



CONTENTS

COVER PAGE.....	i
APPROVAL PAGE	ii
PLAGIARISM STATEMENT	iii
PREFACE	iv
CONTENTS.....	iv
LIST OF FIGURES	ix
LIST OF TABLES	xii
ABSTRACT.....	xiii
CHAPTER I INTRODUCTION.....	1
1.1 Research Background.....	1
1.2 Research Problem	4
1.3 Research Scope	5
1.4 Research Objective.....	5
1.5 Research Benefits.....	5
CHAPTER II LITERATURE REVIEW.....	7
CHAPTER III BASIC THEORY	12
3.1 Technology Used	12
3.2 Preprocessing.....	12
3.3 Word Embeddings.....	13
3.3.1 <i>Word2Vec</i>	13
3.3.2 <i>GloVe</i>	14
3.4 N-grams	16
3.5 Recurrent Neural Network.....	17
3.5.1 <i>Forward Pass</i>	19



3.5.2	<i>Backward Pass</i>	19
3.5.3	<i>Bidirectional RNN</i>	20
3.6	Gated Recurrent Unit Network (GRU)	20
3.7	Performance Evaluation	22
3.7.1	<i>Accuracy</i>	22
3.7.2	<i>AUC (Area-under-the-curve) ROC (Receiver Operating Characteristic) curve</i> 22	
3.7.3	<i>Precision</i>	24
3.7.4	<i>Recall</i>	25
3.7.5	<i>F-1 Score</i>	25
	CHAPTER IV RESEARCH METHODOLOGY	26
4.1	Research Description	26
4.2	Research Phases	26
4.3	Data Acquisition	27
4.4	Data Preprocessing	28
4.5	Modelling	32
4.5.1	Word2Vec – Continious Bag Of Words (CBOW)	33
4.5.2	GloVe	34
4.5.3	RNN Model	34
4.5.4	GRU Model	35
4.6	Model Evaluation	36
4.6.1	Dataset Splitting For Training & Testing Data	36
4.6.2	Training & Validation on RNN & GRU	37
4.6.3	Testing on RNN & GRU	37
4.6.4	Classification Evaluation Results	38
	CHAPTER V IMPLEMENTATION	39



5.1	Specification of Hardware & Software.....	39
5.2	Implementation of Data Acquisition.....	40
5.3	Implementation of Preprocessing.....	40
5.3.1	Case Folding.....	40
5.3.2	Mention, Retweets & URL Removal.....	41
5.3.3	Punctuation Removal	41
5.3.4	Numbers Removal	42
5.3.5	Single Characters Removal.....	42
5.3.6	Whitespace Removal.....	43
5.3.7	Tokenization	43
5.3.8	Padding.....	43
5.3.9	Drop row.....	44
5.4	Modelling	45
5.4.1	Word2Vec – Skip-gram.....	45
5.4.2	Word2Vec – Continuous Bag of Words (CBOW).....	51
5.4.3	GloVe	52
5.4.4	No Word Embedding.....	56
5.5	Model Evaluation	57
5.5.1	Data Splitting	57
5.5.2	Implementation of Training & Validating on RNN Model.....	58
5.5.3	Implementation of Training & Validating on GRU Model.....	61
5.5.4	Implementation Testing on RNN	64
5.5.5	Implementation Testing on GRU	66
5.6	Classification Results	67
	CHAPTER VI RESULTS & DISCUSSION	68
6.1	Results of Preprocessing	68



6.1.1	Results of Data Exploration	69
6.2	Experimental Results.....	69
6.2.1	Using Recurrent Neural Network (RNN)	69
6.2.1.1	Training & Validating Data on RNN.....	69
6.2.1.2	Evaluation Testing on RNN	71
6.2.2	Using Gated Recurrent Unit Network (GRU).....	74
6.2.2.1	Training & Validating Data on GRU.....	74
6.2.2.2	Evaluation Testing on GRU.....	75
	CHAPTER VII CONCLUSIONS AND SUGGESTIONS	80
7.1	Conclusions.....	80
7.2	Suggestions.....	80
	REFERENCES.....	82