

TABLE OF CONTENT

PERNYATAAN	ii
PRAKATA	iii
ABSTRACT	v
TABLE OF CONTENT	vi
LIST OF FIGURE	viii
LIST OF TABLE	ix
CHAPTER I : INTRODUCTION	1
1.1 Background	1
1.2 Problem Statement	3
1.3 Originality	3
1.4 Objectives	6
1.5 Significance	6
CHAPTER II : LITERATURE REVIEW AND THEORY	7
2.1 Literature Review	7
2.2 Theory	9
2.2.1 Clustering	9
2.2.1.1 Definition	11
2.2.1.2 Records and Attributes	13
2.2.1.3 Distance and Similarities	13
2.2.1.4 Clusters, Centers, and Modes	15
2.2.2 Carrot ² Framework	15
2.2.2.1 Carrot ² Architecture	17
2.2.3 Vector Space Model	19
2.2.3.1 Term Weighting	19
2.2.3.2 Query Matching	20

2.2.3.3	An Example	21
2.2.4	Latent Semantic Indexing	23
2.2.4.1	Singular Value Decomposition	23
2.2.4.2	Query Matching	24
2.2.5	Suffix Array	25
2.2.5.1	Data Structure	25
2.2.6	Lingo	26
2.2.6.1	Preprocessing	28
2.2.6.2	Features Extraction	29
2.2.6.3	Cluster Label Induction	30
2.2.6.4	Cluster Content Discovery	33
2.2.6.5	Final Cluster Formation	34
2.3	Research Question	35
CHAPTER III : RESEARCH METHOD		36
3.1	Analysis	36
3.1.1	System Architecture	36
3.1.2	System Requirement Analysis	37
3.2	Design	39
3.2.1	Actor Identification	39
3.2.2	Use Case Diagram	39
3.2.3	Activity Diagram	41
3.2.4	Sequence Diagram	43
3.2.5	Class Diagram	46
CHAPTER IV : RESULTS AND DISCUSSION		49
4.1	Result	49
4.1.1	Managing Tweets	50
4.1.2	Perform Tweets Analysis	54
4.2	Discussion	62
4.2.1	Analysis of Results	62
4.2.2	Accuracy Testing	68



UNIVERSITAS
GADJAH MADA

Sistem Informasi Manajemen Rumah Sakit: Sebuah Studi untuk Dikembangkan di Libya
MOHAMED ABDELMAJID, MOAMAR, Selo, S.T., M.T., M. Sc., Ph.D., Adhistya Erna Permanasari, S.T., M.T., Ph.D.
Universitas Gadjah Mada, 2015 | Diunduh dari <http://etd.repository.ugm.ac.id/>

CHAPTER V : CONCLUTION AND SUGGEST	71
5.1 Conclusion	71
5.2 Suggest	72
REFERENCES	73
APPENDIX	75

LIST OF FIGURE

Figure 2.1	Data Mining Tasks	12
Figure 2.2	Three well-separated center-based clusters in a two-dimensional space	15
Figure 2.3	Two chained clusters in a two-dimensional space	15
Figure 2.4	A special award in the research tools category, recently awarded to Carrot ² by EKMA	17
Figure 2.5	Carrot ² component data flow	17
Figure 2.6	Vector Space Model example – input data	22
Figure 2.7	Vector Space Model example – query matching	22
Figure 2.8	Singular Value Decomposition result matrices	23
Figure 2.9	k-rank matrix approximation	24
Figure 2.10	An example suffix array	25
Figure 2.11	Label induction example – input data	28
Figure 2.12	Label induction example – Singular Value Decomposition result	32
Figure 2.13	Label induction example – the P matrix	32
Figure 2.14	Label induction example – the M matrix	33
Figure 2.15	Cluster content discovery example	34
Figure 3.1	System Architecture	37
Figure 3.2	Use case diagram	40
Figure 3.3	Activity diagram of login	42
Figure 3.4	Activity diagram of processing an analyze	42
Figure 3.5	Activity diagram of accessing tweet from twitter	43
Figure 3.6	Activity diagram of loading tweet from index	43
Figure 3.7	Sequence diagram of accessing tweets from index	44
Figure 3.8	Sequence diagrams of access tweets from index	45
Figure 3.9	Sequence diagram of process an analyze	46

Figure 3.10	Class diagram	47
Figure 4.1	LibyanTweet project's structure	49
Figure 4.2	Tweet.java	50
Figure 4.3	FilterQuery to collect tweets from twitter users in Libya	52
Figure 4.4	Method writeIndex()	52
Figure 4.5	Index	53
Figure 4.6	List of entire tweets stored in index	53
Figure 4.7	GUI's display of tweetAnalyst.java	55
Figure 4.8	Method searchIndex()	56
Figure 4.9	Code to convert search result to be documents	57
Figure 4.10	Clustering of documents	58
Figure 4.11	Method topicalDisplay()	59
Figure 4.12	How tweetAnalyst.java call statisticGraphic.java	59
Figure 4.13	Method createDataset()	60
Figure 4.14	Pie Chart	60
Figure 4.15	Feature extraction phrase	61
Figure 4.16	Cluster scoring	62
Figure 4.17	TreeView of Topical Cluster with query "قبليّة اشتباكات"	63
Figure 4.18	Topic's detail	63
Figure 4.19	Pie chart for query "قبليّة اشتباكات"	65
Figure 4.20	TreeView of Topical Cluster with query "الأضحى عيد"	65
Figure 4.21	Pie chart for query "الأضحى عيد"	66
Figure 4.22	Example of irrelevant topic	68
Figure 4.23	Detail of topic "RayanBanoun"	68