Time Series Based Trend Analysis for Short Term Social Media Posts in Sinhala Language

Abstract **—Due to the openness of social media, different topics are being discussed virtually in social media platforms. This research is aimed to scope down the hate speech performed on social media in short life cycles. Sinhala language and Sinhala words written in English text was considered for analysis. These trends differ with the context of the post. Outcomes of the discussions are profoundly impacted by society and rapidly disseminated among people. This study was performed based on the disseminated content performing time series analysis and mainly focused on hate speech trends. Forecast models are used based on factors like seasonality, trend, cyclic and related factors. In this research, the forecast models were used for seasonality, and the trends were modeled based on exponential smoothing. Hourly sensitive trends analysis and forecasting were performed using the Prophet framework. This experiment tightly coupled with timestamps. The high velocity of interaction in the first half of the post was identified, and the latter part content consists of the outer cycle of the initiators of the comment. Finally, this paper discusses a trend line of hate speech posts in social media platforms, mainly for any outbreaks or short period related social media engagements. This experiment was carried out in the COVID 19 out break in sri lanka to test as short post cycles.**

Keywords—**Trend Analysis, Time Series analysis, Social Media Analysis**

I. INTRODUCTION

COVID 19 is an epidemic that spreads with high impact in many aspects like economy, civil services, community life [1,p.19]. The first victim of the virus was found on January 27, 2020, from Sri Lanka, who was a tourist, and none of the natives were found till the starting of March. Then the virus was spread around the country. The epidemic spread more than three months (2020 March to 2020 May) and counting [2]. Government bodies and health-related bodies decided to have a curfew in several areas in the country to control the situation. The primary purpose of this curfew is to control the spread by blocking the movements of the infected clusters. There are mathematical models to forecast these movements[3]. There are different platforms used by people for discussions in Sri Lanka[4]. Currently, social media is one of the leading platforms among information dissemination. Social media in Sri Lanka heavily used to disseminate information similar to “mouth of the word”. Apart from information dissemination, social media engagements and digital discussions happened in different aspects like political, religions, drugs-related news. Some rules and regulations imposed by the authorities to prevent the spread of the virus distracted religious and political aspects of the natives and created unrest during the period. As a result social media was used by users to express their views on the imposed rules and regulations by the authorities. Further, the asymptomatic carriers related discussions happened during the curfew period in social media platforms. Some of the discussions were highly criticized. Such rebuked discussions are moved towards irresponsibility behaviour and blame among several parties to hold the responsibility of the outbreak.

Two prominent social media platforms under study are, Twitter and YouTube, for data extraction. Both platforms allow us to run open-ended discussions. Hence those data were used for the data analysis. Factors that are post-date, posted time, the number of subscribers, followers, are ignoring—mainly considered about the sensitivity of the comments in Sinhala language or related express that are related to hate speech and having sufficient comments to perform data analysis. More in deep data extraction time-limited to April 01, 2020, to May 15, 2020, and considered a minimum 25 comments for the original post. There were 35 posts under this condition satisfied. Hence analysis was conducted based on that data. There were nearly 1500 comments extracted.

Words that performed as hate speech mainly considered in the experiment. In social media platforms, users are performing textual expressions in posts, comments. In such textual representations, Sri Lankans are using different languages. The Sinhala language mainly considered in this research. Further, the Sinhala language words in the English language text also considered. Emojis with text considered. Just an emoji not considered a hate speech. Textual comments with images considered if it contains sufficient text. Only an image ignored. While experimenting, each word is human interacted and read each word before clustering or defining any further activity. Tamil is another national language in Sri Lanka. However, Tamil language text analysis considered as future work.

The speech expressed encourage of violence is a common factor in any aggressive discussion. The violence that is mainly visible during the outbreak period is political, religious, irresponsible behavior, and etheromaniac. Discussions that are related to "hate" is considered for this research. The main purpose is to identify the discussion trends related to hate speech. This hate speech related discussion considered based on the timestamp of the comments in the research to identify a trend. They are mainly performing a time series analysis to determine the trends of such discussions. Based on the commenting time, an extended time series-based analysis carried out to spot trends in an outbreak situation how social media users are express hate-related words among each other. Time is the sensitive factor that social media engagement can spread. Commenters are the user profiles that are engaging in the discussion. Hate speech time-based discussions on social media platforms are spreading information rapidly. It can be highly effective for the entire social media platform that is a consideration for a specific topic.

