A Classification Model to Analyse the Spread and Emerging Trends of the Zika Virus in Twitter

*ABSTRACT*

*The recent trend in sharing critical information on social networks such as Twitter has been a motivation for us to propose a classification model that classifies tweets related to Zika and thus enables us to extract helpful insights for the community. In this paper, we try to explain the process of data collection from Twitter, the pre-processing of the data, building a model to fit the data and present some useful predictions and insights that will be helpful in the fight against the Zika virus.*

1. INTRODUCTION

The zika virus, is responsible for causing the Zika disease and is primarily carried by the *Aedes* species mosquito. The incubation period of the disease lasts for at most a week and has symptoms such as fever, rashes, headache and conjunctivitis.

In May 2015, the Pan American Health Organization (PAHO) issued an alert regarding the first confirmed Zika virus infection in Brazil. On February 1, 2016, the World Health Organization (WHO) declared Zika virus a Public Health Emergency of International Concern (PHEIC). Local transmission has been reported in many other countries and territories. Zika virus will likely continue to spread to new areas. However, there is growing evidence of a causal link between Zika virus infection in pregnancy and births of a congenital disorder called microcephaly, where the brain of the developing foetus fails to grow normally and babies are born seriously deformed. There is no vaccine or treatment for the Zika virus and thus is a serious health issue.