Chapter V
Support Vector Machines for Business Applications

Brian C. Lovell
NICTA & The University of Queensland, Australia

Christian J. Walder
Max Planck Institute for Biological Cybernetics, Germany

ABSTRACT

This chapter discusses the use of support vector machines (SVM) for business applications. It provides a brief historical background on inductive learning and pattern recognition, and then an intuitive motivation for SVM methods. The method is compared to other approaches, and the tools and background theory required to successfully apply SVM to business applications are introduced. The authors hope that the chapter will help practitioners to understand when the SVM should be the method of choice, as well as how to achieve good results in minimal time.

INTRODUCTION

Recent years have seen an explosive growth in computing power and data storage within business organisations. From a business perspective, this means that most companies now have massive archives of customer and product data and more often than not these archives are far too large for human analysis. An obvious question has therefore arisen, "How can one turn these immense corporate data archives to commercial advantage?" To this end, a number of common applications have arisen, from predicting which products a customer is most likely to purchase, to designing the perfect product based on responses to questionnaires. The theory and development of these processes has grown into a discipline of its own, known as Data Mining, which draws heavily on the related fields of Machine Learning, Pattern Recognition, and Mathematical Statistics.

The Data Mining discipline is still developing, however, and a great deal of sub-optimal and ad hoc analysis is being done. This is partly due to the complexity of the problems, but is also due to the vast number of available techniques. Even the most fundamental task in Data Mining, that of inductive inference, or making predictions based on examples, can be tackled by a great many different techniques. Some of these techniques are very difficult to tailor to a specific problem and require highly skilled human design. Others are more generic in application and can be treated more like the proverbial "black box." One particularly generic and powerful method, known as the Support Vector Machine (SVM) has proven to be both easy to apply and capable of producing results that range from good to excellent in comparison to other methods. While application of the method is relatively straightforward, the practitioner can still benefit greatly from a basic understanding of the underlying machinery.

Unfortunately most available tutorials on SVMs require a very solid mathematical background, so we have written this chapter to make SVM accessible to a wider community. This chapter comprises a basic background on the problem of induction, followed by the main sections. In the first section we introduce the concepts and equations on which the SVM is based, in an intuitive manner, and identify the relationship between the SVM and some of the other popular analysis methods. In the second section we survey some interesting applications of SVMs on practical real-world problems. Finally, the third section provides a set of guidelines and rules of thumb for applying the tool, with a pedagogical example that is designed to demonstrate everything that the SVM newcomer requires in order to immediately apply the tool to a specific problem domain. The chapter is intended as a brief introduction to the field that introduces the ideas, methodologies, as well as a hands-on introduction to freely available software, allowing the reader to rapidly determine the effectiveness of SVMs for their specific domain.

BACKGROUND

SVMs are most commonly applied to the problem of inductive inference, or making predictions based on previously seen examples. To illustrate what is meant by this, let us consider the data presented in Tables 1 and 2. We see here an example of the problem of inductive inference, more specifically that of supervised learning. In supervised learning we are given a set of input data along with their corresponding labels. The input data comprises a number of examples about which several attributes are known (in this case, age, income, etc.). The label indicates which class a particular example belongs to. In the example above, the label tells us whether or not a given person has a broadband Internet connection to their home. This is called a binary classification problem because there are only two possible classes. In the second table, we are given the attributes for a different set of consumers, for whom the true class labels are unknown. Our goal is to infer from the first table the most likely labels for the people in the second table, that is, whether or not they have a broadband Internet connection to their home.

In the field of data mining, we often refer to these sets by the terms test set, training set, validation set, and so on, but there is some confusion in the literature about the exact definitions of these terms. For this reason we avoid this nomenclature, with the exception of the term training set. For our purposes, the training set shall be all that is given to us in order to infer some general correspondence between the input data and labels. We will refer to the set of data for which we would like to predict the labels as the unlabelled set.

17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/support-vector-machines-business-applications/26134

Related Content

African Americans and Planned Resilience: In Search of Ordinary Magic

Hansel Burley, Lucy Barnard-Brak, Valerie McGaha-Garnett, Bolanle A. Olaniranand Aretha Marbley (2012). *Cases on Institutional Research Systems (pp. 305-316).*

www.irma-international.org/chapter/african-americans-planned-resilience/60856

Using "Blackbox" Algorithms Such AS TreeNET and Random Forests for Data-Ming and for Finding Meaningful Patterns, Relationships and Outliers in Complex Ecological Data: An Overview, an Example Using G

Erica Craigand Falk Huettmann (2009). *Intelligent Data Analysis: Developing New Methodologies Through Pattern Discovery and Recovery (pp. 65-84).*

www.irma-international.org/chapter/using-blackbox-algorithms-such-treenet/24212

Mining Clinical Trial Data

Jose Ma. J. Alvirand Javier Cabrera (2007). *Knowledge Discovery and Data Mining: Challenges and Realities (pp. 31-47).*

www.irma-international.org/chapter/mining-clinical-trial-data/24900

Vector DNF for Datasets Classifications: Application to the Financial Timing Decision Problem

Massimo Liquoriand Andrea Scozzari (2008). *Mathematical Methods for Knowledge Discovery and Data Mining (pp. 24-40).*

www.irma-international.org/chapter/vector-dnf-datasets-classifications/26131

Conceptual Data Warehouse Design Methodology for Business Process Intelligence

Svetlana Mansmann, Thomas Neumuth, Oliver Burgertand Matthias Röger (2010). *Complex Data Warehousing and Knowledge Discovery for Advanced Retrieval Development: Innovative Methods and Applications (pp. 129-173).*

www.irma-international.org/chapter/conceptual-data-warehouse-design-methodology/39591