

Predicting Stroke Occurrence Using Ant Colony Optimization

Nurul Shamimi Shahidan¹, Rizauddin Saian²

¹*nushamimiey422@gmail.com*

²*rizauddin@uitm.edu.my*

ABSTRACT

Stroke is rapidly becoming a major public health issue in Malaysia, where it is the third leading cause of death. It has sparked widespread concern among health professionals. Hence, they must put in more effort to detect stroke disease, which is not an easy task. This study used the Ant Colony Optimization algorithm known as the Ant-Miner to develop a classification model to predict stroke disease. This study made use of a dataset provided by a data scientist at Kaggle where the data is a refined subset of the original dataset which is based on the Electronic Health Record (EHR) controlled by McKinsey and the company. This dataset describes the risk factors for stroke in patients. First, the data set was discretized using WEKA software to convert numeric attributes to nominal attributes. Second, the Ant-Miner algorithm will train the data to produce the classification model. In addition, this research used a k-fold cross-validation procedure to validate the performance of the developed classification model. The results show that the predictive accuracy of the developed classification model is at par with the industry-standard classification algorithm such as J48. Furthermore, it produced a classification model with a fewer number of rules and conditions.

Keywords: Ant Colony Optimization, Classification, Ant-Miner, WEKA