Neurocomputing 73 (2010) 1919–1922

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Editorial

European Symposium on Times Series Prediction

1. Introduction

Time series forecasting is a challenge in many fields. In finance, one forecasts stock exchange courses or stock market indices; data processing specialists forecast the flow of information on their networks; producers of electricity forecast the load of the following day. The common point to their problems is the following: how can one analyze and use the past to predict the future? Many techniques exist including linear methods such as ARX or ARMA, and nonlinear ones such as the ones used in the area of machine learning [1].

In general, these methods try to build a model of the process that is to be predicted. The model is then used on the last values of the series to predict future ones. The common difficulty to all methods is the determination of sufficient and necessary information for a good prediction. If the information is insufficient, the forecasting will be poor. On the contrary, if information is useless or redundant, modeling will be difficult or even skewed. In parallel with this determination, a suitable prediction model has to be selected. In order to compare different prediction methods several competitions have been organized, for example, the Santa Fe Competition [2], the CATS Benchmark Competition [3] and the ESTSP'07 Competition [4].

After the competitions, their results have been published and the time series have become widely used benchmarks. The goal of these competitions is the prediction of the subsequent values of a given time series (3–100 values to predict). Unfortunately, the long-term prediction of time series is a very difficult task. Furthermore, after the publication of results, the real values that had to be predicted are also published. Thereafter, it becomes more difficult to trust in new results that are published: knowing the results of a challenge may lead, even unconsciously, to bias the selection of model; some speak about "data snooping". It becomes therefore more difficult to assess newly developed methods, and new competitions have to be organized.

This special issue is based on extended version of papers presented at the joined ESTSP'08 (European Symposium on Time Series Prediction) and AKRR'08 (Adaptive Knowledge Representation and Reasoning) conferences [5,6]. This shared event took place in Porvoo, Finland, from 17th to 19th of September, 2008. The goal of joining these conferences was to create an interdisciplinary forum for researchers who may widen their scope of attention beyond the usual scope of research. The crossfertilization took place, for instance, by offering the attendees shared keynote talks. Prof. Marie Cottrell (Paris University 1) gave a talk on data analysis using Self-Organizing Maps. Prof. José Príncipe (University of Florida) described information theoretic learning and kernel methods. Dr. Harri Valpola (Helsinki University of Technology) explained how to extract abstract concepts from raw data using statistical machine learning methods. One specific shared theme of interest was anticipation, i.e., how an agent makes decisions based on predictions, expectations, or beliefs about the future. Anticipation is an important concept when complex natural cognitive systems are considered [7].

2. ESTSP'08 competition

The goal of the ESTSP'08 competition was to predict the future of three very different Time Series.[1] Firstly, the length and the sampling period of the time series are very different. Secondly, the origin of each time series varies. The data and the origins, i.e., environment, electric load, and internet traffic, are described below in more detail. In order to provide the participants an equal opportunity for success, the origins of the three time series were kept secret until the end of the competition.

2.1. Data sets

Chemical descriptors of environmental condition: This series is part of a multidimensional time series of monthly averages of different chemical descriptors of a certain area of the Baltic Sea. The series is made of 354 samples and spans for 29.5 years. This competition data set is shown in Fig. 1. For this time series, the goal was to predict the next 18 values of the third time series, using the two other one as exogenous variables.

Traffic in a data network: The second dataset from the ESTSP 2008 competition is a univariate time series consisting of 1300 samples that describe the daily average amount of traffic in a data network. The competition data set 2 is shown in Fig. 2. For this time series, the goal was to predict the next 100 values of the time series.

Electric load: The third dataset was a univariate time series consisting of 31 614 samples that describe the daily average amount of electric load. The competition data set 2 is shown in Fig. 3. For this time series, the goal is the prediction of the next 200 values of the time series.

2.2. Results

Twenty sets of predictions have been submitted to the competition. The results in Table 1 present the Normalized Test

[1] The data sets can be downloaded from http://www.cis.hut.fi/projects/tsp/index.php?page=timeseries.

0925-2312/$ - see front matter
doi:10.1016/j.neucom.2010.02.011

Mean Squared Error for the three predictions respectively. We present only the results of the participants that agreed to have their results published. The winners of the competition were Rubio, Herrera, Pomares, Rojas and Guillen [3].

3. Summary of the special issue papers

For this special issue, 17 authors were invited to submit an extended version of their conference paper. Finally, 14 extended paper were accepted. Not all the authors participated to the ESTSP'08 Competition. The list of papers can be classified in 3 distinct categories:

1. The authors that participated to the competition.
2. The papers that presents new methods for the analysis and/or prediction of Time Series but did not participate in the competition.
3. The papers that participated in the ESTSP08-AKRR'08 Special Session on Prediction for Finance organized by Prof. Eric Séverin.