Hate speech in an outbreak is highly sensitive to the time. Mainly it is started after a few minutes of the original post, and it spread a few days. Then the discussion is over. In the same manner, topics change with COVID 19 related information spreading. Especially after a piece of news spread in any media (mainly television or a digital channels), that is starting another topic to discuss. Some discussions are blamed among each other or in some situations it is making some people, party, or another social cluster sarcastically and radically rejecting. While experimenting, few data points were identified. Those are in very irregular data, and no seasonality explicitly found. By considering initial time stamps all posts having a trend, but no seasonality found. Hence data split into two components and experimented. As per the requirements, algorithms use to define the trend of the commented timestamp sequence. Finally, the final solution was a merging trend by smoothing.

II. RELATED WORK

Sri Lanka impacted different social changes for different reasons. Due to even distribution among the end-users, social media are impacting with numeracy facts in both good and bad sides. In such considerations are eventually flow for a high volume of thoughts that are spread[5]. Currently, Sri Lanka is mainly focusing on block entire social media platforms to block the dissemination of fake and hate-related content [6]. Human is an evolution with different stages, from physical to digital. In the same way, the digital age can change a lot of human thinking and interaction with the entire environment [7]. Social media commentary considered the central aspect of extreme speech. Commentary is an impactful lesson to keep a more narrative approach to keep the information as co interactive with the hate speech. Hence it is a detective with a framework commentary [8]. Hate speech extent to have various social settings. In that case, the reach of communication was interpreting with different social media contexts and other related information retrieval solutions. In those cases, communication was essential to have a substantial frequency. Critical factors related to these hate speech also experimented [9].

Hate speech detection based on Lexicon is another aspect in a social media context. Gitari et al. discussed the three main classifiers for a lexicon-based subjective analysis. Rate the polarity of sentiment expressions also considered in the same discussion [10]—constructive comments corpus with different textual representations discussed. Models developed using deep learning models. Moreover, domain-specific and topic-specific discussions conducted. Quality of discussions is also one aspect that they provided[11]. Studying social networks based on different ethnographic is essential. In this case, the majority of social media platforms are discussing hot topics each day. The intercept of human behavior with digital behavior is a critical study. These can easily describe using visual graphs [12]. The effect of anonymity on hate speech is considerable. Mainack Mondal et al. discussed how to measure hate speech in social media. As per the experiment, sentimental design and structure considered. Further based on 'intensity' and 'hate target' considered [13]. Hate speech is some text that contains the offensive text. There are annotations created for this with labels. William Warner introduced the binary anti-Semitic classifier [14]. This classifier is scope with entire text in the world wide web. Hate speech detection in social media based on classification models discussed. The main target was the identification of speech is either hate or not [15]. Sinhala language social media racist analysis model with machine learning was discussed based on the Azure platform. Detection carried out racist speech conducted in the Sinhala language [16].

Predicting social media posts with factors concerning context also discussed. In this case Subjectivity, polarity considered for sentimental analysis. Further, they have used algorithms like linear regression to have a prediction [17]. A collective sentiment based on statistics experimented. It mainly tested on an individual analysis of individual tweets and the structure of the response. These experiments were successful by having predictions based on sentiments [18]. Extraction data from twitter discussed in different aspects. Data extraction and analysis can discuss the forecasting of social interactions among specific topics as threads of messages. The same way machine learning algorithms highly employed for generating results that modeling the forecasting models [19]. Some occurrences lifetime is comparatively slow. Time-sensitive hourly forecasting is another valuable aspect. Mainly time-sensitive applications discussed than ever [20]. Forecasting models that are relevant to social media interaction discussed. Predicting retweets is one of aspect carried out. The number of retweet counts can impact social interactions. Hence forecasting the number of retweets is a significant fact in social media platforms [21]. Sarcasm is another aspect of social media hate speech. In this case, users have mainly interacted with positive hate speech. To analysis such a case a framework developed to detect Sarcasm [22]. The approach was a neural network-based system.

III. METHODOLOGY

Research Questions

RQ1: Identification of the trend of a hate speech with a short period of a post and forecast the commenting behavior using time series analysis.

A. Data set

Novel predictive analysis is based on social media data. Under popular social media platforms, twitter, and YouTube selected. The main reason behind this is the opens to comment and hashtags provide open-ended access for a specific topic. Data collected from 1-April 2020 to May 15 2020. Mainly consider textual comments in the Sinhala language. The authors were targeting whenever the original post is in the Sinhala language. There are comments in the English language and Sinhala words written in English text too.

