



CONTENTS

APPROVAL PAGE	i
STATEMENT	ii
FOREWORD	iii
CONTENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES	ix
ABSTRACT	x
I. INTRODUCTION	1
1.1 Research Background.....	1
1.2 Research Problem.....	3
1.3 Research Scope	4
1.4 Research Objective.....	4
1.5 Research Advantage.....	4
1.6 Research Methodology.....	4
1.7 Thesis Outline	5
II. LITERATURE REVIEW	7
III. BASIC THEORY	14
3.1 Sentiment Analysis.....	14
3.2 Aspect-Based Sentiment Analysis	15
3.3 Natural Language Processing.....	16
3.4 Tokenization.....	16
3.5 Stemming	17
3.6 Stopwords Removal	17
3.7 Bag-of-Words.....	18
3.8 Bag-of-n-Grams	18
3.9 Feature Selection.....	19
3.10 Part of Speech Tagging	19
3.11 Natural Language Toolkit (NLTK).....	20
3.12 Supervised Learning.....	22



3.13	Conditional Random Field	23
3.14	Maximum Entropy Classifier	24
3.15	Performance Measure.....	25
IV.	ANALYSIS AND SYSTEM DESIGN.....	28
4.1	Research Description.....	28
4.2	Material	29
4.3	Research Stages.....	29
4.3.1	Data Collection.....	29
4.3.2	Data Labelling	31
4.3.3	Pre-processing	35
4.3.4	System Design.....	37
4.3.5	Aspect Extraction	37
4.3.6	Aspect Categorization	42
4.3.7	Sentiment Classification.....	47
4.3.8	System Implementation.....	49
4.3.9	Evaluation Design	50
V.	IMPLEMENTATION.....	53
5.1	Main Program Flow	53
5.2	Specification.....	53
5.3	Data Collection and Labeling.....	53
5.4	Preprocessing	54
5.4.1	Case Folding.....	55
5.4.2	Remove Unused Words.....	55
5.4.3	Stop Word Removal	56
5.4.3	Lemmatization.....	57
5.4.4	Stemming.....	57
5.5	Aspect Term Extraction	58
5.6	Aspect Classification.....	61
5.7	Sentiment Analysis.....	63
5.8	Evaluation.....	65
VI.	RESULTS AND DISCUSSIONS.....	67



6.1	Dataset.....	67
6.2	Pre-processing Result.....	67
6.4	Aspect Term Extraction Result.....	68
6.4	Aspect Classification Result.....	70
6.5	Sentiment Analysis Result.....	73
6.6	System Result.....	77
VII.	CONCLUSIONS AND SUGGESTION.....	79
7.1	Conclusion.....	79
7.2	Future Works.....	79
REFERENCE	81