
Contents

1	Machine learning for healthcare technologies – an introduction	1
	<i>David A. Clifton</i>	
1.1	The changing needs of healthcare	1
1.2	Online resources	1
1.3	Survey of contents	2
	Acknowledgement	6
2	Detecting artifactual events in vital signs monitoring data	7
	<i>Partha Lal, Christopher K. I. Williams, Konstantinos Georgatzis, Christopher Hawthorne, Paul McMonagle, Ian Piper and Martin Shaw</i>	
2.1	Introduction	7
2.2	Data collection	8
2.3	Data preprocessing	9
2.4	Data annotation	9
2.5	The effect of data cleaning on data summary measures	13
2.6	Factorial switching linear dynamical system	13
2.6.1	Factors	15
2.6.2	Channels	16
2.6.3	Inference	16
2.7	Discriminative switching linear dynamical system	16
2.7.1	Predicting s_t	17
2.7.2	Predicting x_t	18
2.7.3	Combining the FSLDS and DSLDS predictions for s_t	18
2.8	Stability detection	19
2.9	Real-time implementation	20
2.9.1	Computational efficiency	20
2.9.2	Stability model estimation	20
2.10	Software	21
2.11	Experiments	22
2.12	Conclusions and future work	28
2.13	Appendix: Models for each factor	29
2.13.1	Stability	29
2.13.2	Blood sample events	29
2.13.3	Damped trace events	30

2.13.4	Suction events	30
2.13.5	X-factor	30
2.13.6	Overwriting order of factors	31
	Acknowledgements	31
	References	31
3	Signal processing and feature selection preprocessing for classification in noisy healthcare data	33
	<i>Qiao Li, Chengyu Liu, Julien Oster and Gari D. Clifford</i>	
3.1	Introduction	33
3.2	Preprocessing and database	35
3.2.1	QRS detection	35
3.2.2	Signal quality assessment	35
3.2.3	Datasets	36
3.2.4	Adding realistic noise to known data	37
3.3	Feature extraction	37
3.3.1	Time-domain features	38
3.3.2	Frequency-domain features	38
3.3.3	Nonlinear features	38
3.4	Feature selection	38
3.4.1	Forward likelihood ratio selection for logistic regression	39
3.4.2	Recursive feature elimination for support vector machine	40
3.5	Evaluation metrics	42
3.6	Results	42
3.6.1	Feature results comparison between AF and Non-AF	42
3.6.2	Model development phase	43
3.6.3	Model validation phase	47
3.7	Discussion	53
	Appendix 1	54
	References	56
4	ECG model-based Bayesian filtering	59
	<i>Julien Oster</i>	
4.1	Background	59
4.2	Theory	60
4.2.1	Bayesian filtering	60
4.2.2	Non-linear Bayesian filtering	61
4.2.3	Switching Kalman filters	62
4.3	ECG model	63
4.4	Denoising	65
4.4.1	Problem formulation	65
4.4.2	Parameter initialisation	66
4.4.3	Benchmarking and results	66

4.5	Delineation	67
4.5.1	Problem formulation	67
4.5.2	Benchmarking and results	69
4.6	Source separation	69
4.6.1	Problem formulation	70
4.6.2	Benchmarking and results	71
4.7	Detection of pathological beats	72
4.7.1	Problem formulation	73
4.7.2	Parameter initialisation	74
4.7.3	Benchmarking and results	75
4.8	Discussion	76
	References	78
5	The power of tensor decompositions in biomedical applications	83
	<i>Borbála Hunyadi, Sabine Van Huffel and Maarten De Vos</i>	
5.1	Introduction to tensors	83
5.2	Tensor decomposition techniques	85
5.2.1	Decomposition of matrices	85
5.2.2	Decomposition of tensors	87
5.3	Construction of tensors in biomedical applications	90
5.4	Naturally occurring tensors	90
5.4.1	Genomic data	90
5.4.2	Repeated multichannel measurements	91
5.4.3	Epoched multichannel measurements	91
5.5	Tensor expansion of matrix data	92
5.5.1	Frequency transformation	92
5.5.2	Hankel structure	92
5.5.3	Representation by means of a feature set	93
5.6	Successful decompositions of biomedical data tensors	93
5.7	Unsupervised tensor decompositions	94
5.7.1	Blind source separation	94
5.7.2	Unsupervised classification	97
5.8	Supervised tensor decompositions	97
5.9	Coupled tensor decompositions	98
5.9.1	Coupling of multi-subject data	99
5.9.2	Temporal coupling	100
5.9.3	Spatial coupling	102
5.10	Practical considerations	102
5.11	Parameter selection	102
5.12	Initialization	104
5.13	Tools and algorithms	104
	Acknowledgements	105
	References	105

