

# Classification of Dengue Serotypes Using Gini-index based Feature Selection and Rule Extraction from Neural Network

 Pandiselvam Pandiyarajan and Kathirvalavakumar Thangairulappan

## Abstract

Machine learning algorithms are used to diagnose the dengue based on the symptoms, climate risk factors, patients' records and gene sequence of the patients. These methods are used to diagnose the dengue in later stages. If the structure of the protein is known then it would be easier for the biologist to classify the serotypes based on the function of the protein. However, it is still costly to know the structure of the protein. Sometimes these methods could not correctly classify the dengue serotypes. To overcome these problems, this paper proposes the stable and low-cost method for classifying dengue serotypes based on amino acids in the protein sequences. The proposed method uses Gini-index and information gain for feature selection and rule extraction from the neural network for classifying dengue serotypes. It also identifies the most significant amino acids for the cause of dengue. Results of the experiments show that the proposed method classifies 96% of the dengue serotypes correctly by simple extracted rules and identify the cause amino acids for the dengue. The result of this paper is useful to the drug designer. The proposed method classifies dengue serotypes easily also to the children as it needs only the protein sequence which can be obtained from nail or hair.

 Volume 11 | 04-Special Issue

 Pages: 1620-1629