

TABLE OF CONTENTS

TABLE OF CONTENTS.....	v
LIST OF FIGURES	viii
LIST OF TABLES.....	ix
LIST OF APPENDICES.....	xi
ABSTRACT.....	xii
I INTRODUCTION	1
1.1 Research Background.....	1
1.2 Research Questions	4
1.3 Research Problems	5
1.4 Research Scope.....	5
1.5 Research Objectives	6
1.6 Research Advantages.....	6
1.7 Research Approach.....	6
II LITERATURE REVIEW.....	8
III THEORETICAL FRAMEWORK.....	21
3.1 Emotion and its Models.....	21
3.2 Emotion Classification and Sentiment Analysis	23
3.3 Twitter Posts and Its Attributes.....	24
3.4 Twint.....	25
3.5 Pre-processing	26
3.5.1 Usernames, URLs and RTs Removal	26
3.5.2 Numbers and Special Characters Removal	27
3.5.3 Single and Multiple Characters Removal	28
3.5.4 Whitespaces Removal	29
3.5.5 Lowercase Conversion	29
3.6 Natural Language Processing.....	30
3.6.1 Tokenization	30
3.6.2 Lemmatization	30
3.6.3 Stop-words Removal	31
3.7 Feature Extraction	31
3.8 Machine Learning Classifiers.....	32

3.8.1 Naive Bayes Classifier	33
3.8.2 Support Vector Machines	37
3.9 Cross Validation.....	39
3.10 Evaluation and Performance Measures	39
3.11 Feedback Analysis	41
IV ANALYSIS AND DESIGN.....	43
4.1 Research Description.....	43
4.2 Research Phases	43
4.3 Data Acquisition Design.....	47
4.4 Pre-processing	49
4.5 Feature Extraction Design	49
4.6 Classifiers Implementation Design	50
4.7 Cross Validation Design	52
4.8 Evaluation Design	52
4.9 Feedback Analysis Design.....	53
V IMPLEMENTATION	54
5.1 Tools and Libraries	54
5.2 Data Acquisition	55
5.3 Pre-processing	56
5.4 Feature Extraction	57
5.5 Implementation of Naive Bayes Classifiers	58
5.6 Implementation of Support Vector Machines.....	60
5.7 Cross Validation.....	62
5.8 Evaluation.....	63
5.9 Feedback Analysis	63
VI RESULTS AND DISCUSSION	64
6.1 Data Acquisition Results	64
6.2 Pre-processing and Feature Extraction Results	66
6.3 Evaluation of Naive Bayes Classifiers with Cross Validation.....	66
6.4 Evaluation of Support Vector Machines with Cross-Validation.....	68
6.5 Products' Feedback Analysis	70

VII CONCLUSION.....	75
7.1 Conclusions	75
7.2 Recommendations for Future Researches	75
REFERENCES.....	76
APPENDICES.....	77

LIST OF FIGURES

Figure 2.1 Customer emotions detected in contact centre calls	19
Figure 3.1 Implemented emotion models	23
Figure 3.2 An example of a tweet.....	24
Figure 3.3 K-fold cross validation	39
Figure 4.1 Overall research phases	46
Figure 4.2 Dataset acquisition design	47
Figure 4.3 Pre-processing design.....	49
Figure 4.4 Implementation design	51
Figure 4.5 Pseudocode of OnevsRest	52
Figure 5.1 Snippet of scraping with Twin	55
Figure 5.2 Snippet of pre-processing.....	56
Figure 5.3 Snippet of lemmatization	57
Figure 5.4 Snippet of stop-words removal	57
Figure 5.5 Snippet of feature extraction	58
Figure 5.6 Implementation of label encoding.....	59
Figure 5.7 Implementation of Naive Bayes Classifiers	59
Figure 5.8 Implementation of Support Vector Machines.....	61
Figure 5.9 Implementation of evaluation by Cross Validation	62
Figure 6.1 Calculation results of iPhone 11 dataset (left) and Amazon Alexa dataset (right)	71

LIST OF TABLES

Table 2.1 Models of emotion	9
Table 2.2 Annotated datasets	10
Table 2.3 Related works	14
Table 3.1 Attributes of a tweet	25
Table 3.2 Username removal	26
Table 3.3 URL removal	27
Table 3.4 Hashtag symbol removal	27
Table 3.5 Numbers and special characters removal.....	28
Table 3.6 Single characters removal.....	28
Table 3.7 Multiple characters removal	29
Table 3.8 Lowercase conversion.....	29
Table 3.9 Natural Language Processing	31
Table 3.10 Sample dataset of Naive Bayes representation	33
Table 4.1 An example dataset for products' feedback.....	53
Table 5.1 Tweets from Twint with before and after unnecessary attributes removed ...	55
Table 5.2 Mapping emotion category for iPhone 11 Camera dataset.....	58
Table 6.1 iPhone 11 Camera dataset description	64
Table 6.2 Amazon Alexa dataset description	64
Table 6.4 Word occurrences for Amazon Alexa dataset	65
Table 6.5 Some annotated tweets from the dataset.....	65
Table 6.6 Processing results.....	66
Table 6.7 Results of Naive Bayes Classifiers with Cross Validation for iPhone 11 Camera dataset.....	67
Table 6.8 Results of Naive Bayes Classifiers with Cross Validation for Amazon Alexa dataset	67
Table 6.9 Results of Support Vector Machines with Cross Validation for iPhone 11 Camera dataset (rounded to two decimal points)	68
Table 6.10 Results of Support Vector Machines with Cross Validation for Amazon Alexa dataset (rounded to two decimal points)	68
Table 6.11 Twitter attributes calculations results of iPhone 11 dataset	70

Table 6.12 Twitter attributes calculations results of Amazon Alexa dataset	70
Table 6.13 Some tweets from the dataset	72
Table 6.14 Kylie Lip Kit dataset emotion prediction results	73
Table 6.15 Twitter attributes calculations results of Kylie Lip Kit dataset.....	73
Table 6.16 Root causes of Kylie Lip Kit's tweets.....	74

LIST OF APPENDICES

NBC with Cross Validation results for the iPhone 11 dataset	81
NBC with Cross Validation results for the Amazon Alexa dataset.....	82
SVM with Cross Validation results for the iPhone 11 dataset.....	83
SVM with Cross Validation results for the Amazon Alexa dataset	85