Hathaichanok Kornchee 2008: Calculation of Attribute Weights for Naive Bayes Classifier by Using Data Clustering and Decision Tree. Master of Science (Computer Science), Major Field: Computer Science, Department of Computer Science. Thesis Advisor: Associate Professor Anongnart Srivihok, Ph.D. 92 pages.

The objective of this study is to improve Hall's algorithm by applying a local weighting scheme for optimizing each test data and solving problems of attribute weights which equal to zero. To improve the prediction performance, we purpose a data clustering and decision tree algorithms to calculate weights for each attribute and applied to each node of Naïve Bayes classifier. In this study two clustering methods and decision tree were used to calculate weights of Naïve Bayes. The clustering methods include (1) K-Means algorithm and (2) two step clustering by Cluster Feature Tree and Agglomerative Clustering. Then the performance of the two proposed methods were compared with Hall's method. The measures included percent of correct data prediction and Root Relative Square Error. Four Data sets used in the experiment were obtained from the University of California, Irvin (UCI) machine learning repository. They included (1) Chess End-Game (KR) (2) Balance Scale Weight and Distance Database (Bal_sc) (3) Wave Form and (4) German Credit dataset (Crd).

Result showed that data clustering and decision tree algorithms outperformed Hall's algorithm upon 3 parameters which included number of data (large), number of attribute (large) and number of class (small). However the time complexity of this approach is more than Hall's because it uses more algorithms in calculating attribute weights, clustering algorithm.

_____ _____ ___ / ___ / ___

Student's signature   Thesis Advisor's signature