

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320-088X



IJCSMC, Vol. 2, Issue. 12, December 2013, pg.229 – 238

SURVEY ARTICLE

SURVEY OF CLASSIFICATION RULE MINING TECHNIQUES FOR IDENTIFYING DISEASE CAUSE AND DIAGNOSIS

K.S.Thirunavukkarasu¹, Dr. S.Sugumaran²

¹Ph.D Research Scholar, Manonmaniam Sundaranar University, Assistant Professor, Department of computer science, Nehru Memorial College, Trichy, TamilNadu, India
²Associate Professor, Department of computer science, Erode Arts and Science College, Erode, TamilNadu, India

¹ thirukst@gmail.com; ² prof_sukumar@yahoo.co.in

Abstract-Classification is a supervised learning technique. Classification arises frequently from bioinformatics such as disease classifications using high throughput data like microarrays. Classification rule mining classifies data in constructing a model based on the training set and the values or class labels in a classifying attribute and uses it in classifying new data. Currently, a various modeling techniques are detailed for data mining. The details of data mining and machine knowledge in related and network domains are dependent and comparatively distributed. The technique particularly achieves the statistical belief among occurrences in order to enhance classification accuracy. An attention on dependencies is made where the ability to draw classification accuracy is affected in improving performance of the model.

Data partitioning approaches such as bagging and boosting are greatly handled in multiple classifier systems to improve classification accuracy. Most current data stream classification techniques fails in one essential aspect of stream data i.e. arrival of a class. So, a data stream classification method that merges a class detection system into traditional classifiers is enabled. The automatic detection of classes before true labels arrive is detected. The problem of data stream classification, where the data appear in an abstractly limitless stream and the chance to analyze each record is briefed. The searching of a training set accurately and classifying is difficult while considering a large data set. Even with the classified data set the accuracy of the classification is inefficient with error rates. This paper presents classification based on shared information for diagnosing disease. The medical dataset is analyzed with stroke disease reducing error rates providing classification accuracy. This paper also reviews certain data mining papers on classification rule for disease diagnosis patterns.

Keywords— Bayesian classifier; Rule mining; Random forest; Classification; Review

Full Text: <http://www.ijcsmc.com/docs/papers/December2013/V2I12201329.pdf>