

ABSTRAK

Nama : Adolf Pandapotan
Program Studi : Ilmu Komputer
Judul : Implementasi Algoritma *Clustering* (*Cobweb* dan *Iterate*)
Sebagai Bagian Dari *Data Mining Algorithms Collection*

Tujuan Tugas Akhir ini adalah mengimplementasikan algoritma *clustering* (sebagai bagian dari *Data Mining Algorithms Collection*) menggunakan bahasa pemrograman C++. Ada 2 algoritma *clustering* yang diimplementasikan yaitu *Cobweb* dan *Iterate*. Uji coba dilakukan dengan membandingkan kecepatan eksekusi dari implementasi *Cobweb* dengan *Cobweb* pada WEKA dan implementasi *Iterate*, serta membandingkan kualitas partisi implementasi *Cobweb* dengan *Cobweb* pada WEKA dan implementasi *Iterate*. Ada 2 jenis data uji coba yaitu *dataset* kecil dan *dataset* besar. Hasil uji coba menunjukkan algoritma *Cobweb* pada WEKA bukan algoritma *Cobweb* murni, waktu eksekusi *Cobweb* implementasi lebih cepat dari WEKA namun lebih lambat dari *Iterate* implementasi, urutan data berpengaruh terhadap hasil *Cobweb*, dan kualitas *Iterate* lebih baik dari *Cobweb*.

Kata kunci: *clustering, Cobweb, data mining, dataset, Iterate.*

ABSTRACT

Name : Adolf Pandapotan
Study Program: Computer Science
Title : *The Implementation of Clustering Algorithms (Cobweb and Iterate) as a part of Data Mining Algorithms Collection*

The purpose of this mini thesis is to implement clustering algorithms (as part of Data Mining Algorithms Collection) using C++. There are two clustering algorithms that are implemented, that are Cobweb and Iterate. The experiment is done by comparing the execution speed of Cobweb implementation with Cobweb in WEKA and Iterate implementation, also comparing the partition quality of Cobweb implementation with Cobweb in WEKA and Iterate implementation. There are two kinds of experiment data, which are small dataset and large dataset. The test results show that Cobweb algorithm in WEKA is not pure Cobweb algorithm, the execution time of Cobweb implementation is faster than WEKA but slower than Iterate implementation, the data sorted affected to the Cobweb result and the quality of Iterate is better than Cobweb.

Key words: clustering, Cobweb, data mining, Iterate.