6 Patient physiological monitoring with machine learning	111
<i>Marco A. F. Pimentel and David A. Clifton</i>	
6.1 Introduction	111
6.2 Methodology	113
6.2.1 Dataset	113
6.2.2 Gaussian processes	114
6.2.3 Time-series clustering	117
6.3 Results	119
6.4 Discussion	120
6.5 Conclusion	123
Acknowledgements	123
References	124
 7 A Bayesian model for fusing biomedical labels	 127
<i>Tingting Zhu, Gari D. Clifford and David A. Clifton</i>	
7.1 Background	127
7.2 A generative model of annotators	130
7.2.1 The ground truth model	130
7.2.2 The annotator model	132
7.3 Bayesian probability in parameter estimation	133
7.4 The Bayesian continuous-valued label aggregator	136
7.4.1 The MAP approach of the BCLA model	137
7.4.2 Convergence criteria for the BCLA-MAP model	138
7.4.3 Learning from incomplete data using the BCLA-MAP model	140
7.5 Data description	141
7.5.1 Simulated QT dataset with independent annotators	141
7.5.2 The 2006 PhysioNet challenge QT dataset	142
7.5.3 Methodology of validation and comparison	147
7.6 Results and discussion	148
7.6.1 Simulated dataset	148
7.6.2 PCinC QT dataset	149
7.7 Conclusion and future work	155
Acknowledgement	156
References	156
 8 Incorporating end-user preferences in predictive models	 161
<i>Suchi Saria and Daniel P. Robinson</i>	
8.1 Introduction	161
8.1.1 Background and motivation	164
8.1.2 Related work	165
8.1.3 Key contributions	167

8.2	Regularizers for complex cost structures	167
8.2.1	An example from the ICU	167
8.2.2	The structured regularizer for the general case	171
8.2.3	Relaxations of our exact structured regularizer	172
8.3	Numerical experiments	172
8.3.1	Dataset	173
8.3.2	Experimental setup	173
8.3.3	Model diversity	174
8.3.4	Comparison with the ℓ_1 - and scaled ℓ_1 -norm	175
8.4	Conclusions and discussion	177
	References	178
9	Variational Bayesian non-parametric inference for infectious disease models	181
	<i>James Hensman and Theodore Kypraios</i>	
9.1	Introduction	181
9.1.1	Infectious disease modelling	181
9.1.2	Why non-parametric inference?	181
9.1.3	Previous work	182
9.2	Background	183
9.2.1	Gaussian processes	183
9.2.2	Variational Bayes	185
9.3	Modelling framework	186
9.3.1	SIR model definition	186
9.3.2	Approximating the SIR model with a log Gaussian Cox process	187
9.3.3	Relaxing the parametric assumptions of the SIR model	188
9.3.4	Bayesian inference for an LGCP	188
9.3.5	Sparse variational approximations to GPs	189
9.4	Results	191
9.4.1	Dataset 1: Synthetic data from a homogeneously mixing mass-action SIR model	191
9.4.2	Dataset 2: Synthetic data from a seasonal SIR model	193
9.4.3	Application to the Abakaliki Smallpox data	195
9.5	Conclusions	198
	Acknowledgements	199
	References	199
10	Predicting antibiotic resistance from genomic data	203
	<i>Yang Yang, Katherine E. Niehaus and David A. Clifton</i>	
10.1	Antibiotic resistance	203
10.2	Susceptibility test to antibiotics	205

10.3	Genomic data associated with antibiotic resistance	207
10.3.1	Overview	207
10.3.2	DNA sequencing	207
10.3.3	Pre-processing	209
10.3.4	Direct association	211
10.4	Supervised models	211
10.4.1	Logistic regression	211
10.4.2	Support vector machine	212
10.4.3	Random forest	213
10.4.4	Bayesian naive Bayesian (BNB)	213
10.4.5	Supervised classification for antibiotic resistance prediction	215
10.5	Unsupervised models	216
10.5.1	Mixture model	217
10.5.2	Bayesian mixture model	219
10.5.3	Latent feature model	221
10.6	Summary	223
	Acknowledgements	224
	References	225
11	Machine learning for chronic disease	227
	<i>Katherine E. Niehaus and David A. Clifton</i>	
11.1	Introduction	227
11.2	Data	227
11.2.1	EHR data	228
11.2.2	Genomic data	229
11.3	EVT applied to longitudinal data	231
11.3.1	Classical EVT	232
11.3.2	EVT from a point process perspective	234
11.3.3	Practicalities	235
11.3.4	Application of EVT models to healthcare	236
11.3.5	Advanced topics	238
11.3.6	Conclusions on EVT	239
11.4	Patient clustering	239
11.4.1	Clustering overview	240
11.4.2	Modelling choices applicable to chronic disease applications	242
11.4.3	Clustering extensions	243
11.4.4	Practical considerations in unsupervised clustering	243
11.4.5	Clustering conclusions	245
11.5	Conclusion	246
	References	246

12 Big data and optimisation of treatment strategies	251
<i>Shamim Nemati and Mohammad M. Ghassemi</i>	
12.1 Introduction	251
12.2 Heparin dosing as an illustrative example	252
12.2.1 Medication dosing as a classification problem	254
12.2.2 Medication dosing as a prediction problem	260
12.2.3 Medication dosing as a sequential decision-making problem	263
12.3 Discussion	266
References	268
 13 Decision support systems for home monitoring applications: Classification of activities of daily living and epileptic seizures	 271
<i>Stijn Luca, Lode Vuegen, Hugo Van hamme, Peter Karsmakers and Bart Vanrumste</i>	
13.1 Introduction and overview	271
13.2 Supervised classification	272
13.2.1 Gaussian mixture models for classification	273
13.2.2 Support Vector Machines	274
13.2.3 Classification of activities of daily living	279
13.3 Novelty detection	282
13.3.1 One-class support vector machines	282
13.3.2 Extreme value theory	284
13.3.3 Epileptic seizure detection	287
13.4 Conclusion	290
References	290
 Index	 293