

Proceedings of the 5th International Conference on Electrical, Communication and Computer Engineering (ICECCE), Kuala Lumpur, Malaysia, 30-31 October 2024

Classification of crime-related news titles using machine learning methods

Rangata, Mapitsi R
Council for Scientific and Industrial Research (CSIR)
Meiring Naude Drive, Pretoria, 0184
Email: MRangata@csir.co.za

News classification in the crime-related domain offers a multitude of benefits that extend beyond efficient information retrieval and crime analysis. Classified news articles can be used to educate the public about crime prevention tips, safety measures, and the importance of community involvement in crime reporting. Classifying crime news provides a rich dataset for researchers to study crime trends, patterns, and contributing factors. This evidence-based research can inform the development of effective crime prevention and intervention strategies. This paper proposes classification of crime-related news titles using machine learning models. The data was acquired from live news feeds. Methods such as TF-IDF was applied on the data to generate feature vectors. Machine learning (ML) models were trained and evaluated. ML models such as Decision Tree and Random Forest models exhibit the highest accuracy and F1 scores, suggesting that they are the most effective in predicting crimelated news titles. Both models achieved an accuracy of 99% and F1 scores of 98% and 97%, respectively.