k-Times Markov Sampling for SVMC

B.Tech (IT) 6th semester, Indian Institute of Information Technology, Allahabad

Deepak Katre (IIT2018116), Anurag Makade (IIT2018121), Dhairya Patel (IIT2018122), Akinchan Lunavat (IIT2018123), Kapil Gupta (IIT2018127)

Summary

Support vector machines (SVM) is one of the most widely used learning algorithms for pattern recognition problems. Besides it's good performance in practical applications, SVM classification (SVMC) also has a good theoretical property in universal consistency and learning rates, if the training samples come from an independent and identically distributed (i.i.d.) process. Since independence is a very restrictive concept, such i.i.d. assumption cannot be strictly validated in real-world problems. In this paper they have introduced Markov chains as opposed to i.i.d. samples. As many machine learning applications, such as market prediction, system diagnosis, and speech recognition, are inherently temporal in nature, and consequently not i.i.d. processes. Therefore, the relaxations of such i.i.d. assumptions for SVMC are considered here.The proposed SVMC algorithm based on k-times Markov sampling is compared with the classical SVMC and presents us with three advantages 1) the misclassification rates are smaller; 2) the total time of sampling and training is less; and 3) the obtained classifiers are more sparse. Also we can find that the differences between the classical SVMC and SVMC with k markov samples algorithm are obvious: the classical SVMC with a given training set S is the "batch learning" of m training samples (m is the size of the training set S). The SVMC with Markov sampling introduced in paper 3 is two times "batch learning" of m training samples (m i.i.d. samples and m Markov chain samples). While Algorithm 1 is k + 1 times "batch learning" of N training samples and the total number of training samples is $(k + 1)N \leq m$.

Results

a) K-Markov Sampling for Pascal dataset

Dataset Link - http://host.robots.ox.ac.uk/pascal/VOC/voc2012/

Kernel	KT_SVM
Linear	0.27
RBF	0.31
Hellinger	0.19
X^2	0.23

b) K-Markov Sampling for Letters dataset

Dataset Link - https://archive.ics.uci.edu/ml/datasets/Letter+Recognition

Kernel	KT_SVM
Linear	0.81
RBF	0.88
Hellinger	0.75
X^2	0.85