

Table of Content

DECLARATIONS.....	I
ACKNOWLEDGEMENTS	II
List of Table	I
List of Figure.....	III
Abstrak	V
Abstract	VI
CHAPTER 1.....	2
INTRODUCTION.....	2
1.1 Introduction	2
1.2 Motivation	3
1.2.1 Trending topic	3
1.2.2 Personalization in Social Media	4
1.2.3 Feature Selection for Classification	6
1.3 Problem Statement	6
1.4 Objective	8
1.5 Scope of Research	8
1.6 Contribution	9
1.7 Organization of Thesis	10
CHAPTER 2.....	12
LITERATURE REVIEW	12

2.1 Introduction	12
2.2 Trending Topic	12
2.3 Classification of Trending Topic	14
2.4 Classification of Tweets for Trending Topic	16
2.4.1 Emotion Classification of Tweets.....	16
2.4.2 Topical Classification of Tweets	18
2.4.3 Political Opinions Classification of Tweets	18
2.4.4 Incident Classification of Tweets	19
2.5 Feature Selection for Classification	20
2.5.2 Feature Selection by Part Of Speech (POS).....	21
2.5.3 Feature Selection by TF-IDF.....	22
2.5.4 Similarity Matching for Classification.....	23
2.5.5 Boolean Similarity Matching	24
2.6 The limitation of Related Works.....	24
2.7 Summary	24
CHAPTER 3.....	26
METHODOLOGY.....	26
3.1 Introduction	26
3.2 An Overview of the Proposed Framework	26
3.3 Category Model Construction	29
3.3.1 An Overview of Category	29
3.3.2 Category Model Learning	30

3.3.3 Tokenization & Tagging by POS.....	31
3.3.4 Feature Selection	31
3.4 Classifying Tweets using Category Model.....	35
3.5 Summary	37
CHAPTER 4.....	38
TOPIC-BASED FEATURE SELECTION.....	38
4.1 Introduction	38
4.2 The Limitation of Short Text Message.....	38
4.3 Topic-based Features.....	40
4.4 Incorporating Topic-based Features in Category Model.....	41
4.4.1 Feature Selection from External News Source	41
4.4.2 Enrich Category Model with Topic-based Features	42
4.5 Summary	43
CHAPTER 5.....	44
EVALUATION	44
5.1 Introduction	44
5.2 Dataset.....	44
5.2.1 Tweets Data Collection	44
5.2.2 News Content Collection	45
5.2.3 Feature Analysis	46
5.3 Performance Metric.....	47
5.3.1 Accuracy.....	47

5.3.2 Score of Accuracy Rate	49
5.4 Experimental Result	51
5.4.1 Evaluation on Personalized Tweet Classification	51
5.4.2 Evaluation on the Impact of the Number of Training Text	53
5.4.3 Evaluation on the Impact of Topic-based Features	60
5.5 Finding and Discussion	67
5.5.1 Accuracy of Classification	67
5.5.2 Number of Training Tweets.....	68
5.5.3 Limitations	68
5.6 Summary	69
CHAPTER 6.....	70
CONCLUSION	70
6.1 Summary	70
6.2 Contributions.....	71
6.3 Limitation of the Research	72
6.4 Future Work.....	72
REFERENCE	74
Appendix A	77
Appendix B	83