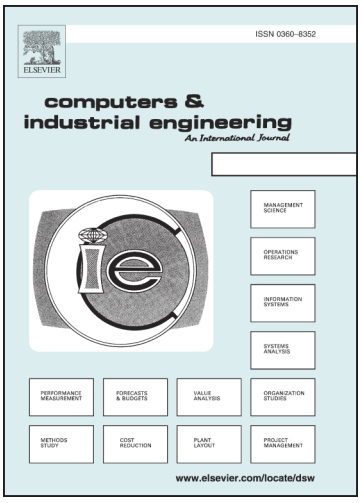
Accepted Manuscript



A support vector machine for model selection in demand forecasting applications.

Marco A. Villegas, Diego J. Pedregal, Juan R. Trapero

PII: S0360-8352(18)30186-4
DOI: https://doi.org/10.1016/j.cie.2018.04.042
Reference: CAIE 5191

To appear in: *Computers & Industrial Engineering*

Received Date: 1 February 2017
Revised Date: 3 February 2018
Accepted Date: 21 April 2018

Please cite this article as: Villegas, M.A., Pedregal, D.J., Trapero, J.R., A support vector machine for model selection in demand forecasting applications., *Computers & Industrial Engineering* (2018), doi: https://doi.org/10.1016/j.cie.2018.04.042

A support vector machine for model selection in demand forecasting applications.

Marco A. Villegas[a,*], Diego J. Pedregal[a], Juan R. Trapero[b]

Universidad de Castilla-La Mancha. Ciudad Real 13071, Spain

[a]*Department of Business Administration. ETSI Industriales.*
[b]*Department of Business Administration. Faculty of Chemical Science and Technology.*

Abstract

Time series forecasting has been an active research area for decades, receiving considerable attention from very different domains, such as econometrics, statistics, engineering, mathematics, medicine and social sciences. Moreover, with the emergence of the big data era, the automatic identification with the appropriate techniques remains an intermediate compulsory stage of any big data implementation with predictive analytics purposes.Extensive research on model selection and combination has revealed the benefits of such techniques in terms of forecast accuracy and reliability. Several criteria for model selection have been proposed and used for decades with very good results. Akaike information criterion and Schwarz Bayesian criterion are two of the most popular criteria. However, research on the combination of several criteria along with other sources of information in a unified methodology remains scarce.

This study proposes a new model selection approach that combines different criteria using a support vector machine (SVM). Given a set of candidate models, rather than considering any individual criterion, an SVM is trained at each forecasting origin to select the best model. This methodology will be particularly interesting for scenarios with highly volatile demand because it allows changing the model when it does not fit the data sufficiently well, thereby reducing the risk of misusing modeling techniques in the automatic processing of large datasets.

The effects of the proposed approach are empirically explored using a set of representative forecasting methods and a dataset of 229 weekly demand series from a leading household and personal care manufacturer in the UK. Our findings suggest that the proposed approach results in more robust predictions with lower mean forecasting error and biases than base forecasts.

Keywords: Demand forecasting, supply chain, SVM, time series analysis, model selection.

1. Introduction

Companies have traditionally adopted forecasting techniques to support decision making on a consistent daily basis, bringing data from different sources into a common data infrastructure. However, the primary focus was mainly the development of reporting tools (Davenport and Harris, 2007). In recent years, the so-called business analytics has introduced a new approach in this domain by leveraging on the latest progress in both computer science (e.g., data mining algorithms) and hardware technology (e.g., cloud computing and in-memory technology), thus enabling the integration of data sources and business operations on a higher stage of abstraction (Sheikh, 2013).

With the emergence of the big data era, the automatic identification of appropriate data techniques is an intermediate compulsory stage of any big data implementation with predictive analytics purposes. Research on model selection and combination has revealed the benefits of such techniques in terms of forecast accuracy and reliability. However, the application of artificial intelligence techniques to this problem is still scarce. Forecasting models are of a strategic nature given that they guide business decisions, ranging from inventory scheduling to strategic management (Petropoulos et al., 2014). Focusing on a supply chain context, automatic model selection is a necessity due to the high number of products whose demand should be forecast (Fildes and

*Corresponding author. Tel.: +34 926 295 300 Ext: 96670.
Email address: marco.villegas@uclm.es (Marco A. Villegas)