SVM classification using k-times Markov sampling

*Sanket Mundra(IIT2018189), Raunak Rathour(IIT2018196), Prateek Mishra(IIT2018199)*

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

***Abstract: Support vector machines (SVM) have been widely used for practical classification tasks . Training complexity of SVM depends on the number of training samples and therefore it increases with large datasets. This paper focuses on the approach in [1], where classification with SVM has been improved based on Markov samplingThe results show that the provided classification approach is better than the SVMC in misclassification rates, training computational time.***

1. **INTRODUCTION**

Pattern recognition applications is a wide domain for SVM based classification algorithms [2]. SVM works well with samples which follow independent and identical distribution but this is not the case with applications which follow a time dependent pattern such as market prediction, speech recognition, which are dependent in nature [3]. In [4] and [5], the authors studied the generalization ability of SVMC with uniformly ergodic Markov chain (u.e.M.c.) samples while optimizing the learning rate of Gaussian kernels SVM. This model of SVMC with Markov sampling still has longer total time than the general SVM classifier model. To reduce this time, SVMC algorithm which is based on k-times Markov sampling is introduced in [1] and presents the numerical studies on its learning performance on well known data sets. In [6] the researchers compare the classic SVMC model with sampling and k-times sampling. The k-times sampling method has smaller misclassification rates;reduced training and sampling time and is sparse.

1. **ALGORITHM DESCRIPTION**

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 ).
7. **RESULTS AND OBSERVATIONS**
8. **CONCLUSION**
9. **REFERENCES**

**[1]** Zou, Bin & Xu, Chen & Lu, Yang & Tang, Yuan & Xu, Jie & You, Xinge. (2017). k-Times Markov Sampling for SVMC. IEEE transactions on neural networks and learning systems. 29. 10.1109/TNNLS.2016.2609441.

**[2]**V. Vapnik, Statistical Learning Theory. New York, NY, USA: Wiley, 1998.

**[3]**T. Steinwart, D. Hush, and C. Scovel, “Learning from dependent

observations,” J. Multivariate Anal., vol. 100, no. 1, pp. 175–194,

Jan. 2009.

**[4]**B. Zou, Z. Peng, and Z. B. Xu, “The learning performance of support

vector machine classification based on Markov sampling,” Sci. China

Inf. Sci., vol. 56, no. 3, pp. 1–16, Mar. 2013.

**[5]**J. Xu, Y. Y. Tang, B. Zou, Z. B. Xu, L. Q. Li, and Y. Lu, “Generalization

performance of Gaussian kernels SVMC based on Markov sampling,”

Neural Netw., vol. 53, pp. 40–51, May 2014.

**[6]** J. Xu et al., “The generalization ability of SVM classification based on

Markov sampling,” IEEE Trans. Cybern., vol. 45, no. 6, pp. 1169–1179,

Jun. 2015.