

Chapter 13

A Review of Kernel Methods Based Approaches to Classification and Clustering of Sequential Patterns, Part I: Sequences of Continuous Feature Vectors

A. D. Dileep

Indian Institute of Technology Madras, India

T. Veena

Indian Institute of Technology Madras, India

C. Chandra Sekhar

Indian Institute of Technology Madras, India

ABSTRACT

Sequential data mining involves analysis of sequential patterns of varying length. Sequential pattern analysis is important for pattern discovery from sequences of discrete symbols as in bioinformatics and text analysis, and from sequences or sets of continuous valued feature vectors as in processing of audio, speech, music, image, and video data. Pattern analysis techniques using kernel methods have been explored for static patterns as well as sequential patterns. The main issue in sequential pattern analysis using kernel methods is the design of a suitable kernel for sequential patterns of varying length. Kernel functions designed for sequential patterns are known as dynamic kernels. In this chapter, we present a brief description of kernel methods for pattern classification and clustering. Then we describe dynamic kernels for sequences of continuous feature vectors. We then present a review of approaches to sequential pattern classification and clustering using dynamic kernels.

DOI: 10.4018/978-1-4666-2455-9.ch013

INTRODUCTION TO SEQUENTIAL PATTERN ANALYSIS USING KERNEL METHODS

Classification and clustering of patterns extracted from sequential data are important for pattern discovery using sequence data mining. Pattern discovery from bio-sequences involves classification and clustering of discrete symbol sequences. Pattern discovery from multimedia data such as audio, speech and video data involves classification and clustering of continuous valued feature vector sequences. Classification and clustering of sequential patterns of varying length have been challenging tasks in pattern recognition. Conventional methods for classification of sequential patterns use discrete hidden Markov models (HMMs) for discrete sequences, and continuous density HMMs for continuous feature vector sequences. Conventional methods for clustering of sequential patterns use distance measures such as edit distance for discrete sequences and dynamic time warping based distance for continuous feature vector sequences. During the past 15 years, kernel methods based approaches such as support vector machines and kernel K -means clustering have been explored for classification and clustering of static patterns and sequential patterns. Kernel methods have been shown to give a good generalization performance. This Chapter presents a review of kernel methods based approaches to classification and clustering of sequential patterns.

Kernel methods for pattern analysis involve performing a nonlinear transformation from the input feature space to a higher dimensional feature space induced by a Mercer kernel function, and then constructing an optimal linear solution in the kernel feature space. Support vector machine for two class pattern classification constructs an optimal hyperplane corresponding to the maximum margin separating hyperplane in the kernel feature space (Burges, 1998). Kernel K -means clustering gives an optimal nonlinear separation of clusters in the input feature space by minimizing the trace

of the within-cluster scatter matrix for the clusters formed in the kernel feature space (Girolami, 2002; Satish, 2005). The choice of the kernel function used in the kernel methods is important for their performance. Several kernel functions have been proposed for static patterns. Kernel methods for sequential pattern analysis adopt one of the following two strategies: (1) Convert a sequential pattern into a static pattern and then use a kernel function defined for static patterns, and (2) Design and use a kernel function for sequential patterns. Kernel functions designed for sequential data are referred to as dynamic kernels or sequence kernels (Wan & Renals, 2002). Examples of dynamic kernels for continuous feature vector sequences are Gaussian mixture model (GMM) supervector kernel (Campbell et al., 2006b) and intermediate matching kernel (Boughorbel et al., 2005). Fisher kernel (Jaakkola et al., 2000) is used for both the discrete observation symbol sequences and sequences of continuous feature vectors. This Chapter discusses the issues in designing dynamic kernels for continuous feature vector sequences and then presents a review of dynamic kernels proposed in the literature.

Dynamic kernels for continuous feature vector sequences belong to the following two main categories: (1) Kernels such as Fisher kernels that capture the sequence information in the feature vector sequences, and (2) Kernels such as GMM supervector kernels and intermediate matching kernels that consider the feature vector sequences as sets of feature vectors. The kernels belonging to the first category have been explored for classification of units of speech such as phonemes, syllables and words in speech recognition. The kernels belonging to the second category have been explored for tasks such as speaker identification and verification, speech emotion recognition and image classification. This chapter presents a review of dynamic kernels based approaches to classification and clustering of sequential patterns.

The organization of the rest of the chapter is as follows: The next section describes the kernel

26 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/review-kernel-methods-based-approaches/73443

Related Content

Computational Intelligence Clustering for Dynamic Video Watermarking

Daw-Tung Lin and Guan-Jhih Liao (2010). *Intelligent Soft Computation and Evolving Data Mining: Integrating Advanced Technologies* (pp. 298-318).

www.irma-international.org/chapter/computational-intelligence-clustering-dynamic-video/42366

Predictive Modeling Versus Regression

Patricia Cerrito (2010). *Text Mining Techniques for Healthcare Provider Quality Determination: Methods for Rank Comparisons* (pp. 110-152).

www.irma-international.org/chapter/predictive-modeling-versus-regression/36635

Machine Learning in Studying the Organism's Functional State of Clinically Healthy Individuals Depending on Their Immune Reactivity

Tatiana V. Sambukova (2013). *Data Mining: Concepts, Methodologies, Tools, and Applications* (pp. 893-919).

www.irma-international.org/chapter/machine-learning-studying-organism-functional/73476

Developing a Competitive City through Healthy Decision-Making

Ori Gudes, Elizabeth Kendall, Tan Yigitcanlar, Jung Hoon Han and Virendra Pathak (2013). *Data Mining: Concepts, Methodologies, Tools, and Applications* (pp. 1545-1558).

www.irma-international.org/chapter/developing-competitive-city-through-healthy/73511

Semantics-Aware Advanced OLAP Visualization of Multidimensional Data Cubes

Alfredo Cuzzocrea, Domenico Sacca and Paolo Serafino (2007). *International Journal of Data Warehousing and Mining* (pp. 1-30).

www.irma-international.org/article/semantics-aware-advanced-olap-visualization/1791