3.1. Competition papers

Crone and Kourentzes [8] propose a data driven, fully automated methodology to specify multilayer perceptrons for time series prediction using a combination of iterative (neural network) filters and wrappers. Their approach is capable of identifying unknown time series frequencies, multiple overlying seasonality, and additional relevant features without human expert intervention. The approach has shown promising performance in forecasting by ranking second in the ESTSP competition.

Pouzols and Barriga [15] deal with an automatic methodology for clustering-based fuzzy inference models. A number of clustering methods are compared and an extension of Improved Clustering for Function Approximation is proposed. The approach yields compact models and its accuracy and speed compare favorably against MLP, LS-SVM and ELM models for a diverse set of time series benchmarks.

Ben Taieb, Sorjamaa and Bontempi [9] present a new multiple-output approaches for Multi-Step-Ahead Time Series Forecasting and compares it to state-of-the-art approaches. The extensive validation made with the series of the NN3 competition shows that the multiple-output paradigm is very promising and able to outperform conventional techniques.

Reservoir Computing has been shown to perform well in chaotic time series prediction. Wyffels and Schrauwen [11] extend these results by a comparison of multiple Reservoir Computing strategies for time series prediction (including research on regularization, influence of reservoir size and decomposition) in the domain of noisy, seasonal time series prediction for industrial purposes. They compare their approach to standard approaches such as ARIMA modeling and NAR modeling using LS-SVMs.

Rubio, Herrera, Pomares, Rojas and Guillen [16] present a kernelized version of the weighted k-nearest neighbors method (KWKNN) for regression problems and address the creation of specific-to-problem kernels for time series data. This unified

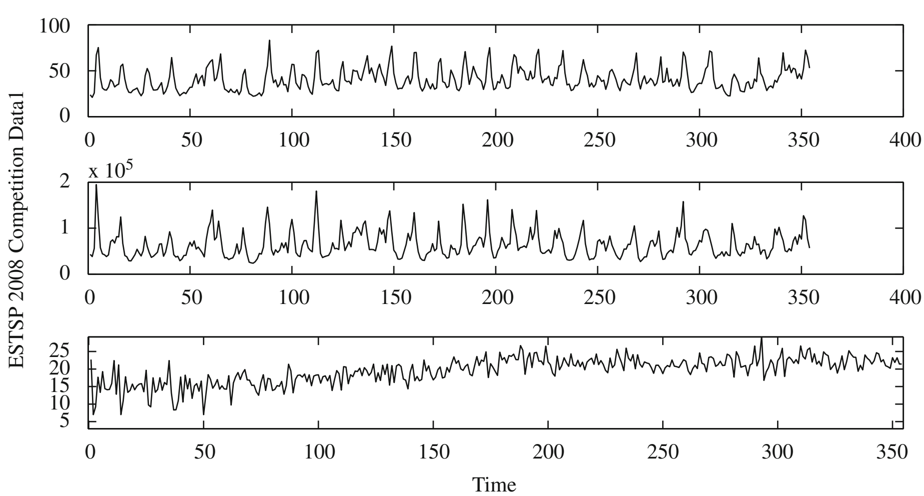

Fig. 1. ESTSP 2008 competition data 1.

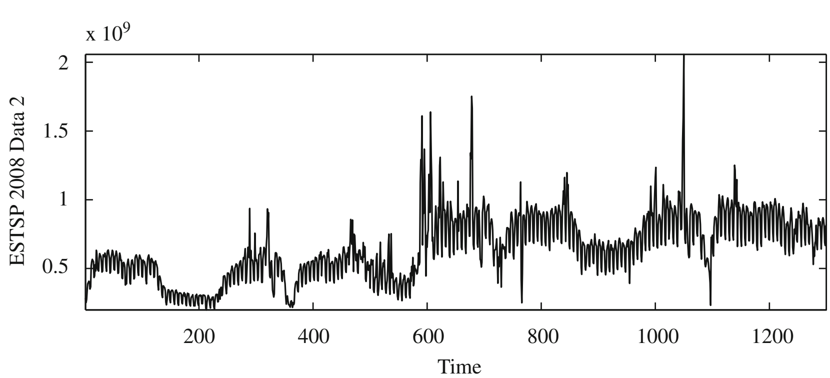

Fig. 2. ESTSP 2008 competition data 2.

