The Role of Social Media (Twitter) in Analysing Home Violence: A Machine Learning Approach

Saleem Adeeba[1], Kuhaneswaran Banujan[2*], B.T.G.S. Kumara[3]

[1] *Department of Computing and Information Systems, Sabaragamuwa University of Sri Lanka,*
sadeeba@std.appsc.sab.ac.lk

[2] *Department of Computing and Information Systems, Sabaragamuwa University of Sri Lanka,*
bhakuha@appsc.sab.ac.lk

[3] *Department of Computing and Information Systems, Sabaragamuwa University of Sri Lanka,*
btgsk2000@gmail.com

Home Violence (HV) has been a persistent issue across the globe, transcending economic status and cultural boundaries. The COVID-19 pandemic has further exacerbated this problem, bringing it to the forefront of public discourse. This study aims to analyse the impact of HV by utilising Twitter data and Machine Learning (ML) techniques, categorising tweets into three groups: (i) HV Incident Tweets, (ii) HV Awareness Tweets, and (iii) HV Shelter Tweets. This categorisation provides several advantages, such as uncovering new or hidden evidence, filling information gaps, and identifying potential suspects. Over 40,000 tweets were collected using the Twitter API between April 2019 and July 2021. Data pre-processing and word embedding were performed to prepare the data for analysis. Initially, tweets were categorised into HV Positive (containing relevant information) and HV Negative (noise or unrelated content) groups. Manually labelled tweets were used for training and testing purposes. Machine learning models, including Support Vector Machines (SVM), Naïve Bayes (NB), Logistic Regression, Decision Tree Classifier, Artificial Neural Networks (ANN), and BERT+LSTM, were employed for this task. Subsequently, HV Positive tweets were classified into the three aforementioned categories. Manually labelled tweets were again used for training and testing. Models such as Tf-IDF+SVM, Tf-IDF+Decision Tree, Tf-IDF+NB, and GloVe+LSTM were utilised. Several evaluation metrics were used to assess the performance of the models. The study's results provide important new understandings of the prevalence, patterns, and causes of HV as they are reported on social media and how the general population reacts to these problems. The research clarifies how social media may help spread knowledge, provide assistance, and link victims to resources. These insights can be instrumental in informing policymakers, non-profit organisations, and researchers as they work to develop targeted interventions and strategies to address HV during and beyond the COVID-19 pandemic.