Further rarely (<1%) Tamil text also in comments. These texts considered for timestamp as commented time. There are related to only COVID-19 and highly coupled topics during the data collection period. While collecting data if there is a reply to more than a single account including the original post, the authors consider it is a reply to the original post. If not, then consider the reply to the user account who comment on the original post. Those posts considered as a sub comment to the original post. Due to data extraction conditions on twitter, authors were unable to extracted replies that are hidden by originated by tweet author. While deriving data, authors were careful to extract data in Indian standard time as a timestamp. The authors have collected 35 publicly available posts. It was considered a minimum 25 reply comments should available in the post.

B. Data Cleaning process

Users are interacting with a post in different ways. From fully explain comments to an emoji can by applying. The same can express with an image too. These are exciting topics to filter explicit comments that required for any flow of intensive collaborative text analysis [23]. The initial consideration is personal information. For any personal identifiers in the data, the authors removed all identifiable records accordingly. Then consider all timestamps. The authors have conducted noise removal for all Tweets. Removed unnecessary comments that are out of the scope and the hidden comments ignored.

C. Data classification process

Data classified into mainly two categories: the content authors post and reply posts. Once reply posts are available, it is considered a reply to the original author or not and then classified each comment into two categories. Reply to hate speech or not. All the hate speech related comments are classified into two main categories as "agreed hate speech" or "disagree hate speech". The classification was carried out based on crowdsourcing methods. While performing this categorization most relevant category was selected. Moderated comments kept as neutral comments. According to the classification, comments annotated with timestamps.

D. Crowdsourcing based classification

For classification of the posts as a hate/not hate or neutral, authors used crowdsourcing. The classification used opinion on a text in the comment with the original post. All-time considered only this couple of the original post and reply comment. Further, it was analyzed with around 50 university students. Based on their comments, classification conducted for more than 50% acceptance.

IV. Analysis and Results

A. Technology selection

To perform framework development, python considered for data analysis. Further software packages related to python programming languages such as pandas, statsmodels, matplotlib highly used for data analysis and visualization. Data analysis was conducted in various time series analysis methods to forecast trends.

B. Experimenting Seasonality

An initial experiment conducted for finding life span and outliers in the data set. Each post started time by the author was considered as the starting point of the post, and the end time was considered the last commented time to the original post. The following graph displays the period of each post. During the analysis "time" was considered as python naive Date Time object.

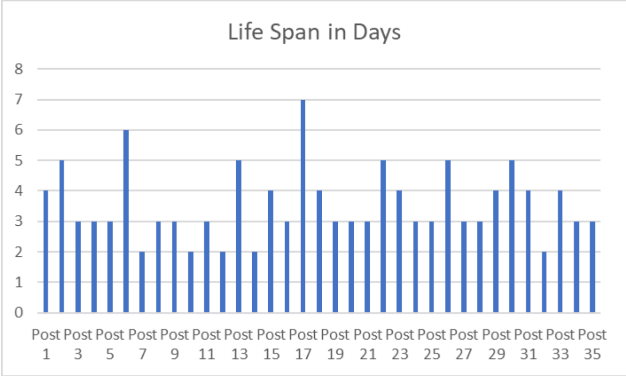


Figure 01 – Life Span of tweets

As per the above graph (Figure 01), it is clear that the time of each post is very short in days. The maximum lifespan is only seven days. So, the trend was considered for a short period. This research is a very time-sensitive project. (i.e. within low timespan users are interactive rapidly) While performing a time series analysis, the minute-based analysis considered. As per the analysis, maximum life span average 3.54 days. Further comments were exploring within the first few hours rapidly and finished within a few days. In this experiment, the authors considered the period regardless of date, time, or day. As per Brendan et al.[24], almost interactive time for a digital post is highly impacting on post-date time. Since this is an outbreak, those conditions are not applicable.

Performing "Summary Statistics" to identify seasonality is a stable method. Hence further analysis performs by using simple statistics. Calculated mean and variance by splitting the data set into two random clusters.

Test	Mean	Variance
Test 1	8.556	2.914
Test 2	3.530	0.720

Table 01 – Summary statistics

Statistical calculations run divided by two main data sets to find out the common statistical values. As per results, the mean in both tests executed in highly independent values are no correlation was found.

