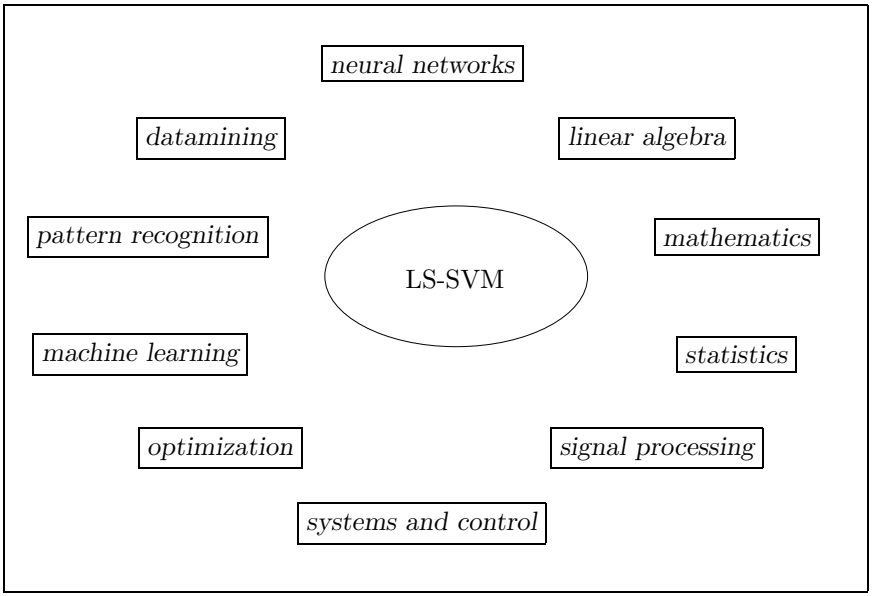
Preface

In recent years there have been many new and exciting developments in kernel based learning, largely stimulated by work in statistical learning theory and support vector machines. Our research in this area started in 1998 around the time that we organized the *International Workshop on Advanced Black-Box Techniques for Nonlinear Modelling* in Leuven, where Vladimir Vapnik presented his important breakthroughs in this area. Solving nonlinear modelling and classification problems by convex optimization without suffering from many local minima sounded indeed very appealing and interesting for a deeper investigation.

Driven by the dream to make the approach as simple as possible (but not simpler) led us to the formulation of least squares support vector machine classifiers, as a first contribution in this area. Many tests and comparisons showed great performance of LS-SVMs on several benchmark data set problems and were very encouraging for further research in this promising direction. At the ESAT-SISTA research division of the Electrical Engineering department of the Katholieke Universiteit Leuven a lot of expertise in the area of mathematical engineering, including neural networks, was available which largely motivated the study of least squares support vector machines from this perspective. Conceptually, the additional explicit *primal-dual* interpretations from the viewpoint of optimization theory turned out to be essential for further developments of least squares support vector machines. In the neural networks area the emphasis has always been on *universal models* with applications within a very broad context such as function estimation, recurrent modelling, classification, control, unsupervised learning, on-line learning and many more. Links between LS-SVMs and regulariza-



LS-SVM: an interdisciplinary topic.

tion networks, Gaussian processes and kernel Fisher discriminant analysis became clear. A Bayesian learning and a robust statistics framework were developed. Support vector machine formulations to principal component analysis and canonical correlation analysis and their kernel versions were made. Towards large scale problems the primal-dual formulations have been exploited in a fixed size LS-SVM algorithm. It turned out that several extensions of the SVM methodology are much easier to formulate in terms of least squares and equality constraints instead of other loss functions or inequality constraints. Therefore, one of the important motivations for this book is to present a *general framework* (in the sense of traditional neural nets) for a class of support vector machines towards supervised and unsupervised learning and feedforward as well as recurrent networks.

