



## TABLE OF CONTENTS

APPROVAL PAGE	iii
STATEMENT PAGE	iv
DEDICATION PAGE	v
FOREWORD	vi
TABLE OF CONTENTS	vii
LIST OF FIGURES	x
TABLES LIST	xiii
ABSTRACT	xiv
CHAPTER 1 INTRODUCTION	1
1.1 Research Problem	3
1.2 Research Scope	3
1.3 Research Objective	3
1.4 Research Advantage	4
CHAPTER 2 LITERATURE STUDIES	5
2.1 Comparison of Research Methods	10
CHAPTER 3 THEORETICAL BASIS	13
3.1 Psychological Theory	13
3.1.1 Briggs-Myers Personality Indicator	13
3.1.2 Keirsey Temperament Sorter II	14
3.2 Data Mining and Classification	15
3.3 Tokenization	17
3.4 Stemming	17
3.5 Feature Extraction	18
3.6 Bag of Words	18
3.7 Word2Vec	18
3.8 Doc2Vec	19
3.9 Term Frequency – Inverse Document Frequency	19
3.10 WordCloud	20
3.11 Random Forest Classifier	20



CHAPTER 4 SYSTEM ANALYSIS AND DESIGN	22
4.1 System Analysis	22
4.1.1 Requirement Analysis	22
4.1.2 Data Acquisition	23
4.1.3 General Description	23
4.1.4 Data Analysis	24
4.2 System Design	25
4.3 Test Subject's Credentials Gathering and Briggs-Myers Test	27
4.4 Briggs-Myers Test Model	28
4.5 Briggs-Myers Test Result Calculation	29
4.6 Twitter Mining and Classification Model	31
4.7 Twitter Mining	32
4.8 Twitter Data Analysis	34
4.8.1 Data Labeling	35
4.8.2 Data Pre Processing, Feature Extraction and WordCloud Generation	36
4.8.3 Twitter Data Classification	44
4.9 Tweet Category Percentage	46
4.10 Data Combination	46
CHAPTER 5 SYSTEM IMPLEMENTATION	48
5.1 Implementation Environment	48
5.2 Subject Permission and Twitter Username Gathering	48
5.3 Briggs-Myers Testing Implementation	50
5.4 Briggs-Myers Test Processing	57
5.4.1 Data Extraction and Representation	57
5.4.2 Briggs-Myers Data Calculation Process	64
5.5 Twitter Mining Implementation	68
5.6 Tweet Labeling Implementation	69
5.7 Tweet Percentage Calculation Process Implementation	70
5.8 Tweet Data Analysis Process Implementation	72
5.8.1 Data Pre Processing Implementation	72



5.8.2 WordCloud Creation Implementation	76
5.8.3 Feature Extraction Implementation	76
5.8.4 Classification, Training and Testing Process Implementation	82
<b>CHAPTER 6 RESULTS AND ANALYSIS</b>	<b>86</b>
6.1 System and Result Testing	86
6.2 Briggs-Myers Test Result	86
6.3 Briggs-Myers Test Analysis Result	88
6.4 Twitter Mining Result	91
6.5 Twitter Data Labeling Result	93
6.6 Twitter Data Analysis Result	94
6.7 WordCloud Result	95
6.8 Twitter Percentage Data Result	97
6.9 Final Data Combination Result	98
<b>CHAPTER 7 CONCLUSSION AND SUGGESTIONS</b>	<b>101</b>
7.1 Conclusion	101
7.2 Suggestion	102
<b>REFERENCE LIST</b>	<b>103</b>
<b>APPENDIX</b>	<b>107</b>