Fuzzy Rough Support Vector Machine for Data Classification

Arindam Chaudhuri, Samsung R&D Institute, Noida, India

ABSTRACT

In this paper, classification task is performed by FRSVM. It is variant of FSVM and MFSVM. Fuzzy rough set takes care of sensitiveness of noisy samples and handles impreciseness. The membership function is developed as function of center and radius of each class in feature space. It plays an important role towards sampling the decision surface. The training samples are either linear or nonlinear separable. In nonlinear training samples, input space is mapped into high dimensional feature space to compute separating surface. The different input points make unique contributions to decision surface. The performance of the classifier is assessed in terms of the number of support vectors. The effect of variability in prediction and generalization of FRSVM is examined with respect to values of C. It effectively resolves imbalance and overlapping class problems, normalizes to unseen data and relaxes dependency between features and labels. Experimental results on both synthetic and real datasets support that FRSVM achieves superior performance in reducing outliers’ effects than existing SVMs.

KEYWORDS

Classification, FRSVM, FSVM, Fuzzy Rough Membership Function, MFSVM, SVM

1. INTRODUCTION

Classification of data (Aggarwal 2014), (Duda, Hart & Stork, 2007) is a common task in machine learning. In this direction support vector machines (SVM) (Burges, 1998) has emerged as a promising pattern classification tool in recent years. It is based on the principle of structural risk minimization (SRM) and statistical learning theory (Vapnik, 1998). SVM was proposed by Vapnik (Vapnik, 1998) and has received much attention from the pattern recognition community (Abe, 2010), (Bishop, 2006), (Duda, Hart & Stork, 2007). It has been widely used in various real life applications with appreciable classification performances (Burges, 1998).

Many complex problems have been solved by SVMs (Abe, 2010). Some notable applications where SVM has been successfully applied are handwritten digit recognition, object recognition, speaker identification, charmed quark detection, face detection, optical character recognition, medical diagnostics, text classification etc. (Abe, 2010). Two important applications where SVM has outperformed other methods are electric load prediction (EUNITE, 2001) and optical character recognition (Tautu & Leon, 2012). For regression estimation SVMs have been compared on benchmark time series prediction tests, the Boston housing problem and (on artificial data) on PET operator inversion problem (Abe, 2010), (Burges, 1998). In most of these cases SVM generalization performance i.e. error rates on test sets either matches or is significantly better than that of the competing methods. The use of SVMs for density estimation and ANOVA decomposition has also been studied (Burges, 1998). Regarding extensions the basic SVMs contain no prior knowledge of the problem. For example, a large class of SVMs for image recognition problem gives the same
results if pixels are first permuted randomly with each image suffering the same permutation, an act of vandalism that would leave the best performing neural networks severely handicapped. Although SVMs have good generalization performance they can be abysmally slow in test phase a problem which has been addressed in (Burges, 1998). Several works have generalized the basic ideas of SVM and have shown connections to regularization theory (Abe, 2010), (Burges, 1998). They have also shown how SVM ideas can be incorporated in a wide range of other algorithms (Abe, 2010), (Burges, 1998), (Chaudhuri, De & Chatterjee, 2008), (Chaudhuri & De, 2011), (Chaudhuri, 2014).

SVM is most widely used nonparametric technique (Chaudhuri, De & Chatterjee, 2008), (Chaudhuri & De, 2011), (Chaudhuri, 2014) and yields accurate results. SVM classification exercise finds hyperplane in possible space for maximizing the distance from hyperplane to data points. This is basically solving a quadratic optimization problem. The solution of strictly convex problems for SVM is unique and global. SVM implements SRM that has high generalization performance. By increasing the number of support vectors the computational complexity of the model increases. This is taken care of by constructing SVM through a trade-off where the number of training errors is decreased and the risk of over-fitting data is increased. However, data dependent SRM for SVM does not rigorously support the argument such that good generalization performance of SVM is attributable to SRM (Abe, 2010), (Burges, 1998). Since SVM captures geometric characteristics of feature space without deriving weights of networks from the training data, it is capable of extracting optimal solution with small training set size. SVMs have flexible structure and produce better classification results than parametric methods. They have attractive properties and produce single solution characterized by global minimum of optimized functional and not multiple solutions associated with local minima. SVM do not rely on heuristics and thus are an arbitrary choice to model various problems.

In this work, we study classification problem using fuzzy rough support vector machine (FRSVM). Making a decision to classify a point correctly yields maximum returns (Burges, 1998). Keeping this in view, the classification here is performed by FRSVM which is a variant of fuzzy support vector machine (FSVM) (Chaudhuri & De, 2011) and modified fuzzy support vector machine (MFSVM) (Chaudhuri, 2014). The success in FRSVM lies in importing fuzzy rough membership function (Dubois & Prade, 1992) which is sensitive to noisy mislabeled samples, handles vagueness and gives robust results. The membership function is constructed by considering center and radius of each class in feature space and is represented with hyperbolic tangent kernel. The larger value of membership function represents the importance of sample to decision surface (Abe, 2010), (Burges, 1998). The input points make different contributions to decision surface. FRSVM extracts useful information from data and improves accuracy. However, choice of appropriate parameter values plays a vital role on its performance. The training samples are either linear or nonlinear separable. FRSVM effectively handles nonlinear classification problem. Its performance is also assessed in light of number of support vectors. The effect of variability in prediction and generalization of FRSVM is studied with respect to values of $C$. The experimental results on both synthetic and real datasets illustrate that FRSVM achieves better performance in reducing outliers’ effects than existing SVMs.

This paper is presented as follows. In section 2 some important concepts on SVM are highlighted. The fuzzy rough sets are presented in section 3. This is followed by a discussion on FRSVM in section 4. In next section experimental results and discussions are highlighted. Finally, in section 6 conclusions are illustrated.
Lightweight Key Management for Adaptive Addressing in Next Generation Internet
*International Journal of Ambient Computing and Intelligence* (pp. 50-69).
[www.igi-global.com/article/lightweight-key-management-for-adaptive-addressing-in-next-generation-internet/176713?camid=4v1a](www.igi-global.com/article/lightweight-key-management-for-adaptive-addressing-in-next-generation-internet/176713?camid=4v1a)