Variance considered in both test sets. Generated variance not coupled highly. Hence, these statistical values considered a non-stationary time series.

Subsequent analysis carried out to find out any seasonality, trend, cyclicity, or residuals can find in the data set. As per the below histogram (Figure 02), it has a Gaussian distribution. It represents the seasonality of the data set. Further analysis carried out to find out hourly based (Based on the timestamps) and considered each post separately with the time stamps to find out any seasonality in the data set.

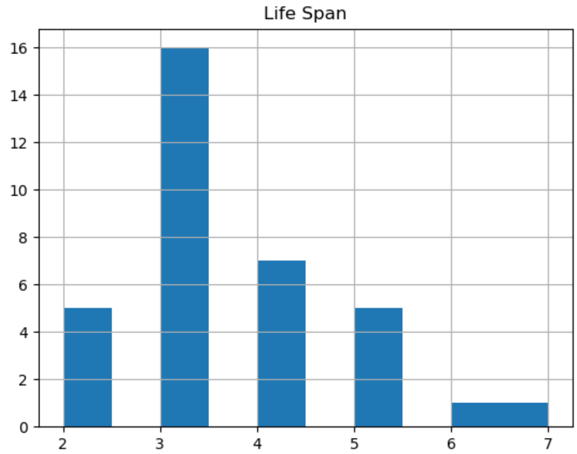


Figure 02- Overall life span of timestamps

Augmented Dickey-Fuller (ADF) Test

Further analysis carried out to find out the Augmented Dickey-Fuller (ADF) test to find out any stationary can represent. Following results (Table 02) generated after performing the ADF test.

ADF Statistic	p-value	Critical Values
6.735665	0.9	1%: -3.639
		5%: -2.951
		10%: -2.614

Table 02 – ADF results

As per the results, the p-value is greater than 0.05. Hence Fail to reject the null hypothesis (H0). Finally, it considered as no seasonality or trend. Further, this data set contains

Figure 03 – Analysis of each post time stamps

few data points with irregular data.

C. Analysis of post-based time stamps

None of the seasonality or trend expressed. As per the below graph (*Figure 03*), none of seasonality or cyclicity. The data pictorial as a random walk.

D. Implementation using exponential smoothing

Selection of algorithms considered with the forecast models that fit for the data. exponential smoothing is the best-fit design for all the above conditions. Hence subsequent analysis carried out to forecast the trend by using exponential smoothing. The authors run the application using python's statsmodels' package. This package can have automatically defined an optimized α value. This method is the recommended approach for implementing the application. The above graph represents smoothing with factor 0.3. This alpha value is the optimum value for time series analysis.

To perform Exponential smoothing data set to split into 70% for training and 30% for testing. After performing the following forecast line generated.

