

## TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	<b>ABSTRACT</b>	<b>iii</b>
	<b>LIST OF TABLES</b>	<b>xii</b>
	<b>LIST OF FIGURES</b>	<b>xv</b>
	<b>LIST OF SYMBOLS AND ABBREVIATIONS</b>	<b>xvii</b>
<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	BACKGROUND	1
1.2	MOTIVATION	4
1.3	RESEARCH GOALS AND CONTRIBUTIONS	6
1.4	THESIS OUTLINE	8
1.5	LITERATURE REVIEW	10
1.5.1	Paraphrase Recognition	10
1.5.1.1	Paraphrase Corpora	11
1.5.1.2	Categories of Features	13
1.5.1.3	Paraphrase Recognition using Machine Learning	16
1.5.1.4	PR Systems using Intermediate Representations	19
1.5.2	Paraphrase Extraction	24
1.5.2.1	Approaches for Paraphrase Extraction	24
1.5.2.2	Fuzzy Clustering	28
1.5.3	Applications	32

<b>CHAPTER NO.</b>	<b>TITLE</b>	<b>PAGE NO.</b>
	1.5.3.1 Student Answer Evaluation Approaches	32
	1.5.3.2 Plagiarism Detection Methods	35
	1.5.3.3 News Headline Clustering Methods	38
	1.5.4 Summary	40
<b>2</b>	<b>PARAPHRASE RECOGNITION USING SUPERVISED LEARNING</b>	<b>42</b>
2.1	RBFNN BASED PARAPHRASE RECOGNIZER	43
2.1.1	Features used for Paraphrase Recognition	43
2.1.1.1	Lexical Features	44
2.1.1.2	Syntactic Features	46
2.1.1.3	Semantic Features	48
2.1.2	Design of RBFNN based PR System	51
2.2	PARAPHRASE RECOGNITION USING SVM	56
2.3	FEATURE SELECTION FOR PARAPHRASE RECOGNITION	60
2.3.1	Genetic Algorithm based Feature Selection	60
2.4	RESULTS AND DISCUSSION	64
2.4.1	Performance Comparison of RBFNN and SVM based PR Systems	65
2.4.2	Evaluation of GA based Feature Selection Approach	72

<b>CHAPTER NO.</b>	<b>TITLE</b>	<b>PAGE NO.</b>
	2.4.3 Performance Improvement using Equivalent Phrases	75
2.5	SUMMARY	79
<b>3</b>	<b>PARAPHRASE RECOGNITION USING INTERMEDIATE REPRESENTATIONS</b>	<b>80</b>
3.1	UNL BASED PARAPHRASE RECOGNITION	81
	3.1.1 Pre-Processing and Co-Reference Resolution	83
	3.1.2 UNL Conversion and Processing	84
	3.1.3 WSD and Similarity Calculation	84
	3.1.4 UNL Matching	85
3.2	PARAPHRASE RECOGNITION USING PA MATCHING	87
	3.2.1 Predicate Argument Representations	89
	3.2.2 Predicate Argument Structure Matching	90
	3.2.3 Classification of Sentence pairs	94
3.3	RESULTS	98
	3.3.1 Performance Evaluation of UNLPR System	98
	3.3.2 Performance Evaluation of PASPR System	101
3.4	SUMMARY	104
<b>4</b>	<b>PARAPHRASE EXTRACTION USING FUZZY CLUSTERING</b>	<b>105</b>
4.1	ARCHITECTURE OF PARAPHRASE EXTRACTION SYSTEM	105

<b>CHAPTER NO.</b>	<b>TITLE</b>	<b>PAGE NO.</b>
4.2	PREPROCESSING	107
4.3	FUZZY AGGLOMERATIVE CLUSTERING BASED ON VERBS	107
4.3.1	Initial Cluster Formation	109
4.3.2	Grouping of Clusters	110
4.3.3	Merging of Clusters	113
4.4	FUZZY DIVISIVE CLUSTERING	115
4.5	IDENTIFICATION AND GROUPING OF PARAPHRASES	118
4.6	EVALUATION METRICS	119
4.7	EXISTING EXTRACTION SYSTEM	122
4.8	RESULTS AND DISCUSSION	124
4.8.1	Experiments on MSRPC	125
4.8.2	Experiments on MSRVDC	130
4.8.3	Discussion	135
4.9	SUMMARY	137
<b>5</b>	<b>APPLICATIONS OF PARAPHRASE RECOGNITION AND EXTRACTION</b>	<b>138</b>
5.1	STUDENT ANSWER EVALUATION	139
5.1.1	Answer Evaluation using Paraphrase Recognition	139
5.1.2	Experiments and Results	140
5.1.2.1	Short Answer Grading Corpus	141
5.1.2.2	User Language Paraphrase Corpus	142
5.2	PLAGIARISM DETECTION	144
5.2.1	Performance Evaluation	146

<b>CHAPTER NO.</b>	<b>TITLE</b>	<b>PAGE NO.</b>
5.3	CROSS LANGUAGE FAQ ACCESS	154
5.4	CLUSTERING NEWS HEADLINES	156
5.5	SUMMARY	158
<b>6</b>	<b>CONCLUSION AND FUTURE WORK</b>	<b>160</b>
6.1	CONCLUSION	160
6.2	FUTURE WORK	162
	<b>APPENDIX 1</b>	<b>163</b>
	<b>APPENDIX 2</b>	<b>166</b>
	<b>REFERENCES</b>	<b>167</b>
	<b>LIST OF PUBLICATIONS</b>	<b>183</b>