INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, ALLAHABAD

VI Semester BTech in Information Technology

Report - Group Assignment 1

Data Mining and Warehousing

Group Members:

- **Mudit Goyal (IIT2018132)**
- **Shiv Kumar (IIT2018134)**
- **Moksh Grover (IIT2018186)**
- **Karan Chatwani (IIT2018194)**
- **Shubham (IIT2018200)**

Title: k-Times Markov Sampling for SVMC
Author: Bin Zou, Chen Xu, Yang Lu, Yuan Yan Tang, Fellow, IEEE, Jie Xu, and Xinge You

INTRODUCTION

Many machine learning applications, such as market prediction, system diagnosis, and speech recognition, are inherently temporal in nature, and consequently not independent and identically distributed. processes . Therefore, the relaxations of such . assumptions for SVMC have to be considered. They studied the generalization ability of SVMC with uniformly ergodic Markov chain samples, but the obtained learning rate is not optimal. The optimal learning rate of Gaussian kernels SVMC with samples by using the strongly mixing property of the Markov chain. The optimal learning rate of SVMC with u.e.M.c. samples and presented the numerical studies on the performance of SVMC with Markov sampling. Although the SVMC with Markov sampling introduced has smaller misclassification rates,its total time of sampling and training is longer compared with the classical SVMC based on randomly independent sampling. The main purpose of using k-times markov sampling SMVC is that we need to reduce the sampling and training time of SVMC with Markov sampling, at the same time keeping its smaller misclassification rates.

Markov sampling has three advantages at the same time compared with the classical SVMC and the SVMC with Markov sampling:
(1) The misclassification rates are smaller
(2) The total time of sampling and training is less
(3) the obtained classifiers are more sparse.

To have a better showing the performance of SVMC with k-times Markov sampling, we also give some discussions for the cases of unbalanced training samples and large-scale training samples.

IMPORTANT TERMINOLOGY:

SVM: SVM was created by Vapnik based on the Structural Risk Minimization principle. It is a binary classification algorithm that transforms data and finds the best boundary between the possible outputs depending on the transformations. This strategy is known as the kernel trick.
The following are some of the reasons why SVMs are important:
• Successful because there are a large number of features and a limited sample size.
• It is possible to learn both basic and complex classification models.
• Using complex statistical methods to prevent overfitting.

Algorithm:

Input: ST , N, k, q, n2
Output: sign(fk)
1: Draw randomly N samples Siid := {zj}N j=1 from ST . Train Siid by SVMC and obtain a eliminary learning model f0. Let i = 0.
2: Let Ni = 0, t = 1.
3: Draw randomly a sample zt from ST , called it the current sample. Let Ni = Ni + 1.
4: Draw randomly another sample z∗ from ST , called it the candidate sample. Calculate the ratio α, α = e−(fi,z∗) /e−(fi,zt)
.5: If α = 1, yt y∗ = 1 accept z∗ with probability α1 = e−y∗ fi /e−yt fi . If α = 1 and yt y∗ = −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∗, Ni = Ni + 1 (if α (or α1,α2) is greater than 1, accept z∗ with probability 1).
6: If Ni < 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 .
7: If i < k, go to Step 2, else output sign(fk)

Result:
Accuracy of k-Times Markov Sampling for SVMC with different Kernels for letter-recognition dataset.

Kernel	KPCA	SVDD	OCSVM	OCSSVM	OCSSVM with smo	k_MS_SVM
Linear	0.02	0.09	0.01	0.07	0.04	0.855

RBF	0.05	0.07	0.14	0.09	0.04	0.94225
Hellinger	0.01	0.02	0.02	0.13	0.10	0.8332
chi_square	0.18	0.0	0.02	0.18	0.17	0.891