Applications of Kernel Methods

Gustavo Camps-Valls
Universitat de València, Spain

Manel Martínez-Ramón
Universidad Carlos III de Madrid, Spain

José Luis Rojo-Álvarez
Universidad Rey Juan Carlos, Spain

INTRODUCTION

In this chapter, we give a survey of applications of the kernel methods introduced in the previous chapter. We focus on different application domains that are particularly active in both direct application of well-known kernel methods, and in new algorithmic developments suited to a particular problem. In particular, we consider the following application fields: biomedical engineering (comprising both biological signal processing and bioinformatics), communications, signal, speech and image processing.

KERNEL METHODS IN BIOMEDICINE AND BIOINFORMATICS

Kernel methods have been extensively used to solve biomedical problems. For instance, a study of prediction of cyclosporine dosage in patients after kidney transplantation using neural networks and kernel-based methods was carried out in (Camps-Valls et al., 2002). Recently, (Osowski, Hoai, & Markiewicz, 2004) proposed a committee of experts formed by several support vector machines (SVM) for the recognition of 13 heart rhythm types.

The most impressive results of kernels have been obtained in genomics and computational biology, due to both the special characteristics of data and the great interest in solving biological problems since the Human genome sequencing. Their ability to work with high dimensional data, to process and efficiently integrate non-vectorial string data, make them very suitable to solve various problems arising in computational biology. Since the early papers using SVM in bioinformatics (Mukherjee et al., 1998), the applications of these methods have grown exponentially, and many novel and powerful methods have been developed (only in 2004, more than 1000 papers have been devoted to this topic). The use of kernel methods in computational biology has been accompanied by new developments to match the specificities and the needs of the field, such as methods for *feature selection* in combination with the classification of high-dimensional data, the introduction of *string kernels* to process biological sequences, or the development of methods to *learn from several kernels simultaneously* ('composite kernels'). The interested reader can find a comprehensive introduction in (Vert, 2006).

KERNEL METHODS IN COMMUNICATIONS

There are four situations that make kernel methods good candidates for use in electromagnetics (Martínez-Ramon, 2006): 1) No close solutions exist, and the only approaches are trial and error methods. In these cases, kernel algorithms can be employed to solve the problem. 2) The application requires operating in real time, and the computation time is limited. In these cases, a kernel algorithm can be trained off-line, and used in test mode in real time. The algorithms can be embedded in any hardware device. 3) Faster convergence rates and smaller errors are required. Kernel algorithms have shown superior performance in generalization ability in many problems. Also, the block optimization and the uniqueness of solutions make kernelized versions of linear algorithms (as SVM) faster than many other methods. 4) Enough measured data exist to train a regression algorithm for prediction and no analytical tools exist. In this case, one can actually use an SVM to solve the part of the problem where no analytical solution exist and combine the solution with other existing analytical and closed form solutions.

The use of kernelized SVMs has been already proposed to solve a variety of digital communications problems. The decision feedback equalizer (Sebald & Buclew, 2000) and the adaptive multi-user detector for Code Division Multiple Access (CDMA) signals in multipath channels (Chen et al., 2001) are addressed by means of binary SVM nonlinear classifiers. In (Rahman et al., 2004) signal equalization and detection for a MultiCarrier (MC)-CDMA system is based on an SVM linear classification algorithm. Koutsogiannis et al. (2002) introduced the use of KPCA for classification and de-noising of communication signals.

KERNEL METHODS IN SIGNAL PROCESSING

Many signal processing supervised and unsupervised schemes such as discriminant analysis, clustering, principal/independent component analysis, or mutual information extraction have been addressed using kernels (see previous chapters). Also, an interesting perspective for signal processing using SVM can be found in (Mattera, 2005), which relies on a different point of view to signal processing.

The use of time series with supervised SVM algorithms has mainly focused on two DSP problems: (1) non-linear system identification of the underlying relationship between two simultaneously recorded discrete-time processes, and (2) time series prediction (Drezet and Harrison 1998; Gretton et al., 2001; Suykens, 2001). In both of them, the conventional SVR considers lagged and buffered samples of the available signals as its input vectors.

These approaches pose several problems and opportunities. First, the statement of linear signal models in the primal problem, which will be called *SVM primal signal models*, will allow us to obtain robust estimators of the model coefficients (Rojo-Álvarez et al., 2005a) in classical DSP problems, such as ARMA modeling, the γ-filter, and spectral analysis (Rojo-Álvarez et al., 2003, Camps-Valls et al., 2004, Rojo-Álvarez et al., 2004). Second, the consideration of nonlinear SVM-DSP algorithms can be addressed from two different approaches: (1) *RKHS signal models*, which state the signal model equation in the feature space (Martínez-Ramón et al., 2005), and (2) *dual signal models*, which are based on the nonlinear regression of each single time instant with appropriate Mercer's kernels (Rojo-Álvarez et al., 2005b).

KERNEL METHODS IN SPEECH PROCESSING

An interesting and active research field is that of speech recognition and speaker verification. First, there have been many attempts to apply SVMs to improve existing speech recognition systems. Ganapathiraju (2002) uses SVMs to estimate Hidden Markov Models state likelihoods, Venkataramani et al. (2003) applied SVMs to refine the decoding search space, and in (Gales and Layton, 2004) statistical models for large vocabulary continuous speech recognition were trained using SVMs. Second, early SVM approaches by Schmidt and Gish (1996), and then by Wan and Campbell (2000), used polynomial and RBF kernels to model the distribution of cepstral input vectors. Further improvements considered mapping to *feature space* using sequence kernels (Fine et al. 2001). In the case of speaker verification, the recent works of Shriberg et al. (2005) for processing high-level stylistic or lexical features are worth mentioning.

Voice processing has been performed by using KPCA. Lima et al. (2005) used sparse KPCA for voice feature extraction and then used them for speech recognition. Mak et al. (2005) used KPCA to introduce speaker adaptation in voice recognition schemes.

KERNEL METHODS IN IMAGE PROCESSING

One of the first works proposing kernel methods in the context of image processing was (Osuna *et al.,* 1997), where a face detection system was proposed. Also, in (Papagiorgiou & Poggio, 2000) a face, pedestrian, and car detection method based on SVMs and Haar wavelets to represent images was presented.

The previous global approaches demonstrated good results for detecting objects under fixed viewing conditions. However, problems occur when the viewpoint and pose vary. Different methods have been built to

Related Content

Projected Clustering for Biological Data Analysis
Ping Deng, Qingkai Maand Weili Wu (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 1617-1622).*
www.irma-international.org/chapter/projected-clustering-biological-data-analysis/11035

Clustering of Time Series Data
Anne Denton (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 258-263).*
www.irma-international.org/chapter/clustering-time-series-data/10830

Program Comprehension through Data Mining
Ioannis N. Kouris (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 1603-1609).*
www.irma-international.org/chapter/program-comprehension-through-data-mining/11033

Clustering Data in Peer-to-Peer Systems
Mei Liand Wang-Chien Lee (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 251-257).*
www.irma-international.org/chapter/clustering-data-peer-peer-systems/10829

Process Mining to Analyze the Behaviour of Specific Users
Laura Maruster (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 1589-1597).*
www.irma-international.org/chapter/process-mining-analyze-behaviour-specific/11031