

# Time Series Based Trend Analysis for Hate Speech in Twitter During COVID 19 Pandemic

H.M.M.Caldera<sup>1</sup>, G.S.N.Meedin<sup>2</sup>, Indika Perera<sup>3</sup>

Department of Computer Science and Engineering  
University of Moratuwa  
Sri Lanka

<sup>1</sup>maneesac@cse.mrt.ac.lk

<sup>2</sup>gsnadeerameedin@gmail.com

<sup>3</sup>indika@cse.mrt.ac.lk

**Abstract**—In Sri Lanka social media platforms were used as information propagation medium stupendously during COVID 19 outbreak period. The context of the post creates a high impact on the Social media engagement trends. This research primarily aims at how hate speech was propagated in twitter platform during the COVID 19 outbreak period. This paper describes the results of the study performed on how contents in the Sinhala language, and Sinhala language words written using English text were disseminated. Moreover, and the paper discusses the trend analysis techniques used in performing the trends analysis to identify the hate speech propagation trends over the period. It was observed that there was a rapid interaction with social media platforms during the first half of the posts identified, and in the latter part, the interactions were gradually decreased. Finally, this abstract paper discusses a trend line of hate speech posts identified on twitter data.

**Keywords**— *Trend Analysis, Time Series analysis, Covid-19 outbreak, Hate Speech*

## I. INTRODUCTION

At the time of publishing this paper COVID 19 epidemic is spreading at an exponential rate while impacting public life, including the economy, civil services, community life. The first victim identified in Sri Lanka is a tourist, and the epidemic is been propagated since 2020 March [1]. Researchers have used mathematical models such as SEIR conceptual model and its modified version to control the virus dispersal in Sri Lanka [2]. Twitter users have been using the Twitter platform for open discussions and the primary purpose of this research is to identify the “discussion trends” related to hate speech during the pandemic period. Few approaches taken up globally and locally are briefed in the next section.

## II. RELATED WORK

Sri Lanka performed a ban to the entire social media platforms from restraining the dissemination of fake and hate-related content during the Easter attack in Sri Lanka in April 2019[3] as a preventive action. Studies have been carried out to test the widespread belief that social media platforms have a high level of hate speech in the opinion of survey respondents and to test the idea that encountering perceived hate speech is related to avoiding political talk [4]. Moreover, William Warner [5] has performed pilot classification experiments in which they classify anti-Semitic speech, and the classifier is scope with the entire

text in the world wide web. Simple keyword spotting techniques cannot be used to accurately identify the exact intent of a comment and to address this issue a text analytics model with two-class support vector machine has been used to filter racist comments in Sinhala language and tested with selected Facebook comments. The next section explains the methodology carried out in this research to analyze the trends related to hate.

## III. METHODOLOGY

The research was carried out to identify the trends of hate speech during a short period of Twitter posts and to forecast the commenting behavior using time series analysis. Data collection was performed during the period of 1-April 2020 to May 15, 2020 and the study was carried out only with the Twitter posts used Sinhala language. The posts consists of English language and Sinhala words written using English text ,personal identifiers and outliers were removed from the data set. Next, the data set was classified into two categories: the posts or content by a user and replies. Labelling and the classification of the posts as hate, not hate or neutral was done using a crowdsourcing approach manually comment analysis was performed after distributing the comments among 50 university students and each post was distributed among five students to get the comments labelled Finally, the comments with the label of hate content was used for the experiment.

## IV. ANALYSIS AND RESULTS

Initially the study was conducted on seasonality and the life span to create the experiment’s boundaries , and the maximum life span is seven days as shown in Fig. 1.

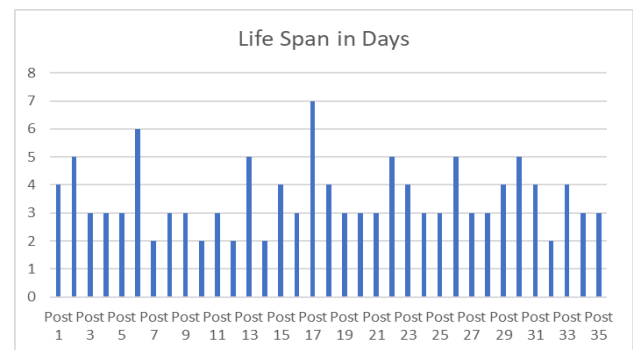


Fig. 1 – Life Span of Tweets

The trend defines for a short period, and research represents a time-sensitive experiment. (i.e., Within the low timespan, users are interactive rapidly). While performing the time series analysis, a minute-based analysis was considered and as per the analysis, the life span average is 3.54 days. A subsequent ‘‘Summary Statistics’’ analysis was performed to identify the seasonality besides a subsequent analysis was conducted to identify any seasonality, trend, cyclicity, or residuals can find in the data set and the results shown in Table I.

