted from Tweets and the actual Crime incidents reported.

Alternative hypothesis Ha: Publicly available data from Twitter do not include features that can portray a correlation between Crime pattern predicted from Tweets and the actual crime reported.

p-value: The p-value tells us if the result of an experiment is statistically significant (significance level=0.05). The p-value is calculated using a t-distribution, with (n-2) degree of freedom.

$$t\text{-test Statistics} = \left\{ \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}} \right\}$$

Since the p-value is larger than 0.05 as shown in Table IV, we fail to reject null hypothesis and we cannot conclude that a significant difference exists.

VIII. CONCLUSION

In this paper, we have tried to predict crime pattern using geo-tagged tweets from five regions of India. We hypothesized that publicly available data from Twitter may include features that can portray a correlation between Tweets and the Crime pattern using Data Mining. We have further applied Semantic Sentiment Analysis using BiLSTM and feed forward neural network to the tweets to determine the crime intensity across a region. BiLSTM is a variant of LSTM and is more powerful than LSTM as it overcomes the problem of gradient explosion that occurs in LSTM. The purpose of combining these two approaches was to exploit the strength of BiLSTM and feed forward neural network. The performance of the classifier is 84.74 for each class of sentiment. The results showed correlation between crime pattern predicted from Tweets and actual crime incidents reported. The main limitation of our study was unavailability of geo-tagged tweets as more than half of twitter users prefer to conceal their location due to privacy issues. We hope to further make our research effective by using open mapping from Google. The data used in the research is available on-line on Twitter to support further investigation.

REFERENCES

- [1] M. A. Russell, Mining the Social Web. OReilly, 2nd ed. October 2013.
- [2] M. Suresh, D. Nagendrababu, S. Nakkiran, and K. Vanjinathan, “A Survey on Personality Prediction Using Digital Footprints in Social Media”, International Research Journal of Engineering and Technology, vol. 03, no. 02, pp. 1787–1793, 2016.
- [3] Y. Wang , “Using Social Media Mining Technology to Assist in Price Prediction of Stock Market”, IEEE International Conference on Big Data and Analytics ICBDA, vol. 07, pp.143-147, 2016.
- [4] J. Ramteke and D. Godhia, “Election result prediction using Twitter sentiment analysis,” International Conference on Inventive Computation Technologies (ICICT), pp.122-128 ,2016.
- [5] S. Shim and M. Pourhomayoun, “Predicting Movie Market Revenue Using Social Media Data”, IEEE International Conference on Information Reuse and Integration, vol. 04, no.1, 2017.
- [6] S. Sathyadevan, M. S. Devan, and S.S.Gagadharan, “Crime analysis and prediction using data mining”, International Conference on Soft Computation. ICNSC 2014., vol. 14, pp. 406–412, 201.
- [7] J. Chan and L.B. Moses, “Is Big Data challenging criminology?,” Theoretical Criminology, vol. 20, issue 1, pp. 21-39, 2016.

- [8] M. S. Gerber, "Predicting crime using Twitter and kernel density estimation", *Decision Support System*, vol. 61, no.1, pp. 115–125, 2014.
- [9] X. Wang, M.S. Gerber and D.E. Brown, "Automatic crime prediction using events extracted from twitter posts, in *Social Computing, Behavioral-Cultural Modeling and Prediction*," Springer, 2012, pp. 231–238.
- [10] X. Chen, Y. Cho and S.Y. Jang, "Crime prediction using twitter and weather," *Systems and Information Engineering Design Symposium (SIEDS)*, IEEE, 2015, pp. 63–68.
- [11] T. Brandt, J. Bendler, and D. Neumann, "Information & Management Social media analytics and value creation in urban smart tourism ecosystems," *Information and Management*, vol. 54, no. 6, pp. 703–713, 2017.
- [12] N. Malleson and M.A. Andresen, "The impact of using social media data in crime rate calculations: shifting hot spots and changing spatial patterns," *Cartography and Geographic Information Science*, 42(2), pp.112–121, 2015.
- [13] N. Zainuddin, A. Selamat, and R. Ibrahim, "Improving Twitter AspectBased Sentiment Analysis Using Hybrid Approach," *Intelligent Information and Database Systems*, vol. 9621, N. T. Nguyen, B. Trawiński, H. Fujita, and T.-P. Hong, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 151–160.
- [14] B. Gokulakrishnan, P. Priyanthan, T. Ragavan, N. Prasath, and As. Perera, "Opinion mining and sentiment analysis on a Twitter data stream," 2012, pp. 182–188.
- [15] V. N. Patodkar and S. I.R, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," *IJARCCCE*, vol. 5, no. 12, pp. 320–322, Dec. 2016.
- [16] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in *Proceedings of the ACL-02 conference on Empirical methods in natural language processing* vol 10, 2002, pp. 79–86.
- [17] S. Aghababaei and M. Makrehchi, "Mining Social Media content for crime Prediction," *IEEE/WIC/ACM International Conference of Web Intelligence*, vol. 13, pp. 526–531, 2016.
- [18] A. Almealmadi, "Language Usage on Twitter Predicts Crime Rates," *Security of Information and Networks 2017*, pp. 307–310.
- [19] A. Ristea, M. Leitner, and M. A. Andresen, "Opinion mining from Twitter and spatial crime distribution for hockey events in Vancouver," *AGILE 2018*, pp. 1–7.
- [20] P. Siriaraya, Y. Wang, and A. Jatowt, "Witnessing Crime through Tweets :A Crime Investigation Tool based on Social Media," *International Conference on Advances in Geographic Information Systems 2019*, pp. 568-571.
- [21] L. Alfantoukh and A. Durresi, "Techniques for collecting data in social networks", *International Conference on Network-Based Information Systems NBIS*, vol. 20, pp. 336–341, 2014.
- [22] A. Graves and J. Schmidhuber, "Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures," *Neural Networks*, vol. 18, no. 5-6, pp. 602–610, June/July 2005.