



ChemTech

International Journal of ChemTech Research

CODEN (USA): IJCRGG, ISSN: 0974-4290, ISSN(Online):2455-9555
Vol.10 No.5, pp 590-594, 2017

Resolving Class Imbalance in Meteorological Datasets for Predicting Dengue Outbreak

Vivek Jagadeesan Sharavanan, Arun venkatesh Aapakudal
Venkataraman, Muthusaravanan Sivaramakrishnan and Ram Kothandan*

Department of Biotechnology, Kumaraguru College of technology,
Coimbatore, Tamilnadu, India.

Abstract : The vector-borne diseases are highly contagious and responsible for high rate of mortality especially in tropical and subtropical region. The rapid urbanization and change in climatic pattern restricts the elimination of vector-related diseases. The weather has a direct effect on outbreak of vector-borne diseases. In this study, an attempt has been made to predict prior occurrence of dengue infection in North Tamilnadu region by constructing a machine learning based model. The training set utilized in this study was collected manually using text mining approaches from various meteorological sources. The major pitfall during the data preparation is the class imbalance problem that exists with the outbreak of disease in the various monsoon seasons. In this present study, we employed cost-sensitive based approach to overcome the class imbalance in the meteorological dataset. Based on the prediction obtained, it was observed that both feature ranking and cost-sensitive method improved the prediction performance. Prediction in performance obtained in the study was based on F-measure, MCC and ROC value.

Keywords : vector-borne diseases, meteorological data, machine learning, class imbalance, cost sensitive.

Ram Kothandan *et al* /International Journal of ChemTech Research, 2017,10(5): 590-594.
