DMW Assignment-1

SVM classification using k-times Markov sampling.

***G. Shashank(IIT2018106), Akhil Naik (IIT2018143), Rahul Naidu(IIT2018192),Gurtej Singh(IIT2018193)***

*Semester VI, Department of IT, Indian Institute of Information Technology, Allahabad, India.*

**Introduction**: In the field of machine learning one of the most widely used algorithms is SVM which is used for the classification techniques.SVM stands for the support vector machine.It is difficult to select the kernels for the SVM in classification problem and the complexity of SVM increases when the size of the dataset increases.so in this paper we are going with the approach of K-times Markov sampling.Some experimental results shows that the combination of SVM with the k-times markov sampling can have the less time of sampling and training and can have smaller misclassification rates, but also the obtained classifier is more sparse compared with the classical SVMC and the previously known SVMC algorithm based on Markov sampling.

**Description about the dataset:**

We use Letter Dataset[2], it has 16 different features relating to 26 alphabets to be recognized.First we segment the dataset into a train and test set with 14000 samples for training and 6000 for testing. We use k times-markov sampling (explained next) to choose samples from the training set that forms a markov chain.

**K-times Markov Sampling Algorithm**

Input: ST , N, k, q, n2

Output: sign( fk )

1. Draw randomly N samples Siid := {zj}j=1N from ST. Train Siid by SVMC and obtain a preliminary

learning model f0. Let i = 0.

1. Let N+ = 0, N− = 0, t = 1.
2. Draw randomly a sample zt from ST , called it the current sample. Let N+ = N++1 if the label of zt is +1, or let N− = N− + 1 if the label of zt is −1.
3. Draw randomly another sample z∗ from ST , called it the candidate sample, and calculate the ratio α, α = e−( fi ,z∗)/e−( fi ,zt).
4. If α ≥ 1, yty∗ = 1 accept z∗ with probability α1 = e−y∗ fi /e−yt fi. If α = 1 and yty∗ = −1 or α < 1, accept z∗ with probability α. If there are n2 candidate samples can not be accepted continually, then set α2 = qα and accept z∗ with probability α2. If z∗ is not accepted, go to Step 4, else let zt+1 = z∗, N+ = N+ + 1 if the label of zt+1 is +1 and N+ < N/2, or let zt+1 = z∗, N− = N−+1 if the label of zt+1 is −1 and N− < N/2 (if the value α (or α1, α2) is bigger than 1, accept the candidate sample z∗ with probability 1 ).
5. If N++N− < N, return to Step 4, else we obtain N Markov chain samples SMar. Let i = i + 1. Train SMar by SVMC and obtain a learning model fi .
6. If i < k, go to Step 2, else output sign( fk ).

Now we will train the SVM with the different types of kernels using the approach k-times markov sampling.the classifier is tested and the performance of each kernels are recorded on the datasets.

***Observation***

Algorithm is run for k-Markov Sampling with k = 1 and q = 1.2

For each iteration of k sample of 2000 randomly selected from

Data set

Accuracy on Linear Kernel SVM -0.77

Accuracy on RBF Kernel SVM - 0.801

Accuracy on Polynomial Kernel SVM - 0.678

For k=5 and q=1.2

Accuracy on Linear Kernel SVM - 0.87

Accuracy on RBF Kernel SVM - 0.912

Accuracy on Polynomial Kernel SVM - 0.841

For k = 10 and q = 1.2

Accuracy on Linear Kernel SVM - 0.8366

Accuracy on RBF Kernel SVM - 0.939

Accuracy on Polynomial Kernel SVM - 0.8908

**CONCLUSION**

So, generalization ability of the SVMCs which are based on the Markov Sampling is improved, as the number of training samples are very large. And for more iteration for SVMCs we can get maximum accuracy

**References:**

[1]k-Times Markov Sampling for SVMC, Bin Zou, Chen Xu, Yang Lu, Yuan Yan Tang, Fellow, IEEE, Jie Xu, and Xinge You, Senior Member, IEEE.

[2] Letter Dataset - https://archive.ics.uci.edu/ml/datasets/Letter+Recognition