

Contents

Cover Page.....	i
Certificate.....	ii
Declaration.....	iii
Copyright Transfer Certificate.....	iv
Acknowledgements.....	v
Preface.....	vii
TOC.....	x
List of Figures.....	xiv
List of Tables.....	xv
Chapter 1 Introduction.....	1
1. Introduction.....	1
1.1 Background and Motivation	2
1.2 Objectives	5
1.3 Methodology.....	6
1.4 Interdisciplinary relevance of the work	8
1.5 Outline of the Thesis.....	10
1.6 Conclusion	13
Chapter 2 An Overview of Stammering	14
2. Introduction.....	14
2.1 What is stammering?	14
2.2.1 By Age of Onset.....	17
2.2.2 By Persistence.....	18
2.2.3 By Severity.....	19
2.3 Prevalence and Demographics.....	21

2.4	Etiology and Risk Factors.....	26
2.5	Impact of Stammering	28
2.6	Assessment and Diagnosis.....	31
2.7	Treatment Approaches.....	33
2.8	Outcome Measures and Treatment Efficacy	35
2.9	Linguistic Aspects of Stammering	38
2.10	Conclusion	41
Chapter 3 Review of Literature.....		43
3.	Introduction.....	43
3.1	Linguistic Aspect.....	43
3.2	Computational Aspect	53
3.4	Conclusion	58
Chapter 4 Data collection and Annotation		60
4.	Introduction.....	60
4.1	Participants	61
4.2	Data collection.....	62
4.2	Tag Set.....	66
4.3	Knowledge representation in corpus	70
4.4	Conclusion.....	73
Chapter 5 Linguistic Analysis.....		76
5.	Introduction.....	76
5.1	Linguistic analysis of stammering.....	77
5.1.1	Phonological analysis.....	78
5.1.2	Morphological analysis.....	84
5.1.3	Syntactic analysis.....	86
5.1.4	Semantic and Pragmatic analysis.....	87
5.1.5	Socio-Linguistic Analysis	88
5.2	Conclusion.....	90
Chapter 6 Computational Experiments.....		91
6.	Introduction.....	91
6.1	Machine learning	93
6.2	Classification	94

6.3	Audio to Spectrogram Conversion	95
6.4	Spectrogram as Image	96
6.5	Audio classification	97
6.6	Pros and cons analysis	99
6.7	Rational behind selecting image-based classification	102
6.8	Data augmentation.....	104
6.9	Data splitting.....	106
6.10	Spectrograms Generation	108
6.10.1	Training Data	109
6.11	Model architecture.....	110
6.12	Optimizer and loss function.....	112
6.13	Rescaling	115
6.14	Color as a hyperparameter.....	116
6.15	Training Patience.....	117
6.16	Gradual increase of training data.....	119
6.17	Usages.....	120
6.18	Results	121
6.19	Wake word.....	122
6.19.1	Problems with Wake Words	124
6.19.2	Gaps in Wake Word Technology.....	125
6.19.3	Proposing a "Sleep Word" for Improved Voice Assistant Functionality	127
6.19.4	Key Features and Functionality	128
6.20	Conclusion.....	130
Chapter 7 Discussion and Conclusion		132
7.	Summary of the work.....	132
7.1	Linguistic Aspect.....	134
7.2	Computational Aspect	134
7.3	Forensic Science Aspect.....	135
7.4	Use cases and Applications	136
7.4.1	Speech therapy for Hindi Speakers.....	137
7.4.2	Content Filtering.....	140
7.4.3	Transcription services	142

7.4.4	Language learning apps	145
7.4.5	Accessibility features	147
7.4.6	Multilingual expansion	149
7.5	Limitations.....	152
7.6	Challenges	153
7.7	Future Plan.....	153
	Bibliography.....	156
	Appendix 1 List of Scopus Indexed Journal Published Papers.....	164
	Appendix 2 Sentences used in textual data collection.....	165

List of Figures

Figure 1: A sentence with three-tier analysis in PRAAT	72
Figure 2: Deviation instances in the non-random corpus	80
Figure 3: Deviation instances in the random corpus.....	81
Figure 4: Deviation trends based on the phonetic analysis.....	82
Figure 5: Deviation trends based on the word length	83
Figure 6: A sample of Spectrogram generated from the audio data	109
Figure 7: Comparison of training-validation accuracy and Training and validation loss.....	116
Figure 8: Comparison of average precision, precision and recall of both models.....	117
Figure 9: Comparison of training-validation accuracy, and training-validation loss of both models.....	117
Figure 10: Comparison of training epochs of colored and grayscale model	118
Figure 11: Correlation between training data size and number of epochs.....	119
Figure 12: Accuracy results on changing the training data size	120
Figure 13: Binary classifier pipeline.....	121
Figure 14: Confusion matrix derived from our experiments involving the top-performing colored and grayscale models.....	122

List of Tables

Table 1: Existing Annotation Scheme	67
Table 2: Difference in Intended and perceived words	79
Table 3: Syllable classes basis their structure and class contribution in stammering.....	79
Table 4: Sounds and their phonetic properties.....	80
Table 5: Syllabic Analysis.....	83
Table 6: Number of participants and count of phrase repetitions.....	87
Table 7: Model's architecture, with the layers and their corresponding output shapes.....	111