

**MACHINE LEARNING TECHNIQUE FOR EARLY
DETECTION OF SOCIAL UNREST ACTIVITIES IN SRI
LANKA USING TWITTER DATA**

Panchali Denipitiya

A dissertation submitted in partial fulfilment of the requirement for Master of Science degree in
Business Analytics

School of Computing

**Informatics Institute of Technology, Sri Lanka in collaboration with
University of Westminster, UK**

2023

Abstract

Protests hold a long-standing legacy of precipitating transformative shifts within societies and cultures. Throughout history, these demonstrations have often ignited profound changes in political systems and, in some instances, escalated into full-fledged conflicts and wars. A historical archetype is the series of uprisings against communism in Central and Eastern Europe during 1989, culminating in the fall of the Berlin Wall. More recently, the Arab Spring protests that were initiated on December 14th, 2010, in Tunisia serve as a poignant testament. These protests catalyzed the resignation of the Tunisian president on January 14th, 2011, setting in motion a cascade of events that reverberated across neighboring nations.

In yet another instance, the year 2022 bore witness to significant protests in Sri Lanka, a country grappling with the reverberations of governmental economic policies and debt mismanagement. Widespread protests surged, decrying these factors, particularly the rapid devaluation of the Sri Lankan rupee. This movement culminated with the resignation of the then-president, Gotabaya Rajapaksa, alongside 26 ministers and Prime Minister Mahinda Rajapaksa.

Collectively, protests can dismantle political regimes to shape socio-economic landscapes. But adverse effects of protests include loss of lives, injuries, and property damage. While quantifying these losses monetarily is challenging, their impact on the businesses is profound including disruptions to manufacturing and services. Sri Lanka is a country that is home to various manufacturing industries. In 2021, Sri Lanka exported a total of USD 14.1 Billion worth of goods and services that accumulated to a 16.58% of Sri Lankan GDP. And the main export goods are apparel & textiles, Ceylon tea & Ceylon spices, rubber & latex products, and fish & seafood. Sri Lanka is also a nation that has experienced a notable series of prominent demonstrations in the recent past. With its economy leaning heavily on exports, any protest capable of immobilizing a key manufacturing hub due to widespread protests and resultant curfews could markedly undermine the overall manufacturing productivity. Hence enabling businesses to receive advance alerts about potential protests that could impact manufacturing provides an opportunity for these businesses to implement pre-emptive measures and strategies to mitigate the ensuing disruptions.

While prior research has explored many aspects of protests in a global scale including early detection, a critical void remained in investigating protest and its early detection using social media

within the specific context of Sri Lanka. The research was undertaken to bridge this lack through a comprehensive analysis of social media data that could help Sri Lankan businesses to confront effects that rises from unexpected protests. Hence, this study conducted an in-depth analysis of Twitter data from the 2022 Sri Lankan Protests, integrating keywords, location mentions, and sentiment analysis. In response to the archival data constraints of Twitter, a dataset was procured from a research student. Rigorous preprocessing and filtration led to the selection of 5666 Tweets. Employing text analytics, manual curation identified keywords, and location mentions. Furthermore, the study calculated positive and negative sentiment percentages. These parameters served as inputs for four machine learning algorithms—Logistic Regression, Naive Bayes, XG Boost, Random Forest and the model performances were calculated. Despite the limitations posed by a limited dataset, this study successfully uncovered that certain Twitter keywords and the negative sentiment expressed in Tweets do exert an influence on the incidence of protests in Sri Lanka, to a certain degree. However, it was also acknowledged that relying solely on a single social media platform is impractical and insufficient, as it overlooks the broader spectrum of information dissemination channels such as other social media platforms and other popular methods such as newspapers, and posters which to date are heavily utilized in Sri Lanka.