One Class SVM Vs SVM Classification

Divya Rana

University School of Information and Communication Technology

Abstract: *One class classification distinguishes the target class from all other classes using only training data from the target class. One class classification is suitable for those situations where outliers are not represented well in the training set. One-class learning, or unsupervised SVM, aims at separating data from the origin in the high-dimensional, predictor space (not the original predictor space), and is an algorithm used for outlier detection. Support vector machine is a machine learning method that is widely used for data analyzing and pattern recognizing. Support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. In this paper we will review the difference between both these classes.*

Keywords: One Class, SVM, outliers, SVDD, Hyperplane

1. Introduction

The previous traditional classification methods were not efficient enough to categorize the data models. In conventional classification, data from two classes are available; the decision boundary is supported from both sides by example objects[1]. Most conventional classifiers assume more or less equally balanced data classes and do not work well when one class is severely under sampled or even completely absent. However, many real world situations are such that they only provide information of only one class, the target class; information of the other class, the outliers, is very difficult to obtain. The problems encountered in the conventional classification problems, such as the definition of the error, atypical training data, measuring the complexity of a solution, the curse of dimensionality, the generalization of the method. In one-class classification only data of one class is available, only one side of the boundary can be determined. It is tough to decide on the basis of just one class how tightly the boundary should fit in each of the directions around the data. It is even tougher to decide which features should be used to find the best separation of the target and outlier class. One-class classification [2] has seen a rise in application over the years, for example, in the use of document classification [3], typist recognition [4] and compliance verification of the CTBT [5].The goal of one class classification is to create a description of one class of objects called target objects and distinguish this class from all other objects, not belonging to this class called outliers. One class classification is particularly suitable where data about outliers is not defined in the training set. There are some applications in which one class is difficult to sample such as detecting credit card fraud, intrusions, or a rare disease. Compared to SVM classification, data of two classes are available. The boundary between the two classes has to be estimated from data of both the classes. Classifying data has been one of the major parts in machine learning. The idea of support vector machine is to create a hyper plane in between data sets to indicate which class it belongs to. The challenge is to train the machine to understand structure from data and mapping with the right class label, for the best result, the hyper plane has the largest distance to the nearest training data points of any class.

2. One Class Classification: Overview

One-class classification has also been known as novelty or outlier detection [6]. Different from normal classification, it samples data from only one class, called the target class, are well characterized, while there are no or few samples from the other class (also called the outlier class). The one-class classification problem differs in one essential aspect from the conventional classification problem. In one-class classification it is assumed that only information of one of the classes, the target class, is available. This means that just example objects of the target class can be used and that no information about the other class of outlier objects is present. The boundary between the two classes has to be estimated from data of only the normal, genuine class. The task is to define a boundary around the target class, such that it accepts as much of the target objects as possible, while it minimizes the chance of accepting outlier objects. A classifier, i.e. a function which outputs a class label for each input object, cannot be constructed from known rules. Therefore, in pattern recognition or machine learning, one tries to infer a classifier from a (limited) set of training examples. The use of examples thus elevates the need to explicitly state the rules for the classification by the user. The goal is to obtain models and learning rules to learn from the examples and predict the labels of future objects.

The goal of the One-Class Classification is to distinguish between a set of target objects and all other possible objects. It is mainly used to detect new objects that resemble a known set of objects. When a new object does not resemble the data, it is likely to be an outlier or a novelty [6]. When it is accepted by the data description, it can be used with higher condence in a subsequent classification Different methods have been developed to make a data description. In most cases the probability density of the target set is modeled. This requires a large number of samples to overcome the curse of dimensionality. Other techniques than estimating a probability density estimate exists. It is possible to use the distance to model or just to estimate the boundary around the class without estimating a probability density.

3. One Class Classification Methods

There are four simple models[7], the support vector data description, k-means clustering, k-center method and an auto-encoder neural network. Here a descriptive model is fitted to the data and the resemblance to this model is used. In the SVDD a hyper sphere is put around the data. By applying the kernel trick the model becomes more flexible to follow the characteristics in the data. Instead of the target density the distance to the center of the hyper sphere is used. In the k-means and k-center method the data is clustered, and the distance to the nearest prototype is used. Finally in the auto-encoder network the network is trained to represent the input pattern at the output layer. The network contains one bottleneck layer to force it to learn a (nonlinear) subspace through the data. The reconstruction error of the object in the output layer is used as distance to the model. The use of a data domain description method is inspired by the support vector machine by, called the Support Vector Domain Description (SVDD)[8]. This method can be used for novelty or outlier detection. A spherically shaped decision boundary around a set of objects is constructed by a set of support vectors describing the sphere boundary. It has the possibility of transforming the data to new feature spaces without much extra computational cost. By using the transformed data, this SVDD can obtain more flexible and more accurate data descriptions. The error of the first kind, the fraction of the training objects which will be rejected, can be estimated immediately from the description without the use of an independent test set, which makes this method data efficient. The Support Vector Domain Description is compared with other outlier detection methods on real data.

