

## **Analysis of road accident patterns in Sri Lanka using k-means clustering**

Kumari W.M.S.K.\* , Britto K.K.S.N., Kannangara K.K.S.V., Jayethileke H.L., Hansani K.G.P.

*Department of Mathematics, University of Ruhuna, Matara, Sri Lanka*

Road accidents are one of the most highly discussed topics in the world, as their severity is the loss of thousands of lives and considerable property damage. The trends in road accidents are intended to analyze and investigate the root causes of such occurrences and may be useful in mitigating the risk. Hence, this study aims to identify the road accident patterns in different locations across Sri Lanka using the K-means clustering technique with principal component Analysis (PCA). Euclidean distance is used to calculate dissimilarity between data points and the quality measure for the clustering algorithm is compared along with the Dunn index (DI), and average silhouette coefficient(S). The dataset covers the 24 hours for a particular year from January 2018 to December 2022 occurring in 40 police divisions in Sri Lanka. The optimal number of clusters is obtained as three based on the Elbow method and the analysis of clustering indicated that the high-risk areas for road accidents are in Colombo, Nugegoda, Gampaha, Mount Lavinia, Kelaniya, Kandy, Kurunagala, and Rathnapura at nightfall. Finally, the accuracy of the model was evaluated utilizing the correlation coefficient and Root Mean Squared Error (RMSE). The model demonstrated acceptable accuracy with a correlation coefficient closer to one and 0.9240 RMSE. These findings are useful in elaborating to strengthen road safety in high-risk areas at nightfall. Further, this study has the potential to identify the various factors behind road accidents that occur at observed times and durations as future work.

**Keywords:** Dunn index, Elbow Method, K-means Clustering, Road accidents, Silhouette coefficient

**Acknowledgement:** *The authors are thankful to the officers of the Police Headquarters of Sri Lanka for providing data.*

\*Corresponding author: samadhikalpani97@gmail.com