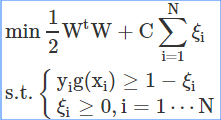
Support vector machines (SVM) is a supervised machine learning classification algorithm, which can be used for classification problem in multi-dimension space.

It is widely used in mining researches, such as text mining and opinion mining, and it has a great result. Comparing to other classification algorithms, SVM algorithm usually has better result in higher dimension, where all number of features is quite large and the data is sparse.

SVM uses   as the linear separation hyperplane, where w is the weight vector, b is the bias,  is a set of high dimensional non-linear transformation function, w and b is determined by training data that optimize the following formulas:

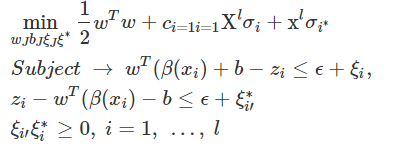


whereis the slack variables, C is the penalty coefficient, for all the training samples ().

Support Vector Regression (SVR) is using the SVM algorithm on regression problem. The goal of SVM is to find the separation hyperplane; and the goal of SVR is to find the regression hyperplane.

For the given training set: {(), …, ()}

where ∈Rn is a feature vector, and ∊ *R1* is the target output. In order to find the hyperplane, two parameters C>0 and ε>0 must be given and the support vector regression can be defined:



In our experiment, we use a free SVM toolkit, sklearn, C = 1to train the SVR model.