4. SVM Classification

A support vector machine [9] constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

Table 1: Difference between One Class and SVM Classification

One Class Classification	SVM Classification
One Class contains data from only one class, target class.	SVM contains data of two or more classes.
Goal is to create a description of one class of objects and distinguish from outliers.	Goal is to create hyperplane with maximum margin between two classes.
Decision boundary is estimated in all directions in the feature space around the target class.	Hyperplane is created in between datasets to indicate which class it belongs to.
Software finds an appropriate bias term such that outlier fraction of the observations in the training set	Software attempts to remove 100*outlier fraction% of observations when the optimizations algorithm converges.
One Class Classifier is traditional classifier.	SVM Classifier is linear classifier.
Used for Outlier detection, Novelty detection	Used for Classification, Regression.
Less parameters, Less training data	More parameters, More training data
Term coined by Moya and Hush	Term coined by Vladimir Vapnik

5. Experimental Framework

I performed one class classification and SVM classification on same dataset and dataset is calculated using BMI that is, Body Mass Index
BMI < 25 → Underweight
BMI > 25 → Obese

Dataset

6. Experimental Results

As one class classification and SVM classification implemented on MATLAB. The result of one class classification is shown in figure (i) and result of SVM classification is shown in figure (ii).
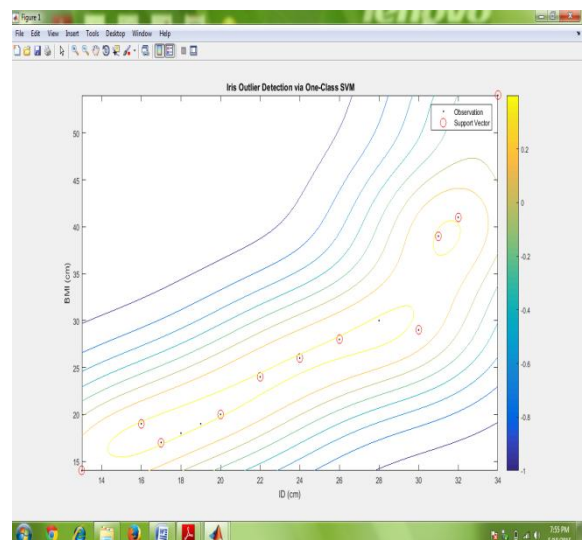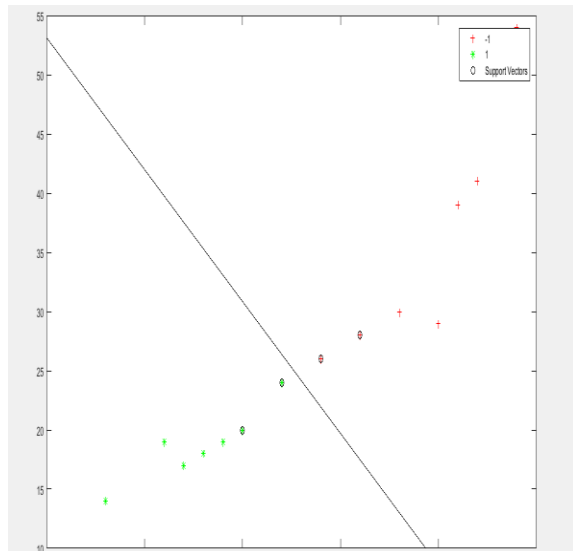

Figure (i)

Paper ID: SUB155491

Figure (ii)

7. Conclusions

I conclude that this paper shows the difference between one class SVM and SVM classification and results show that SVM classification is more efficient than one class classification. Results shows as the number of parameters increases in One Class SVM, the performance decreases of one class classifiers. SVM classifiers, on the other hand, display stable performance and hence SVM classifiers are more appealing.

References

[1] Tax, D., Duin, R.: Uniform object generation for optimizing one-class classifiers. J. Machine Learning Research 2 (2001) 155-173

[2] Tax, D.: One Class Classification. PhD thesis, Delft University of Technology (2001)

[3] Ritter, G., Gallegos, M.: Outliers in statistical pattern recognition and an application to automatic chromosome classification. Pattern Recognition Letters 18 (1997) 525-539

[4] Bishop, C.: Novelty detection and neural network validation. IEEE Proceedings on Vision, Image and Signal Processing. Special Issue on Applications of Neural Networks 141(4) (1994) 217-222

[5] Japkowicz, N.: Concept-Learning in the absence of counterexamples: An autoassociation-based approach to classification. PhD thesis, New Brunswick Rutgers, The State University of New Jersey (1999)

[6] D.M.J. Tax, "One-class classification: Concept-learning in the absence of counter-examples," Delft University of Technology. 2001

[7] Wang, Q., Lopes, L.S., Tax, D.J.: Visual object recognition through one-class learning. In: International Conference on Image Analysis and Recognition. (2004) 463-470

[8] Schoelkopf Support Vector Learning. PhD thesis, Technischen Universitat Berlin, 1997.

[9] Cortes, C.; Vapnik, V. (1995). "Support-Vector Networks". Machine Learning

Paper ID: SUB155491

1352