

4037540 SCCS/M : MAJOR : COMPUTER SCIENCE ; M.Sc. (COMPUTER SCIENCE)

KEY WORDS : KNOWLEDGE DISCOVERY IN DATABASE / DISCRETIZATION / KOHONEN SELF-ORGANIZING MAP

SURAJIT NUDDAMONGKOL : KNOWLEDGE DISCOVERY IN DATABASE BASED ON KOHONEN SELF-ORGANIZING MAP ALGORITHM.

THESIS ADVISORS : SUPACHAI TANGWONGSAN Ph.D., JARERNSRI L. MITRANONT Ph.D. 85 p. ISBN 974-662-915-8

Knowledge Discovery in Database (KDD) is a new methodology for searching and extracting potential knowledge from database. Data mining is a key element of KDD process to seek knowledge. However, some of the data mining algorithms are unable to handle numerical value. This problem recently can be solved by adding a process to reduce details of data before data mining, so-called discretization. Discretization is data transformation so that it has less detail by dividing the data value into a set of sub-range in order to decrease processing time of data mining step. However, conventional methods are not quite efficient to manage the data with numerical value because they consider only the value, not others, resulting in inefficient knowledge produced.

This research presents a new approach of discretization on numerical data by changing from value-based to content-based approach. Content-based discretization is based on data content and conforms to Kohonen Self-Organizing Map (KSOM).

KSOM is applied for solving the above problem. It is able to learn structure and relationship of data by self learning on sole input data and to perform data discretization afterward. From research and test, we observe that the result is more accurate and reliable and also takes less processing time than other methods. This research can be further applied for KDD, which gives more meaningful and accurate knowledge.