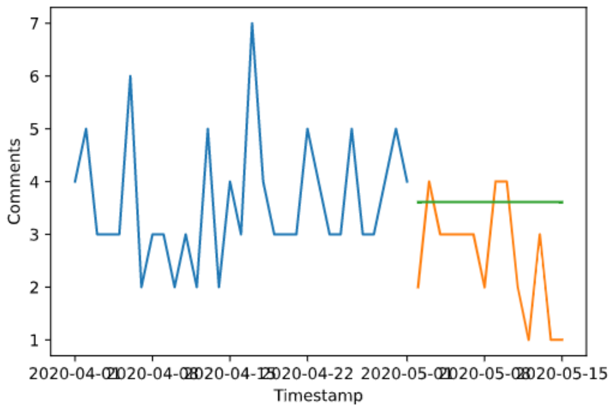
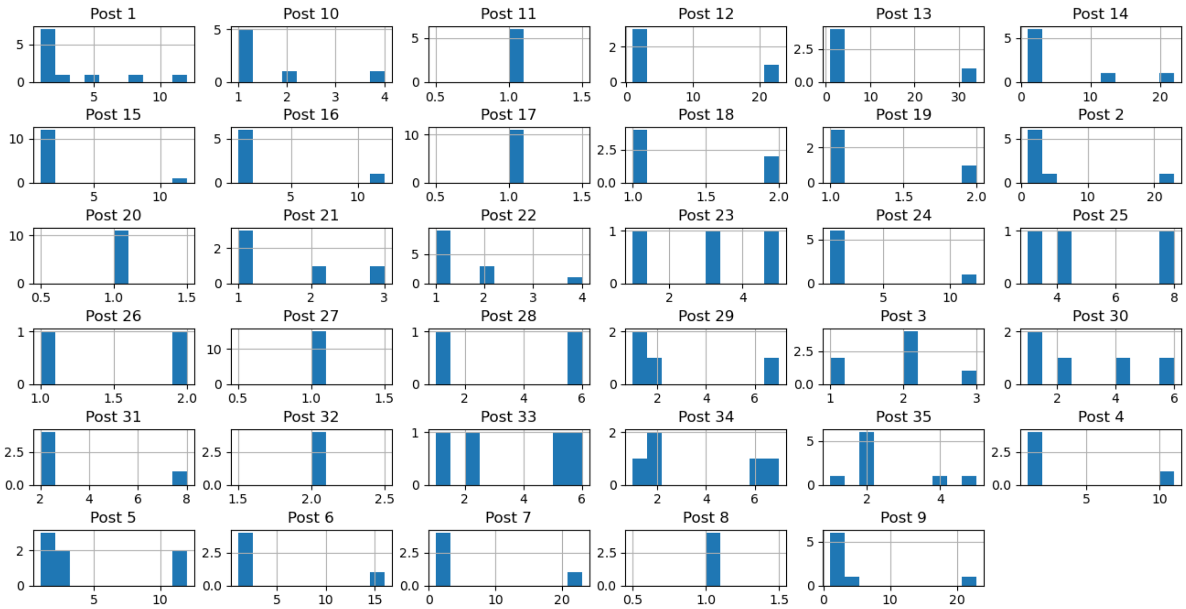


Figure 04 – Exponential smoothing trend line

As per the analysis, it is forecast next comments related to hate speech will downwards trend is available. Training data used training for trend analysis of the textural analysis.



The trend line is horizontal. This trend line mainly represents two points. The first one is the short-term trend for the short life cycle. Plotted graph with the "null" value represents a forecasted range which might be next week. On the other hand trend line under the exponential smoothing provides a sign of highly interactive timely sensitive data in the data set. The trend line is proving in the same way as figure 02 and figure 03. Both are representing a short time and textural representation with the trend forecasting. Further analysis carried out hourly based more sensitive trend line.

E. Implementation using Prophet

Facebook provides an extensive framework called "Prophet" which is an open-source framework[25]. It provides a time series of data analysis—best for an additive model. As per the above data analysis authors are found non-linear trends available in the data set. "Prophet" provides mainly two-time series analysis. The first one is the logistic forecasting model and regression. Authors are experimented both from one day to the maximum trend for nine days trend. It is a more succuss trend line and represented the high granularity of the graph with nearly cyclic effect. The same data set used to predict using Facebook Prophet. Here authors are mainly considered about the liner forecasting model for the next 9 days.

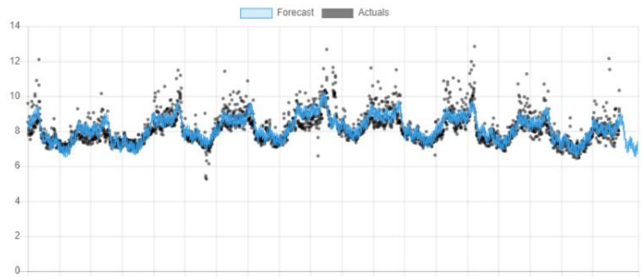


Figure 04 – Hourly based analysis (liner forecasting)

After analysis by using the forecasted hate speech for the next 9 days will get reduced. At the end of the graph, it showed forecast values only. Further Prophet based analysis is more robust than the standard distribution methods due to high sensitivity with the hourly based changes. More analysis carried out for the forecasting based on logistic forecasting model with "+8" as upper saturation points and one as the lower saturation point. "+8" selected because of the maximum period of a post was "+7". The lower point selected as the value of "+1". The key reason is at least one point for the application as a comment.

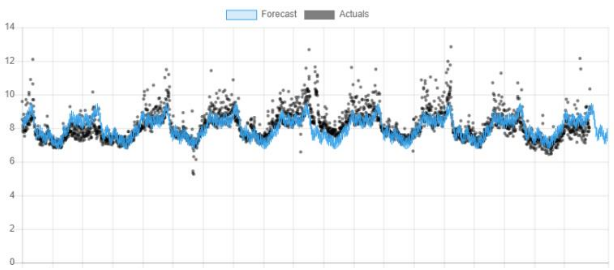


Figure 05 – Hourly based analysis (logistic forecasting)

The trend was a bit upper than the linear forecasting model. However, still, the trend line is expressed to a decreasing in forecasted 9 days.