The topic of least squares support vector machines is also very interdisciplinary. The world in which we are living is characterized by fragmentation, but at the same time the emergence of several new technologies requires different fields to interact with each other. Given these two sometimes conflicting faces of reality, it is important to have common languages for

transferring ideas between different fields and translating them into novel applications. Therefore another motivation for this book is to offer an *interdisciplinary forum* where different fields can meet, ranging from neural networks, machine learning, mathematics, statistics, optimization, pattern recognition, signal processing, circuits-, systems- and control theory to many applications areas. At this point we were stimulated by invitations and organizations of special sessions at international conferences and the opportunities for invited talks, tutorials and mini-courses such as for *IJCNN, ESANN, ECCTD, ISCAS, AMS, IMTC, ECC, FoCM, ICRM*. In parallel with the writing of this book, a *NATO Advanced Study Institute on Learning Theory and Practice* was organized in Leuven July 2002, which has been particularly motivating as well.

More specifically, for stimulating discussions, invitations and organization of joint meetings, we would like to thank Peter Bartlett, Sankar Basu, Jan Beirlant, Colin Campbell, Vladimir Cherkassky, Nello Cristianini, Felipe Cucker, Mark Embrechts, Lee Feldkamp, Martin Hasler, Simon Haykin, Gabor Horvath, Sathiya Keerthi, Anthony Kuh, Chih-Jen Lin, Lennart Ljung, Charlie Micchelli, Tommy Poggio, Massi Pontil, Danil Prokhorov, Johan Schoukens, Steve Smale, Stefan Vandewalle, Paul Van Dooren, Vladimir Vapnik, Mathukumalli Vidyasagar, Grace Wahba, Paul Werbos, Yu Yi, and many others.

All of us also highly appreciated the great efforts made by Bernhard Schölkopf and Alex Smola for setting up a website on kernel machines www.kernel-machines.org in recent years. For least squares support vector machines a Matlab/C toolbox called LS-SVMlab is available at

http : //www.esat.kuleuven.ac.be/sista/lssvmlab/.

We are especially grateful to Kristiaan Pelckmans for this development and to his colleagues Lukas, Bart Hamers and former master student Emmanuel Lambert.

Besides theoretical and algorithmical contributions on least squares support vector machines, several applications studies have been made with joint projects at K.U. Leuven. In the area of *bio-informatics* with applications to microarray data and textmining, we are grateful to Peter Antal, Tijl Debie, Frank De Smet, Patrick Glenisson, Bart Hamers, Kathleen Marchal, Janick Mathys, Yves Moreau and Gert Thijs. For *biomedical applications* we enjoyed our cooperation with Sabine Van Huffel in the projects

on classification of brain tumours from magnetic resonance spectroscopy signals (with Andy Devos, Lukas, Rene Intzand, Leentje Vanhamme, and Rosemary Tate (University of Sussex)), detection of ovarian cancer (with Chuan Lu, and Dirk Timmerman, Ignace Vergote (K.U. Leuven Hospitals)) and prediction of mental development of preterm newborns (with Lieveke Ameye, and Hans Daniels, Gunnar Naulaers, Hugo Devlieger (K.U. Leuven Hospitals)). In the area of *nonlinear system identification and control* we are grateful to Luc Hoegaerts, Jeroen Buijs, several master students and Jakobus Barnard, Chris Aldrich (University of Stellenbosh) for our cooperation on the problem of prediction of air pollution. For *benchmarking studies and marketing applications* we want to thank Bart Baesens, Stijn Viaene, Jan Vanthienen, Guido Dedene (K.U. Leuven Applied Economic Sciences). For studies in *financial engineering* we are grateful to Dirk Baestaens (Fortis Bank Brussels) and several master students, in particular also Gert Lanckriet for contributions to the Bayesian LS-SVM framework.

Furthermore we are grateful to our many colleagues at other universities within the interuniversity poles of attraction in Belgium, the interdisciplinary center of neural networks of the K.U. Leuven, all the members of our (continuously growing) research group ESAT-SCD-SISTA and the interaction with its spin-off companies. Thanks to all of you for the great atmosphere!

Neural networks have often been presented as a universal solution to many real-life problems (sometimes even as a "miracle solution") and one has come in a stage now where it is important to understand the limits of intelligence, both artificial and human as stated by Steve Smale in his *mathematical problems for the next century* [218] (Problem 18). We are very grateful to him and others as Tommy Poggio and Felipe Cucker for invitations to special meetings on learning theory and hope that this book, together with several breakthroughs in this area, may further contribute towards understanding these exciting problems.

Johan Suykens
Tony Van Gestel
Jos De Brabanter
Bart De Moor
Joos Vandewalle
Leuven, June 2002