INVESTIGATING ON HOW BETTER DATA QUALITY AFFECTING THE HIERARCHICAL CLUSTERING PROCESS TIME

By

Michael

A Bachelor's Thesis
Submitted to the Faculty of

INFORMATION TECHNOLOGY

in partial fulfillment of the
requirements for the Degree of

BACHELOR OF SCIENCES
WITH A MAJOR IN INFORMATION TECHNOLOGY

SWISS GERMAN UNIVERSITY
Campus German Centre
Bumi Serpong Damai – 15321
Island of Java, Indonesia
www.sgu.ac.id

July 2010

Revision after the Thesis Defense on 29 July 2010

SGU®
SWISS GERMAN UNIVERSITY

STATEMENT BY THE AUTHOR

I hereby declare that this submission is my own work and to the best of my knowledge, it contains no material previously published or written by another person, not material which to a substantial extent has been accepted for the award of many other degree or diploma at any educational institution, except where due acknowledgement is made in the thesis.

SGU ® SWISS GERMAN UNIVERSITY

______________________________ ________________

Michael Date

Approved by:

______________________________ ________________

Ir. Basuki Setio Msc Date

______________________________ ________________

Chairman of the Examination Steering Committee Date

ABSTRACT

INVESTIGATING ON HOW BETTER DATA QUALITY AFFECTING THE HIERARCHICAL CLUSTERING PROCESS TIME

By

Michael

SWISS GERMAN UNIVERSITY
Bumi Serpong Damai

Ir. Basuki Setio Msc

SWISS GERMAN UNIVERSITY

Hierarchical Clustering is one of many clustering methods used to exploring the relationship in statistical data. It cluster data based on distance, similarities and correlation. The result will be shown as Dendogram.

When developing a hierarchical clustering application, most of the problems occur on how you manage the data and how you process it so it will not use too many resources in your computer in order to reduce the running time.

The application will be developed using C# with .Net 3.5 frameworks in Visual Studio 2008 as our IDE (Integrated Development Environment) and Windows Vista as the operating system.

Keywords – Hierarchical Clustering, Data, Dendogram, Running time

DEDICATION

I dedicate this thesis to both of my parent, Hendra Satyo and Dian Budiman.

SGU®
SWISS GERMAN UNIVERSITY

ACKNOWLEDGMENTS

First of all, I want to give the utmost thanks to God for His help and blessing from the start until I finish writing this thesis.

I would also like to give special thanks to Mr. Basuki Setio for he has helped and support me in developed this application from a scratch with his advice and guidance to make this application run well.

In addition, I also give thanks to Mr. Anto Satriyo Nugroho for helping me creating clear understanding about how hierarchical clustering works and how we implemented it.

Finally, I would like to thank my friends who always support me and helping me during my Thesis work.

SGU®
SWISS GERMAN UNIVERSITY

TABLE OF CONTENTS

SWISS GERMAN UNIVERSITY

SWISS GERMAN UNIVERSITY

SWISS GERMAN UNIVERSITY