List of Figures

1.1 Physiology of speech production. ............................................. 5
1.2 Speaker diarization principle. ................................................. 8
3.1 Extraction of MFCC from speech signal. ................................. 46
3.2 Mel scale filter bank. ........................................................... 48
3.3 Principle of support vector machines. ...................................... 51
3.4 (a) Speech signal. (b) LP residual signal. (c) Hilbert transform of residual signal. (d) Hilbert envelope. (e) Residual phase. ................. 55
3.5 (a) Speech signal. (b) and (c) Misclassification rates of residual phase. (d) and (e) Misclassification rates of MFCC. (f) and (g) Misclassification rates of residual phase with MFCC. ......................... 59
3.6 Performance of residual phase, MFCC, and residual phase with MFCC. ... 60
3.7 Snapshot of the speaker segmentation using SVM. ..................... 61
4.1 A five layer AANN model. ................................................... 64
4.2 Distribution capturing ability of AANN model. (a) Artificial 2 dimensional data. (b) 2 dimensional output of AANN model. (c) Probability surfaces realized by the network. ............................. 65
4.3 Effect of epochs on the confidence score. ................................ 70
4.4 Performance of the algorithm for various analysis window size. ....... 71
4.5 Effect of adjustable parameter (a) on the performance of the algorithm. ... 72
4.6 Snapshot of the speaker segmentation using AANN. .................... 74
5.1 Flow chart of the proposed speaker diarization system. 84
5.2 Snapshot of the speaker diarization using AANN. 85
5.3 Snapshot of the result of speaker diarization using AANN. 86

6.1 Calculation of the shifted delta feature vectors. 90
6.2 Effect of acoustic features for 3 sec. test utterance for HMM based LID system. 96
6.3 Effect of acoustic features for 10 sec. test utterance for HMM based LID system. 97
6.4 Effect of acoustic features for 30 sec. test utterance for HMM based LID system. 97
6.5 Effect of training set size for MFCC features for HMM based LID system. 98
6.6 Effect of training set size for MFCC with delta and acceleration features for HMM based LID system. 99
6.7 Effect of training set size for SDC features for HMM based LID system. 100
6.8 Effect of test utterance duration for 200 sec. training set for HMM based LID system. 101
6.9 Effect of test utterance duration for 300 sec. training set for HMM based LID system. 103
6.10 Effect of test utterance duration for 400 sec. training set for HMM based LID system. 104
6.11 Effect of acoustic features for 3 sec. test utterance for GMM based LID system. 105
6.12 Effect of acoustic features for 10 sec. test utterance for GMM based LID system. 106
6.13 Effect of acoustic features for 30 sec. test utterance for GMM based LID system. ........................................ 107
6.14 Effect of training set size for MFCC features for GMM based LID system. ........................................ 109
6.15 Effect of training set size for MFCC with delta and acceleration features for GMM based LID system. .......... 110
6.16 Effect of training set size for SDC features for GMM based LID system. ........................................ 111
6.17 Effect of test utterance duration for 200 sec. training set for GMM based LID system. .................................... 112
6.18 Effect of test utterance duration for 300 sec. training set for GMM based LID system. .................................... 113
6.19 Effect of test utterance duration for 400 sec. training set for GMM based LID system. .................................... 114
6.20 Snapshot of the language identification for selecting the model. ................................................ 115
6.21 Snapshot of the language identification for training the model with the existing speech files. ...................... 116
6.22 Snapshot of the language identification for training the model by adding new files to the existing speech files. ................................................ 117
6.23 Snapshot of the language identification for training the model by new set of training files. .............................. 118
6.24 Snapshot of the result obtained for Dravidian family from language identification system. ...................... 119
6.25 Snapshot of the result obtained for Aryan family from language identification system. .............................. 120