Preface

In the last few years there have been very significant developments in the theoretical understanding of Support Vector Machines (SVMs) as well as algorithmic strategies for implementing them, and applications of the approach to practical problems. We believe that the topic has reached the point at which it should perhaps be viewed as its own subfield of machine learning, a subfield which promises much in both theoretical insights and practical usefulness. Despite reaching this stage of development, we were aware that no organic integrated introduction to the subject had yet been attempted. Presenting a comprehensive introduction to SVMs requires the synthesis of a surprisingly wide range of material, including dual representations, feature spaces, learning theory, optimisation theory, and algorithmics. Though active research is still being pursued in all of these areas, there are stable foundations in each that together form the basis for the SVM concept. By building from those stable foundations, this book attempts a measured and accessible introduction to the subject of Support Vector Machines.

The book is intended for machine learning students and practitioners who want a gentle but rigorous introduction to this new class of learning systems. It is organised as a textbook that can be used either as a central text for a course on SVMs, or as an additional text in a neural networks, machine learning, or pattern recognition class. Despite its organisation as a textbook, we have kept the presentation self-contained to ensure that it is suitable for the interested scientific reader not necessarily working directly in machine learning or computer science. In this way the book should give readers from other scientific disciplines a practical introduction to Support Vector Machines enabling them to apply the approach to problems from their own domain. We have attempted to provide the reader with a route map through the rigorous derivation of the material. For this reason we have only included proofs or proof sketches where they are accessible and where we feel that they enhance the understanding of the main ideas. Readers who are interested in the detailed proofs of the quoted results are referred to the original articles.

Exercises are provided at the end of the chapters, as well as pointers to relevant literature and on-line software and articles. Given the potential instability of on-line material, in some cases the book points to a dedicated website, where the relevant links will be kept updated, hence ensuring that readers can continue to

access on-line software and articles. We have always endeavoured to make clear who is responsible for the material even if the pointer to it is an indirect one. We hope that authors will not be offended by these occasional indirect pointers to their work. Each chapter finishes with a section entitled Further Reading and Advanced Topics, which fulfils two functions. First by moving all the references into this section we have kept the main text as uncluttered as possible. Again we ask for the indulgence of those who have contributed to this field when we quote their work but delay giving a reference until this section. Secondly, the section is intended to provide a starting point for readers who wish to delve further into the topics covered in that chapter. The references will also be held and kept up to date on the website. A further motivation for moving the references out of the main body of text is the fact that the field has now reached a stage of maturity which justifies our unified presentation. The two exceptions we have made to this rule are firstly for theorems which are generally known by the name of the original author such as Mercer's theorem, and secondly in Chapter 8 which describes specific experiments reported in the research literature.

The fundamental principle that guided the writing of the book is that it should be accessible to students and practitioners who would prefer to avoid complicated proofs and definitions on their way to using SVMs. We believe that by developing the material in intuitively appealing but rigorous stages, in fact SVMs appear as simple and natural systems. Where possible we first introduce concepts in a simple example, only then showing how they are used in more complex cases. The book is self-contained, with an appendix providing any necessary mathematical tools beyond basic linear algebra and probability. This makes it suitable for a very interdisciplinary audience.

Much of the material was presented in five hours of tutorials on SVMs and large margin generalisation held at the University of California at Santa Cruz during 1999, and most of the feedback received from these was incorporated into the book. Part of this book was written while Nello was visiting the University of California at Santa Cruz, a wonderful place to work thanks to both his hosts and the environment of the campus. During the writing of the book, Nello made frequent and long visits to Royal Holloway, University of London. Nello would like to thank Lynda and her family for hosting him during these visits. Together with John he would also like to thank Alex Gammerman, the technical and administrative staff, and academic colleagues of the Department of Computer Science at Royal Holloway for providing a supportive and relaxed working environment, allowing them the opportunity to concentrate on the writing.

Many people have contributed to the shape and substance of the book, both indirectly through discussions and directly through comments on early versions of the manuscript. We would like to thank Kristin Bennett, Colin Campbell, Nicolo Cesa-Bianchi, David Haussler, Ralf Herbrich, Ulrich Kockelkorn, John Platt, Tomaso Poggio, Bernhard Schölkopf, Alex Smola, Chris Watkins, Manfred Warmuth, Chris Williams, and Bob Williamson.

We would also like to thank David Tranah and Cambridge University Press for being so supportive and helpful in the processing of the book. Alessio Cristianini assisted in the establishment of the website. Kostantinos Veropoulos

helped to create the pictures for Chapter 6 which were generated using his software package at the University of Bristol. We would like to thank John Platt for providing the SMO pseudocode included in Appendix A.

Nello would like to thank the EPSRC for supporting his research and Colin Campbell for being a very understanding and helpful supervisor. John would like to thank the European Commission for support through the NeuroCOLT2 Working Group, EP27150.

Since the first edition appeared a small number of errors have been brought to our attention, and we have endeavoured to ensure that they were all corrected before reprinting. We would be grateful if anyone discovering further problems contact us through the feedback facility on the book's web page **www.support-vector.net**.

Nello Cristianini and John Shawe-Taylor
June, 2000