

DMW C2 Assignment-1

Shubham Kumar IIT2018115

Raktim Bijoyपुरi IIT2018125

Aditya Kamble IIT2018126

Aakashdeep IIT2018128

Tejas Mane IIT2018135

VI Semester, Department of Information Technology,

Indian Institute of Information Technology, Allahabad, Prayagraj.

Introduction:

Introducing the comparisons and discussions on the various performance of support vector machine classification algorithms with k-times markov sampling for the sample of large scale training samples and unbalanced training samples. We present support - vector machine classification which is specifically based on k-times Markov sampling.

Algorithm:

SVM Classification Algorithm

SVM is a widely used algorithm in pattern classification problems. SVM classifier classifies the data points into two different classes with a decision boundary or a hyperplane. Minimum distance of a data point from the hyperplane is called the margin and those points which are closest to the hyperplane are called the support vectors. Sometimes, it is easier to find a linear hyperplane to separate the training data points into two different classes but sometimes the data points are distributed in such a way that it becomes very difficult to find a linear hyperplane which separates the data points into two different classes. In such cases soft margin or kernel trick is used. In the case of soft margin, we find a linear hyperplane with minimum number of misclassifications. In the case of kernel trick, we do not find a linear hyperplane, instead, some features are generated from within the dataset and then a non-linear hyperplane is found which separates the data points in two different classes.

Uniformly Ergodic Markov Chains

Markov chain is basically a sequence of Random Variables with some conditions. Suppose space is (Z, X) , which is measurable, random variables $\{z_t\}_{t \geq 0}$ are part of sequence in Markov chain along with a set of probability values $P^n(x, y)$ such that $x \in Z$ and $y \in X$. $P^n(x, y)$ is the value of probability that after n steps of time, the state x will belong to the set y . A markov chain is called a uniform ergodic Markov chain if the following condition holds,

$$\sup \|P^n(z, A) - \pi(A)\|_{TV} = d(n) \rightarrow 0 \text{ as } n \rightarrow \infty$$

Where n is a natural number.

SVMC with K-times Markov Sampling

After performing many experiments results shows that the support vector machine algorithm not have less time of sampling and training ,smaller misclassification rates, but also it provides the obtained classifier is more limited compared with other techniques of classical support vector machine classification(SVMC).

Result -

Kernel	Accuracy
linear	94.44
rbf	93.37
poly	94.95
hellinger	96.42
chi2_kernel	87.42