F. Experimenting with Trend Projection

While performing trend analysis, the authors were interested in performing experimenting with "Trend Projection". Hence data was projected with the graphical method as the initial point for the trend projection. As per the analysis, it was not a highly considerable solution for a trend projection. Further analysis was carried out using the Box-Jenkins Forecasting Method. However, in this case, it was more towards a stable time series with low volatility. Hence a derivation was considered by differencing and moving average.

Figure 06 – Moving Average Analysis

As per the above graph (Figure 06) moving average shows the same line of trend with low demand forecasting. The key reason for the more significant curve in the tail of the graph is that it represents less interactive uses while it is expanding. The authors considered the mapped trend line for the complete forecasting solution. Authors were decomposing time series into two parts that can show as having a trend. Then the time series decomposed a head and a tail. Head is the aggressive comments time stamps. Authors mainly targeted to find out the standard mathematical function after performing this decomposition. Nevertheless, it does not generate a mathematical function for this time-sensitive experiment.

Authors were not interested in "Holt's Linear Smoothing" and “Holt’s Damped Trend.” The reason is that the low period of a post and the trend in constant do not highly useful for the predicted days. In other words, this is a short forecast only considered. Further, from the initial point onwards, there was no trend shown in the data set.

V. Conclusion and Future Work

Modern forecasting trendline for an outbreak is quite essential. Hate speech highly rejected in general society. Minor hate speech and hate speech generators are mainly forcing on different hate speech-related topics that are related to politics, religion, racism. Forecasting trend for a short lifespan commonly noted in the experiment. Though it is a short lifetime, the impact might be higher due to the high velocity of interactive reactors. Commenting some posts is a reaction of a statement with mainly agreed, disagreed, or opinionated text. Hate speech is the text that contains a violation of someone's or some group's natural philosophy. Outbreak related information is mainly expansion the same context, that is, new rules/regulations, or some associated ethics introduced. In this case, different ideas are generating. Such a reaction is disrupting many social media users, and the isolated clusters can also aggressively interact with a social problem. Hence this experiment is quite essential to have a good understanding of the trend of a hate speech in a specific period. As per the analysis, it is clear, hate comments have no specific trend, or seasonally, Hence, the trend line is defined by using Exponential smoothing.

Further forecasting model applies to the prophet framework too. Results are displayed any hate speech during an outbreak is in decreasing interest with time. The trend with time series analysis provided a solid forecast model. The research did not become highly subservient on the natural language processing techniques used for text processing. Instead of used crowdsourcing directly. The usual effects in time series analysis (holiday effects) assumed to have no much effect due to the short period of lifetime of the post. The main aim of the research is experimenting on timestamps, the above conditions assumed as fixed and no effect.

As future work, authors are willing to explore more outbreaks in the country and generate more novel forecasting models. Further analysis of reactions, such as "Likes" to a post is another trend forecasting generated model that can implement in the future. Authors wish to provide a mathematical function as a future work for this application by filling the projection gap of the data and adding more related data.

VI. REFERENCES

[1] G. Bonaccorsi *et al.*, "Economic and social consequences of human mobility restrictions under COVID-19," p. 38.

[2] Gayan Dilantha Illeperuma, "Outbreak of COVID-19, A simple statistical analysis, Focused on Sri Lanka," 2020, doi: 10.13140/RG.2.2.20378.39365.

[3] T. Wickramaarachchi, S. Perera, and S. Jayasinghe, "COVID-19 epidemic in Sri Lanka: A mathematical and computational modeling approach to control," Epidemiology, preprint, Apr. 2020. doi: 10.1101/2020.04.21.20073734.

[4] C. Durugbo, "Analysing communication channels for social networks," in *2011 International Conference on Computational Aspects of Social Networks (CASoN)*, Salamanca, Spain, Oct. 2011, pp. 273–277, doi: 10.1109/CASON.2011.6085957.

[5] S. Hattotuwa, "Digital Blooms: Social Media and Violence in Sri Lanka," no. 28, p. 12.

[6] A. Gunasingham, "Easter Attacks In Sri Lanka," p. 4.

