The **SVM** is originally developed by Vapnik in 1990s and has been applied in many practical problems since then. The advantages of SVM comes from two good characteristics: the maximized margin and the kernel trick. These good characteristics can provide high testing accuracy of classification and avoid the limits of dimensionality. The SVM solves the optimization problems in below equations:



*K* is the kernel functions used in SVM to solve non-linear classification. The very popular kernel functions include the linear function, polynomial function and the Gaussian radius basis function (RBF). Here we adopts the …

**K-means** is a classical clustering algorithm in the unsupervised machine learning. It partitions a data set into a pre-specified number of distinct, non-overlapping clusters. Here, we solve the problem:

 ,

,

Where the  represents the number of observations in the *k*th cluster.

Compression rate:

