

{tag} International Journal of Computer Applications
Foundation of Computer Science (FCS), NY, USA

[Volume 132](#)

-
[Number 4](#)

Year of Publication: 2015

Authors:

Dima Alberg

10.5120/ijca2015907398

{bibtex}2015907398.bib{/bibtex}

Abstract

Interval prediction can be more useful than single value prediction in many continuous data streams. This paper introduces a novel Interval Prediction Tree IP3 algorithm for interval prediction of numerical target variables from temporal mean-variance aggregated continuous data. This algorithm characterized by: processing incoming mean-variance aggregated multivariate temporal data, splitting each of the continuous features of the input according to the best mean-variance and making stable interval predictions of a target numerical variable with a given degree of statistical confidence. As shown by empirical evaluations in forest fires data set the proposed method provides better performance than existing regression tree models.

References

1. Alberg, D., Last, M., & Kandel, A. (2012). Knowledge Discovery in Data Streams with Regression Tree Methods. WIREs Data Mining Knowledge Discovery 2012 , 2, 69-78.
2. Bishop, C. (2006). Pattern Recognition and Machine Learning. New York: Springer.

3. Breiman, L., & Friedman, J. (1985). Estimating Optimal Transformations for Multiple Regression and Correlation. *Journal of American Statistic Association* , 80, 580 - 597.
4. Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and Regression Trees*. Belmont: Wadsworth & Brooks/Cole. Pacific Grove.
5. Ceci, M., Appice, A., & Malerba, D. (2003). Comparing Simplification Methods for Model Trees with Regression and Splitting Nodes. *Foundations of Intelligent Systems*, 14th International LNAI Symposium, 2871, pp. 49 - 56.
6. Cortez, P., & Morais, A. (2007). A Data Mining Approach to Predict Forest Fires using Meteorological Data. In M. F. In J. Neves (Ed.), *New Trends in Artificial Intelligence, Proceedings of the 13th EPIA 2007 - Portuguese Conference on Artificial Intelligence*, (pp. 512-523). Guimaraes, Portugal.
7. Friedman, J. (1991). Multivariate Adaptive Regression Splines. In *Annals of Statistics* , 1 - 19.
8. Friedman, J. (2002). Stochastic Gradient Boosting. *Computational Statistics & Data Analysis* , 38 (4), 367 - 378.
9. Ikonovska, E., Gama, J., Sebastiao, R., & Gjorgjevik, D. (2009). Regression Trees from Data Streams with Drift Detection. *Discovery Science*, (pp. 121 - 135).
10. Karalic, A. (1992). Linear Regression in Regression Tree Leaves. In *Proceedings of International School for Synthesis of Expert Knowledge*, 10(3), pp. 151 - 162.
11. Loh, W. (2009). Regression by Parts: Fitting Visually Interpretable Models with GUIDE. (W. H. in C. Chen, Ed.)
12. Quinlan, J. (1992). Learning with Continuous Classes. In *Proceedings of the 5th Australian Joint Conference on Artificial Intelligence*. World Scientific.
13. Vens, C., & Blockeel, H. (2006). A Simple Regression Based Heuristic for Learning Model Trees. *Intelligent Data Analysis* , 10(3), 215 - 236.
14. Wang, Y., & Witten, I. (1997). Inducing of Model Trees for Predicting Continuous Classes. In *Proceedings of the 9th European Conference on Machine Learning* (pp. 128 - 137). Springer-Verlag.

Index Terms

Computer Science

Information Sciences

Keywords

Interval Prediction, Mean-Variance Aggregation, Prediction Tree, Forest Fires.