[7] "Human-Computer Interaction and Emerging Technologies: Workshop Proceedings from the INTERACT 2019 Workshops.," in *Human-Computer Interaction and Emerging Technologies: Adjunct Proceedings from the INTERACT 2019 Workshops*, May 2020.

[8] M. Pohjonen, "A Comparative Approach to Social Media Extreme Speech: Online Hate Speech as Media Commentary," p. 16, 2019.

[9] M. Barnidge, B. Kim, L. A. Sherrill, Ž. Luknar, and J. Zhang, "Perceived exposure to and avoidance of hate speech in various communication settings," *Telemat. Inform.*, vol. 44, p. 101263, Nov. 2019, doi: 10.1016/j.tele.2019.101263.

[10] N. D. Gitari, Z. Zuping, H. Damien, and J. Long, "A Lexicon-based Approach for Hate Speech Detection," *Int. J. Multimed. Ubiquitous Eng.*, p. 17, 2015.

[11] V. Kolhatkar, N. Thain, J. Sorensen, L. Dixon, and M. Taboada, "Classifying Constructive Comments," *ArXiv200405476 Cs*, Apr. 2020, Accessed: May 29, 2020. [Online]. Available: http://arxiv.org/abs/2004.05476.

[12] A. Abraham, *Computational Social Networks: Mining and Visualization*. Springer Science & Business Media, 2012.

[13] M. Mondal, L. A. Silva, and F. Benevenuto, "A Measurement Study of Hate Speech in Social Media," in *Proceedings of the 28th ACM Conference on Hypertext and Social Media - HT '17*, Prague, Czech Republic, 2017, pp. 85–94, doi: 10.1145/3078714.3078723.

[14] W. Warner, "Detecting Hate on the World Wide Web," p. 6.

[15] Harvard Medical School, Boston, MA, United States, S. Malmasi, M. Zampieri, and University of Wolverhampton, Wolverhampton, United Kingdom, "Detecting Hate Speech in Social Media," in *RANLP 2017 - Recent Advances in Natural Language Processing Meet Deep Learning*, Nov. 2017, pp. 467–472, doi: 10.26615/978-954-452-049-6_062.

[16] D. S. Dias, M. D. Welikala, and N. G. J. Dias, "Identifying Racist Social Media Comments in Sinhala Language Using Text Analytics Models with Machine Learning," *Int. Conf. Adv. ICT Emerg. Reg. ICTer*, p. 7, 2018.

[17] S. Asur and B. A. Huberman, "Predicting the Future with Social Media," *Appl. Energy*, vol. 112, pp. 1536–1543, Dec. 2013, doi: 10.1016/j.apenergy.2013.03.027.

[18] L. T. Nguyen, P. Wu, W. Chan, W. Peng, and Y. Zhang, "Predicting collective sentiment dynamics from time-series social media," in *Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining - WISDOM '12*, Beijing, China, 2012, pp. 1–8, doi: 10.1145/2346676.2346682.

[19] M. Arias, A. Arratia, and R. Xuriguera, "Forecasting with twitter data," *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 1, pp. 1–24, Dec. 2013, doi: 10.1145/2542182.2542190.

[20] A. Choudhury, "Hourly Forecasting of Emergency Department Arrivals – Time Series Analysis," *SSRN Electron. J.*, 2019, doi: 10.2139/ssrn.3311030.

[21] S. Gransee, R. McAfee, and A. Wilson, "Twitter Retweet Prediction," p. 2.

[22] L. H. Son, A. Kumar, S. R. Sangwan, A. Arora, A. Nayyar, and M. Abdel-Basset, "Sarcasm Detection Using Soft Attention-Based Bidirectional Long Short-

Term Memory Model With Convolution Network," *IEEE Access*, vol. 7, pp. 23319–23328, 2019, doi: 10.1109/ACCESS.2019.2899260.

[23] E. Agichtein, C. Castillo, D. Donato, A. Gionis, and G. Mishne, "Finding high-quality content in social media," in *Proceedings of the international conference on Web search and web data mining - WSDM '08*, Palo Alto, California, USA, 2008, p. 183, doi: 10.1145/1341531.1341557.

[24] B. O'Connor, "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series," p. 8.

[25] "Prophet - Prophet is a procedure for forecasting time series," *Prophet*. https://facebook.github.io/prophet/ (accessed May 15, 2020).