

SVR-Mediana: método e aplicação

CAIO BATALHA DIAS OLIVEIRA E ANDERSON ARA *

Abstract

A descoberta de conhecimento em conjunto dados é um processo semiautomático de extração de informações úteis a partir de um conjunto de dados, sendo que sua dimensionalidade inviabiliza que tal processo seja realizado manualmente [1]. O termo semiautomático é caracterizado pelo uso de ferramentas computacionais para o processo de descoberta, sendo indispensável a orientação do analista. As informações obtidas pelo processo de descoberta geralmente assumem a forma de padrões explicativos, muitas vezes referidos como modelos.

Uma nova classe de modelos no início dos anos 90, conhecida como máquina de vetores de suporte, ou *support vector machine* (SVM) [2]. Esta classe de modelos baseia-se na construção de hiperplanos para classificar rótulos - originalmente binários - bem como utiliza a kernelização para aumentar a flexibilidade do modelo, sendo a kernelização Gaussiana tida com alta capacidade preditiva e frequentemente utilizada [3]. Excelentes desempenhos desta metodologia foram obtidos para classificação, regressão e previsão de séries temporais [4]. Em específico, o desenvolvimento das máquinas de suporte de vetores para problemas de regressão denomina-se vetores de suporte para regressão, ou *support vector regression* (SVR).

Um modelo híbrido que integrou o método médias móveis (MA) com SVR para previsões de uma série temporal, através da diferença entre valor da variável Y_t e suas respectivas mediana de médias móveis semanais, a este método os autores chamaram de SVR-mediana [5].

References

- [1] Hamel, L. H. (2011). Knowledge discovery with support vector machines (Vol. 3). John Wiley & Sons.
- [2] Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992, July). A training algorithm for optimal margin classifiers. In Proceedings of the fifth annual workshop on Computational learning theory (pp. 144-152). ACM.
- [3] Wang, W., Xu, Z., Lu, W., & Zhang, X. (2003). Determination of the spread parameter in the Gaussian kernel for classification and regression. *Neurocomputing*, 55(3-4), 643-663.
- [4] Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing*, 14(3), 199-222.

*e-mail: caiobdioliv@gmail.com e anderson.ara@ufba.br

- [5] Hung, C., Hung, C. N., & Lin, S. Y. (2014). Predicting time series using integration of moving average and support vector regression. *International Journal of Machine Learning and Computing*, 4(6), 491.