

## TABLE OF CONTENT

<b>PERNYATAAN</b> .....	<b>iii</b>
<b>PRAKATA</b> .....	<b>iv</b>
<b>TABLE OF CONTENT</b> .....	<b>v</b>
<b>LIST OF TABLES</b> .....	<b>viii</b>
<b>LIST OF FIGURES</b> .....	<b>ix</b>
<b>INTISARI</b> .....	<b>xi</b>
<b>ABSTRACT</b> .....	<b>xii</b>
<b>CHAPTER I INTRODUCTION</b> .....	<b>1</b>
1.1 Research Background.....	1
1.2 Research Problem.....	3
1.3 Research Objective.....	3
1.4 Research Scopes.....	3
1.5 Research Advantage.....	4
1.6 Research Schematic .....	4
<b>CHAPTER II LITERATURE REVIEW</b> .....	<b>6</b>
<b>CHAPTER III THEORETICAL BASIS</b> .....	<b>11</b>
3.1 Gaussian Mixture Models (GMM) .....	11
3.2 Support Vector Machines (SVM) .....	14
3.3 Mel Frequency Cepstral Coefficients (MFCC).....	18
3.4 NearMiss Algorithm .....	20
3.5 Hybrid Method .....	22
3.6 Linear Weighted Combination .....	22
3.7 Classification Metrics .....	23
3.7.1 Confusion Matrix .....	24
3.7.2 Accuracy Score .....	25
3.7.3 Precision.....	25
3.7.4 Recall .....	26
3.7.5 F-Score .....	26
<b>CHAPTER IV RESEARCH METHODOLOGY</b> .....	<b>27</b>
4.1 Research Description .....	27
4.2 Data Acquisition.....	28
4.2.1 Data Acquisition of VoxCeleb Dataset.....	29
4.2.2 Data Acquisition of Free ST American English Corpus (SLR45) Dataset .....	29

4.3	Data Preprocessing.....	30
4.3.1	Data Organization .....	30
4.3.2	File Format.....	30
4.3.3	Normalizing Audio .....	31
4.3.4	Trimming Silence in Audio .....	33
4.3.5	Padding and Trimming Audio .....	35
4.3.6	Noise Reduction in Audio.....	36
4.4	Data Splitting .....	37
4.5	MFCC Feature Extraction .....	38
4.6	Undersampling .....	39
4.7	Sequential Model .....	40
4.7.1	GMM for Sequential Model.....	41
4.7.1.1	Likelihood Scoring for Sequential Model .....	43
4.7.2	SVM Training for Sequential Model .....	44
4.8	Parallel Model .....	45
4.8.1	GMM Training for Parallel Model .....	45
4.8.2	SVM Training for Parallel Model.....	46
4.8.3	Linear Weighted Combination.....	47
4.9	Model Evaluation .....	48
4.9.1	Sequential Model Evaluation .....	48
4.9.2	Parallel Model Evaluation.....	48
4.9.3	Comparison .....	48
	<b>CHAPTER V RESEARCH IMPLEMENTATION.....</b>	<b>49</b>
5.1	Research Environment .....	49
5.2	Importing Libraries .....	49
5.3	Dataset Initialisation.....	50
5.3.1	Data Pre-Processing .....	50
5.3.2	Data Splitting .....	52
5.3.3	Feature Extraction .....	54
5.3.4	Data Undersampling .....	56
5.6	Sequential Model Implementation .....	59
5.6.1	GMM Training.....	59
5.6.2	Likelihood Scoring.....	60
5.6.3	SVM Training .....	62
5.7	Parallel Model Implementation.....	65
5.7.1	GMM Training.....	65
5.7.2	SVM Training .....	65
5.7.3	Linear Weighted Combination.....	67
5.8	Model Evaluation .....	70
5.8.1	Sequential Model Evaluation .....	70
5.8.2	Parallel Model Evaluation.....	72
	<b>CHAPTER VI RESULT AND DISCUSSION.....</b>	<b>74</b>
6.1	Data Gathering .....	74
6.2	Pre-Processing.....	75



6.3 Data Splitting .....	78
6.4 MFCC Feature Extraction .....	79
6.5 Undersampling .....	81
6.6 Hyper parameter Tuning .....	82
6.6.1 Sequential Model .....	82
6.6.2 Parallel Model .....	84
6.7 Model Evaluation .....	86
6.7.1 Sequential Model Evaluation .....	86
6.7.2 Parallel Model Evaluation .....	88
6.8 Comparison .....	90
6.9 Analysis .....	91
<b>CHAPTER VII CONCLUSION .....</b>	<b>93</b>
7.1 Conclusion .....	93
7.2 Future Work .....	94
<b>REFERENCES .....</b>	<b>96</b>