On the use of Bayesian network classifiers to classify patients with peptic ulcer among upper gastrointestinal bleeding patients

ABSTRACT

A Bayesian network classifier is one type of graphical probabilistic models that is capable of representing relationship between variables in a given domain under study. We consider the naive Bayes, tree augmented naive Bayes (TAN) and boosted augmented naive Bayes (BAN) to classify patients with peptic ulcer disease among upper gastro intestinal bleeding patients. We compare their performance with IBk and C4.5. To identify relevant variables for peptic ulcer disease, we use some methodologies for attributes subset selection. Results show that, blood urea nitrogen, hemoglobin and gastric malignancy are important for classification. BAN achieves the best accuracy of 77.3 and AUC of (0.81) followed by TAN with 72.4 and 0.76 respectively among Bayesian classifiers. While the accuracy of the TAN is improved with attribute selection, the BAN and IBK are better off without attribute selection.

Keyword: Bayesian network classifiers; Classification; Feature selection; Gastro intestinal bleeding; Peptic ulcer disease