

Contents

1	Introduction	1
1.1	Time-Series Prediction	1
1.2	Machine Learning Techniques for Time-Series Prediction	3
1.3	Importance of Feature Selection in Machine Learning	5
1.4	Towards Automatic Feature Generation	8
1.5	Contributions to Formalising a Better Hypothesis	9
1.6	Chapter Organisation	10
	References	10
2	Feature Selection	13
2.1	Introduction	13
2.1.1	Redundant and Irrelevant Features	14
2.1.2	Feature Interaction	14
2.1.3	Over-Fitting and the Optimal Number of Features	14
2.1.4	Feature Quality Evaluation	15
2.2	Feature Selection Models	15
2.2.1	Filter Models	16
2.2.2	Wrapper Models	19
2.2.3	Embedded Models	21
	References	23
3	Grammatical Evolution	25
3.1	Introduction	25
3.2	Genetic Algorithms	25
3.3	Context-Free Grammar	27
3.4	Generating Features Using GE	27
3.5	gramEvol Package	29
	References	33

4 Grammar Based Feature Generation	35
4.1 Automatic Feature Generation	35
4.2 Feature Generation Using Context-Free Grammars	35
4.2.1 Grammar Based Feature Generation Systems	36
4.2.2 Feature Generation Framework Architecture	37
4.3 Wavelet Based Grammar for Electricity Load Time-Series Prediction	38
4.4 Grammar Families	40
4.4.1 Technical Indicators	40
4.4.2 Generating Technical Indicators Using Grammar Families	41
4.5 Implementation	44
4.5.1 Pruning Strategies	45
4.5.2 Implementation Issues	45
4.6 Feature Generation Using GE	47
References	50
5 Application of Grammar Framework to Time-Series Prediction . . .	51
5.1 Expert-Suggested Features	51
5.1.1 Stock Index Time-Series Prediction	53
5.1.2 Electricity Load Demand Time-Series Prediction	54
5.1.3 Better Feature Combinations Than Common Input Features?	55
5.2 Data Partitioning for Time-Series Prediction	55
5.2.1 Data Preprocessing	56
5.2.2 Model Selection and Parameter Tuning	56
5.3 Feature Selection Techniques	57
5.3.1 Dimensionality Reduction Using PCA	58
5.3.2 Feature Selection Using Integer Genetic Algorithms	59
5.3.3 Wrapper-Based Feature Subset Evaluation	60
5.4 Avoiding Peeking	61
References	61
6 Case Studies	63
6.1 Predicting Stock Indices	63
6.2 Predicting Peak Electricity Load Data	70
6.3 Predicting Foreign Exchange Client Trade Volume	77
References	82
7 Conclusion	85

Contents	xi
Appendix A: Wavelet Transformations	89
Appendix B: Production Rules for Some Standard Technical Indicators	91
Appendix C: Supplementary Results	95

Grammar-Based Feature Generation for Time-Series
Prediction

De Silva, A.M.; Leong, P.H.W.

2015, XI, 99 p. 28 illus., Softcover

ISBN: 978-981-287-410-8