

**Universiti Teknologi MARA**

**Sentiment Mining Using Immune Network  
Algorithm**

**Raja Muhammad Hafiz Bin Raja Kamarudin**

Thesis submitted in fulfillment of the requirements for  
**Bachelor of Computer Science (Hons)**  
**Faculty of Computer & Mathematical Sciences**

July 2012

## DECLARATION

I hereby declare that the work in this thesis is my own except for the ideas or quotation are from the work of other people and published. Otherwise all the ideas or quotation are fully acknowledged in accordance with the standard referring practices of the discipline.

JULY 2012

  
RAJA MUHAMMAD HAFIZ RAJA KAMARUDIN

2010240622

## ABSTRACT

With the abundant of knowledge accessible via Internet and the likes, many have ventured into using that knowledge in order to understand deeply on how the available knowledge affected the whole world directly or indirectly. Sentiment analysis or opinion mining is an application that tries to identify and extract subjective information from source materials. Many business applied sentiment analysis in their marketing and others practical usage. In short, sentiment analysis has become an important aspect in our today's life with or without we realizing it. Currently, there are many techniques available in order to harvest necessary sentiment values. However, there are very few researches have been done with the intention to utilize biological-inspired computing in sentiment analysis. This research tries to utilize biological-inspired computing algorithm, Immune Network algorithm, in this sentiment analysis field. This research also studies the effect of preprocessing on the data which has shown an unexpected result, the data with no-preprocessing done produces more accurate result than the data with preprocessing albeit very small in percentage,  $< 2\%$ . The results obtained by utilizing Immune Network in sentiment mining are not very impressive compared to other Machine Learning algorithms. The algorithms can be improved furthermore by including other disciplines such as Natural Language Processing (NLP) or others. However, this is a step stone towards developing a biological-inspired Sentiment Mining algorithm.

## TABLE OF CONTENTS

CONTENTS	PAGE
<b>DECLARATION</b>	ii
<b>ACKNOWLEDGEMENT</b>	iii
<b>ABSTRACT</b>	iv
<b>APPROVAL</b>	v
<b>TABLE OF CONTENTS</b>	vi
<b>LIST OF FIGURES</b>	ix
<b>LIST OF TABLES</b>	xi
<b>CHAPTER ONE: INTRODUCTION</b>	1
1.0 Introduction	1
1.1 Research Background	1
1.2 Problem statements	2
1.3 Project Objectives	2
1.4 Project Scope	3
1.5 Project Significant	3
1.6 Summary	4
<b>CHAPTER TWO: LITERATURE REVIEW</b>	5
2.0 Introduction	5
2.1 Text Mining and Sentiment Analysis	6
2.1.1 Text Mining	6

2.1.2	Sentiment Analysis	6
2.1.3	Stages in Sentiment Analysis	7
2.1.4	Existing Work in Sentiment Analysis	9
2.2	Data Preprocessing	13
2.2.1	Stop Word Removal	13
2.2.2	Stemming	15
2.3	Artificial Immune System	25
2.3.1	Theoretical Background	25
2.3.2	Incremental Immune Inspired Algorithm	28
2.4	Summary	29
 <b>CHAPTER THREE : METHODOLOGY</b>		 30
3.0	Introduction	30
3.1	Background Study	32
3.1.1	Preliminary Study	32
3.1.2	Knowledge Acquisition and comprehension	33
3.2	Data Preparation	33
3.2.1	Data Collection	34
3.2.2	Data Set Management	34
3.3	Prototype Design	35
3.3.1	Engine Design	36
3.3.2	System Design	57
3.4	Prototype Development	57
3.4.1	Develop Prototype	57
3.4.2	Software and Hardware Requirement	60
3.5	Documentation	61
3.6	Summary	61