

*Basilio de Bragança Pereira
Calyampudi Radhakrishna Rao
Fábio Borges de Oliveira*

***Statistical Learning Using
Neural Networks: A Guide for
Statisticians and Data Scientists***



To our families.



Contents

Preface	xi
Acknowledgment	xiii
1 Introduction	1
2 Fundamental Concepts on Neural Networks	3
2.1 Artificial Intelligence: Symbolist & Connectionist	3
2.2 The Brain and Neural Networks	4
2.3 Artificial Neural Networks and Diagrams	5
2.4 Activation Functions	6
2.5 Network Architectures	7
2.6 Network Training	7
2.7 Kolmogorov Theorem	8
2.8 Model Choice	10
2.8.1 Generalization	10
2.8.2 Bias-variance Trade-off: Early Stopping Method of Training	10
2.8.3 Choice of Structure	11
2.8.4 Network Growing	11
2.8.5 Network Pruning	11
2.9 Mcculloch-pitt Neuron	13
2.10 Rosenblatt Perceptron	13
2.11 Widrow's Adaline and Madaline	15
2.12 Terminology	16
2.13 Running Python in a Nutshell	20
3 Some Common Neural Network Models	35
3.1 Multilayer Feedforward Networks	35
3.2 Associative and Hopfield Networks	39
3.3 Radial Basis Function Networks	46
3.4 Wavelet Neural Networks	50
3.4.1 Wavelets	50
3.4.2 Wavelet Networks and Radial Basis Wavelet Networks.	53
3.5 Mixture-of-Experts Networks	56
3.6 Neural Network and Statistical Models Interface	56
3.7 Some Common Neural Network in Python	58
3.7.1 Fitting Data	58

3.7.2	Classification	61
3.7.3	Hopfield Networks	64
4	Multivariate Statistics Neural Network Models	71
4.1	Cluster and Scaling Networks	71
4.1.1	Competitive networks	71
4.1.2	Learning Vector Quantization (LVQ)	74
4.1.3	Adaptive Resonance Theory (ART) Networks	75
4.1.4	Self-Organizing Maps (SOM) Networks	82
4.2	Dimensional Reduction Networks	92
4.2.1	Basic Structure of the Data Matrix	93
4.2.2	Mechanics of Some Dimensional Reduction Techniques	95
4.2.2.1	Principal Components Analysis (PCA)	95
4.2.2.2	Non-linear Principal Components	95
4.2.2.3	Factor Analysis - FA	96
4.2.2.4	Correspondence Analysis - CA	96
4.2.2.5	Multidimensional Scaling	97
4.2.2.6	Independent Component Analysis (ICA)	98
4.2.3	PCA Networks	101
4.2.4	Non-linear PCA Networks	106
4.2.5	FA Networks	106
4.2.6	Correspondence Analysis (CA) Networks	109
4.2.7	Independent Component Analysis (ICA) Networks	109
4.3	Classification Networks	112
4.4	Multivariate Statistics Neural Network Models with Python	123
4.4.1	Clustering	124
4.4.2	Fitting Data	132
5	Regression Neural Network Models	135
5.1	Generalized Linear Model Networks (GLIMN)	135
5.1.1	Logistic Regression Networks	136
5.1.2	Regression Network	140
5.2	Nonparametric Regression and Classification Networks	143
5.2.1	Probabilistic Neural Networks (PNN)	143
5.2.2	General Regression Neural Networks (GRNN)	144
5.2.3	Generalized Additive Model Networks	145
5.2.4	Regression and Classification Tree Network	147
5.2.5	Projection Pursuit and Feed-Forward Networks	149
5.2.6	Example	150
5.3	Regression Neural Network Models with Python	151
6	Survival Analysis and other Networks	155
6.1	Survival Analysis Networks	155
6.2	Time Series Forecasting	162
6.2.1	Forecasting with Neural Networks	167

<i>Contents</i>	ix
6.3 Control Chart Networks	169
6.4 Some Statistical Inference Results	175
6.4.1 Estimation Methods	175
6.4.2 Bayesian Methods	176
6.4.3 Interval Estimation	179
6.4.4 Statistical Tests	180
6.5 Forecasting with Python	181
A Command Reference	187
Bibliography	221



Preface

This book is intended to integrate two data science modeling techniques: algorithmic modeling and data modeling. This book approaches the first technique with neural networks and the second with statistical methods.

The first author's (BBP) interest in neural networks began in view of the parallel relations of neural networks and statistical methods for data analysis. Such interest motivated him to write short lecture notes on the subject for a minicourse at the SINAPE (Brazilian National Symposium in Probability and Statistics) in 1998.

The second author (CRR), in his visit to Brazil in 1999, took notice of the notes and invited BBP to visit the Statistics Department of PSU (Penn State University) for short periods in 1999 and 2000, and finally for a year in 2003, when they wrote a first version of this book as a Technical Report of the Center of Multivariate Analysis at PSU.

The interest in the technical report caught the attention of Chapman & Hall/CRC, which offered to publish the manuscript provided the authors would add the algorithmic and computational approaches, including examples. In 2018, BBP became a courtesy researcher at LNCC (Brazilian National Laboratory for Scientific Computing) and started to collaborate with the third author (FBO), who is a professor at LNCC and has research interests in computing and neural networks, especially in the areas of security and privacy.



Acknowledgment

The authors thank CAPES (Agency of the Brazilian Ministry of Education) for a year of support grant to BBP during his visit at Penn State University and to UFRJ for a leave of absence and FAPERJ (Carlos Chagas Filho Foundation for Research Support of the State of Rio de Janeiro) for a visiting grant to work in this book at LNCC.

The authors thank Claudio Téllez and Matheus Aranha for reviewing the manuscript. The authors also thank Renata and Lawrence Hamtil for proofreading.

The authors are grateful to David Grubbs from Taylor and Francis for being supportive and patient editor.