

# Contents

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	What is a Speaker Verification System? . . . . .	1
1.2	What is Machine Learning? . . . . .	2
1.3	Road Map of the Thesis . . . . .	2
<b>2</b>	<b>Text-Independent Speaker Verification Systems</b>	<b>5</b>
2.1	Machine Learning Tools . . . . .	6
2.1.1	Diagonal Covariance Gaussian Mixture Models . . . . .	6
2.1.2	Support Vector Machines . . . . .	8
2.2	GMM Based System . . . . .	10
2.3	Feature Extraction . . . . .	11
2.3.1	Cepstral Parameters . . . . .	12
2.3.2	Additional Transformations . . . . .	14
2.3.3	Silence/Speech Detector . . . . .	14
2.4	Score Normalization . . . . .	15
2.4.1	T-norm . . . . .	16
2.4.2	Z-norm . . . . .	16
2.5	SVM and GLDS Kernel . . . . .	17
2.6	Conclusion . . . . .	19
<b>3</b>	<b>Performance Measures for Speaker Verification</b>	<b>21</b>
3.1	Common Measures . . . . .	22
3.2	Expected Performance Curve . . . . .	25
3.2.1	Cautious Interpretation of ROC and BEP . . . . .	28
3.2.2	Expected Performance Curve: an “a priori” Performance Curve . . . . .	30

---

3.3	Statistical Tests . . . . .	34
3.3.1	The Z-Test on Proportions . . . . .	34
3.3.2	$Z_{\text{HTER}}$ -Test: a Statistical Test for HTERs . . . . .	37
3.3.3	Other Statistical Tests . . . . .	40
3.3.4	Analysis . . . . .	42
3.4	Methodology and Presentation of Results . . . . .	48
3.5	Conclusion . . . . .	49
<b>4</b>	<b>Experimental Methodology</b>	<b>53</b>
4.1	Methodology . . . . .	53
4.2	Databases . . . . .	54
4.2.1	Banca . . . . .	54
4.2.2	Polyvar . . . . .	57
4.2.3	NIST . . . . .	58
4.3	Conclusion . . . . .	60
<b>5</b>	<b>Text-Independent Speaker Verification: a Machine Learning Perspective</b>	<b>63</b>
5.1	Framework . . . . .	64
5.1.1	Statistical Framework . . . . .	65
5.1.2	A Score Based Framework . . . . .	68
5.2	Are GMMs Discriminant? . . . . .	69
5.2.1	GMMs: a Mixture of Linear Classifiers . . . . .	69
5.2.2	Experimental Results . . . . .	72
5.2.3	Discussion . . . . .	74
5.3	Score Normalization . . . . .	74
5.3.1	Unified Framework for Score Normalization . . . . .	75
5.3.2	Relation to Existing Normalization Techniques . . . . .	76
5.3.3	Comparison Between New and Classical Z- and T-norm . . . . .	79
5.3.4	Experiments . . . . .	79
5.3.5	T-norm for SVM . . . . .	80
5.4	Conclusion . . . . .	83
<b>6</b>	<b>GMMs and Discriminant Models</b>	<b>87</b>
6.1	Learning the Decision Function . . . . .	88
6.2	HTER Cost Function . . . . .	89
6.3	GMM LLR . . . . .	90
6.4	GMM Gaussian LLR . . . . .	91
6.5	Posterior Based Approach . . . . .	92

---

6.6	Conclusion . . . . .	96
<b>7</b>	<b>Sequence Kernel Based Speaker Verification</b>	<b>97</b>
7.1	Mean Operator Kernel . . . . .	98
7.2	Max Operator Kernel . . . . .	102
7.3	Non-Mercer Kernels . . . . .	104
7.4	Experimental Results on Polyvar and Banca Databases . . . . .	105
7.4.1	Polyvar . . . . .	106
7.4.2	Banca . . . . .	107
7.5	Smoothing the Max Kernel . . . . .	109
7.6	Clustering Techniques . . . . .	110
7.7	Experimental Results on the NIST Database . . . . .	114
7.8	Conclusion . . . . .	117
<b>8</b>	<b>A New Perspective: Working on the Distance Measure</b>	<b>119</b>
8.1	Unbalanced SVM Criteria . . . . .	120
8.2	Class Dependent RBF Kernel . . . . .	121
8.3	Conclusion . . . . .	125
<b>9</b>	<b>Conclusion</b>	<b>127</b>
9.1	Contribution of the Thesis . . . . .	127
9.2	Other contributions . . . . .	133