

Contents

List of Tables	IX
List of Figures	XI
List of Algorithms	XIII
1 Introduction	1
1.1 Structure of the Book	3
2 Supervised Sequence Labelling	5
2.1 Supervised Learning	5
2.2 Pattern Classification	6
2.2.1 Probabilistic Classification	6
2.2.2 Training Probabilistic Classifiers	7
2.2.3 Generative and Discriminative Methods	8
2.3 Sequence Labelling	9
2.3.1 Sequence Classification	11
2.3.2 Segment Classification	11
2.3.3 Temporal Classification	12
3 Neural Networks	15
3.1 Multilayer Perceptrons	15
3.1.1 Forward Pass	16
3.1.2 Output Layers	18
3.1.3 Loss Functions	19
3.1.4 Backward Pass	20
3.2 Recurrent Neural Networks	22
3.2.1 Forward Pass	22
3.2.2 Backward Pass	23
3.2.3 Unfolding	24
3.2.4 Bidirectional Networks	24
3.2.5 Sequential Jacobian	27
3.3 Network Training	29
3.3.1 Gradient Descent Algorithms	29
3.3.2 Generalisation	30

3.3.3	Input Representation	34
3.3.4	Weight Initialisation	35
4	Long Short-Term Memory	37
4.1	Network Architecture	37
4.2	Influence of Preprocessing	41
4.3	Gradient Calculation	42
4.4	Architectural Variants	42
4.5	Bidirectional Long Short-Term Memory	43
4.6	Network Equations	43
4.6.1	Forward Pass	44
4.6.2	Backward Pass	44
5	A Comparison of Network Architectures	47
5.1	Experimental Setup	47
5.2	Network Architectures	48
5.2.1	Computational Complexity	49
5.2.2	Range of Context	49
5.2.3	Output Layers	49
5.3	Network Training	51
5.3.1	Retraining	51
5.4	Results	52
5.4.1	Previous Work	54
5.4.2	Effect of Increased Context	54
5.4.3	Weighted Error	56
6	Hidden Markov Model Hybrids	57
6.1	Background	57
6.2	Experiment: Phoneme Recognition	58
6.2.1	Experimental Setup	59
6.2.2	Results	59
7	Connectionist Temporal Classification	61
7.1	Background	61
7.2	From Outputs to Labellings	63
7.2.1	Role of the Blank Labels	64
7.2.2	Bidirectional and Unidirectional Networks	64
7.3	Forward-Backward Algorithm	64
7.3.1	Log Scale	68
7.4	Loss Function	68
7.4.1	Loss Gradient	69
7.5	Decoding	71
7.5.1	Best Path Decoding	71
7.5.2	Prefix Search Decoding	72
7.5.3	Constrained Decoding	73
7.6	Experiments	78

7.6.1	Phoneme Recognition 1	79
7.6.2	Phoneme Recognition 2	80
7.6.3	Keyword Spotting	81
7.6.4	Online Handwriting Recognition	85
7.6.5	Offline Handwriting Recognition	88
7.7	Discussion	91
8	Multidimensional Networks	95
8.1	Background	95
8.2	Network Architecture	97
8.2.1	Multidirectional Networks	100
8.2.2	Multidimensional Long Short-Term Memory	102
8.3	Experiments	104
8.3.1	Air Freight Data	104
8.3.2	MNIST Data	105
8.3.3	Analysis	107
9	Hierarchical Subsampling Networks	109
9.1	Network Architecture	110
9.1.1	Subsampling Window Sizes	112
9.1.2	Hidden Layer Sizes	112
9.1.3	Number of Levels	113
9.1.4	Multidimensional Networks	113
9.1.5	Output Layers	115
9.1.6	Complete System	116
9.2	Experiments	117
9.2.1	Offline Arabic Handwriting Recognition	119
9.2.2	Online Arabic Handwriting Recognition	122
9.2.3	French Handwriting Recognition	124
9.2.4	Farsi/Arabic Character Classification	126
9.2.5	Phoneme Recognition	127
	References	133
	Acknowledgements	141