TABLE I. SUMMARY STATISTICS

Test	Mean	Variance
Test 1	8.556	2.914
Test 2	3.530	0.720

Furthermore, an Augmented Dickey-Fuller (ADF) test was performed the result represents no seasonality or trend as shown in Table II.

TABLE II. ADF RESULTS

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

Additionally, the experiment was evaluated by an implementation using exponential smoothing, and as per the analysis, it forecasts the next comments related to hate speech with downwards trend identified. The ‘‘Prophet’’ framework provides a time series of data analysis—best for an additive model, logistic forecasting(Fig. 3), and regression forecasting models.

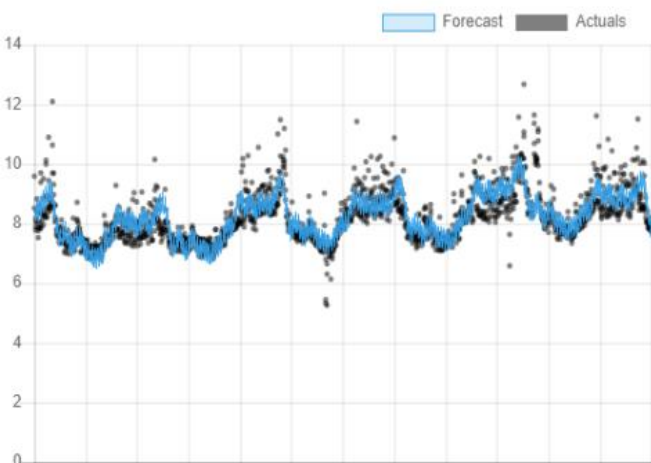


Fig. 2 Hourly based analysis (liner forecasting)

Finally, the analysis carried out using the Box-Jenkins forecasting method, and it was towards a stable time series with low volatility. Hence, a derivation considers by differencing and moving average. The moving average provides a strong relationship with the previous data points and the trend line’s total movement. Hence moving average can produce a strong association of the commenting behavior in tweeter platform.

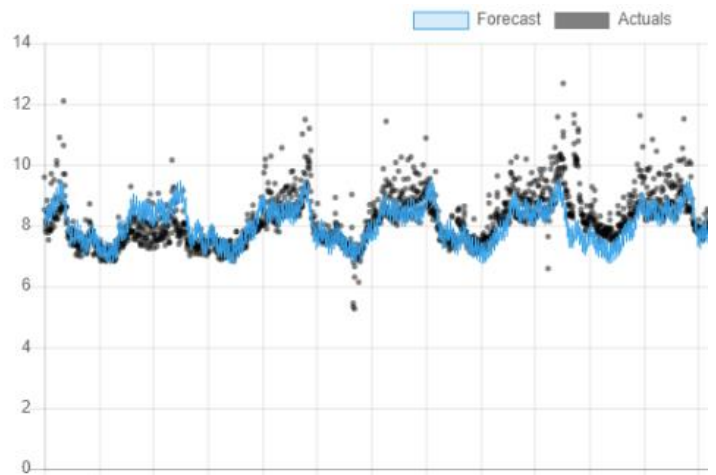


Fig. 3 Hourly based analysis (logistic forecasting)

### V. CONCLUSION AND FUTURE WORK

Through forecasting trend line for an outbreak is viral, and this experiment depicts the nature of the short-term trending trend, and as per the results, the life span is a short lifetime depicts in all cases. Hence, this experiment is essential to understand the trend of a hate speech in a specific period. Moreover, as per the analysis, hate comments that have no specific seasonally. Furthermore, the impact might be higher due to the high velocity of interactive reactors. Any hate speech during an outbreak is in decreasing with time. The trend with time series analysis provided a solid forecast model, and as future work, authors are willing to generate novel forecasting models for short term trend analysis.

### ACKNOWLEDGMENT

This research was supported by the Accelerating Higher Education Expansion and Development (AHEAD) Operation of the Ministry of Higher Education funded by the World Bank.

### REFERENCES

- [1] G. D. Illeperuma, ‘‘Outbreak of COVID-19, A simple statistical analysis, Focused on Sri Lanka,’’ Mar, 2020, doi: 10.13140/RG.2.2.20378.39365.
- [2] 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.
- [3] A. Gunasingham, ‘‘Easter Attacks In Sri Lanka,’’ *Counter Terrorist Trends and Analyses*, vol. 11, no. 6, 2019, pp. 8–13. JSTOR, [Online]. Available: <https://www.jstor.org/stable/26662255>.
- [4] 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.
- [5] W. Warner, J. Hirschberg, ‘‘Detecting hate speech on the world wide web’’, in *Proceedings of the Second Workshop on Language in Social Media*. Montreal, Canada, Jun 2012, p. 19-26, [Online]. Available: <https://www.aclweb.org/anthology/W12-21.pdf>