

List of Tables

Table No.	Table Caption	Page No.
Table 1.1	Derivational Morpheme	4
Table 1.2	Derivational Morpheme for Noun	4
Table 1.3	Double Suffixation words	8
Table 1.4	History of IE, IR and NLP	11
Table 2.1	Sample List of Suffices and Recoding Rules of Lovins	24
Table 2.2	Sample Set of Output of Lovins Stemmer	24
Table 2.3	Sample Set of Output of Porter Stemmer	25
Table 2.4	Sample Set of Output of Paice Stemmer	26
Table 2.5	Sample Output of KStem Stemmer	27
Table 2.6	Similarity Matrix Based on Dice's similarity	29
Table 2.7	Overview of popular Rule-based and Statistical Stemming Techniques	32
Table 2.8	Research Studies related to String Distance Measures	33
Table 2.9	Summarization of Existing Popular Stemming Algorithms	36
Table 2.10	Different String Similarity Measures	43
Table 2.11	Word Pairs with Their Distance Measures	45
Table 3.1	Advantage and Disadvantage of Affix-Removal Stemmers	48
Table 3.2	Number of Stem Word Generation for Set of Wordlist	49
Table 3.3	Number of Stem Word Generation for Set of Wordlist (Elaboration of Table 3.2)	49
Table 3.4	Comparative Result of Affix-Removal Stemmers	50
Table 3.5	Output for Nominalized/Derived Words by Existing Online Lemmatizers	55
Table 4.1	Derivative Noun Suffix Rules from Verbs	64
Table 4.2	Derivative Verb Suffix Rules from Noun	65
Table 4.3	Derivative Adjective Suffix Rules from Verb	65
Table 4.4	Derivative Adjective Suffix Rules from Adjective	65
Table 4.5	Derivative Adjective Suffix Rules from Noun	65
Table 4.6	Derivative Noun Suffix Rules from Verb/Noun	66
Table 4.7	Derivative Verb Suffix Rules from Verb	68
Table 4.8	Derivative Adverb Suffix Rules from Adjective	68
Table 4.9	Derivative Noun Suffix Rules from Noun	68
Table 4.10	Sample Input Text for POS Tagging	75
Table 4.11	Output of Stanford POS-Tagger	75
Table 4.12	Sample of POS Tag Set	75
Table 4.13	Sample Data Set of Derived and Inflected Words (Mixed POS) with Lemmas	76
Table 4.14	Sample Data Set of Derived Noun Words with Lemmas	78
Table 4.15	Sample Data Set of Derived Adjective Words with Lemma	79
Table 4.16	Sample Data Set of Mixed POS Words with Lemmas	81

Table 4.17 Sample Data Set of Derived Adverb Words with Lemmas_____	82
Table 4.18 Comparative Analysis between LemmaChase and Stanford Lemmatizer_____	83
Table 5.1 Sample Set of Suffix Rules_____	98
Table 5.2 Sample of Double suffixation Derived words and Lemmas_____	98
Table 5.3 A. Sample of Input Text of Brown Corpus_____	99
Table 5.4 Sample of POS tagged Text_____	99
Table 5.5 POS tagged Unique Sample Word List_____	99
Table 5.6 List of Word-Pair with their distance value_____	100
Table 5.7 Groups with Lemma List_____	100
Table 5.8 Discrete Word List_____	101
Table 5.9 Input Words with their Lemmas_____	101
Table 5.10 B. Sample of Input Text of Brown_____	102
Table 5.11 POS tagged Text_____	103
Table 5.12 POS tagged Unique Sample Word List_____	103
Table 5.13 List of Word-Pair with Their Distance Value _____	103
Table 5.14 List of Groups with Lemmas_____	104
Table 5.15 Discrete Word List_____	104
Table 5.16 Input Words with Their Lemmas_____	105
Table 5.17 Sample Input-DUC Text_____	105
Table 5.18 Sample POS-tagged Text_____	105
Table 5.19 Unique Words List_____	106
Table 5.20 Word-Pair with Distance Value_____	106
Table 5.21 Groups with Lemmas_____	106
Table 5.22 Output: Collection of Lemmas of Discrete Words_____	106
Table 5.23 Sample Text of Morphed words_____	107
Table 5.24 Group-wise Lemma _____	107
Table 5.25 Oxford Dictionary Sample Words with Lemmas_____	107
Table 5.26 Nominalized Adverb Words with Lemmas_____	108
Table 5.27 Nominalized/Derived Words with Lemmas_____	108
Table 5.28 Double/Triple Suffixed Nominalized/Derived Words with Lemmas_____	109
Table 5.29 Comparative Study between lemmatizers_____	110
Table 5.30 Resources with Word and Lemma Count_____	111
Table 5.31 Word, Lemma Count with Error Count_____	112
Table 5.32 Comparative Output of LemmaQuest vs. LemmaChase vs. Stanford Lemmatizer_____	115
Table 5.33 Precision, Recall and F-Score value for LemmaQuest_____	116