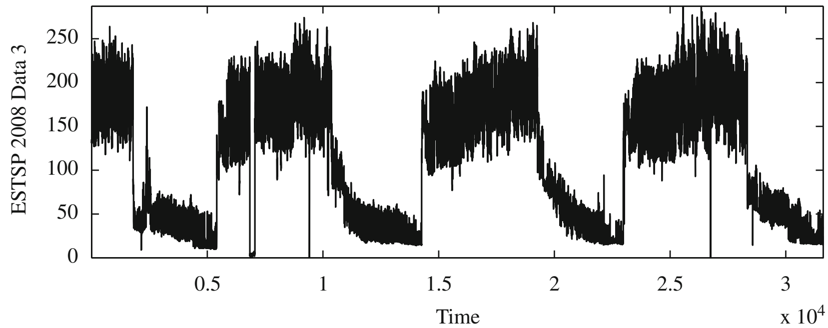

Fig. 3. ESTSP 2008 competition data 3.

Table 1
Competition results: test NMSE for each data set.

	Kourentzes [8]	Bontempi [9]	Olteanu [10]	Wyffels [11]	Espinoza [12]	Adeodato [13]	Rubio [14]	Montesino [15]
Data 1	0.07	0.12	0.19	0.157	0.112	0.151	0.079	0.16
Data 2	0.212	0.431	0.359	0.529	0.266	0.49	0.208	0.4
Data 3	0.25	1.802	1.655	1.582	0.464	1.611	0.036	1.344
Total	0.178	0.785	0.735	0.756	0.281	0.751	0.107	0.635

framework for kernel and *k*-nearest neighbors methods allows for a comparison of KWKNN with LSSVM using time series prediction examples with interesting results. Additionally, a parallel implementation of KWKNN, developed in order to speed up the method and make it practical for large datasets, is proposed and applied to a large scale problem.

3.2. General papers

Sovilj, Sorjamaa, Yu, Miche and Séverin [17] present a methodology for long-term time series prediction that can also be applied to standard regression tasks. The methodology consists of two main steps: (1) input variable scaling or projection with Delta Test, optimized with Genetic Algorithm, and (2) prediction on the projected data using two models, Optimally-Pruned Extreme Learning Machine and Optimally-Pruned k-Nearest Neighbors. The methodology is tested on two time series prediction tasks and one financial regression problem.

Nybo [18] provides an applied perspective from the petroleum industry. Normally conservative, this industry nonetheless shows an increasing interest in machine learning and data mining. The paper gives a taste of the new opportunities in this industry and goes on to show how a successful choice of machine learning algorithms becomes governed by the industry's work processes and the user's behavioral mode.

Souza and Barreto [19] provide a comprehensive performance evaluation of the use of vector quantization (VQ) algorithms to building local models for inverse system identification. Statistical hypothesis testing is carried out through the Kolmogorov–Smirnov test in order to study the influence of the VQ algorithms on the performances of the local models. Tests on four benchmarking input–output time series reveal that the resulting local models achieve performances superior to standard global MLP-based model.

Lemke and Gabrys [20] describe how the performance of the time series forecasting algorithms differ depending on the data set used. However, for a limited data set of similar time series, it can be possible to determine one particular method or combination of methods that performs best. Following this idea, the article presents an empirical study extracting characteristics of time series in order to generate domain knowledge. This knowledge is then used to dynamically select or combine different forecasting algorithms.

Mateo, Sovilj and Gadea [21] present a method that uses genetic algorithms to select an optimum set of input variables that minimizes the Delta Test on a dataset. The nearest neighbor computation has been speeded up by using an approximate method. The scaling and projection of variables has been addressed to improve the interpretability.

Guillen, Herrera, Rubio, Pomares, Lendasse and Rojas [14] present a totally new approach for the problem of filtering the outliers, reducing the noise and defining a good subset of samples. The approach is based in the concept of Mutual Information with the advantage of just having one parameter to be tuned. The simple idea is efficient and easy to implement, providing satisfactory results within a wide range of problems.

Korpela, Mäkinen, Nöjd, Hollmén and Sulkava [22] present a Markov-switching autoregressive model. Its performance is compared with other statistical and machine learning methods in a new kind of real-world change detection problem with environmental time-series.

3.3. Financial prediction papers

du Jardin [23] presents two main results. It is shown that a neural-network-based model for predicting bankruptcy performs better when designed with appropriate variable selection techniques than when designed with methods commonly used in the financial literature. Furthermore, it has been found that there is a relationship between the structure of a prediction model and its ability to reduce Type I errors.

Séverin [24] deals with the advantages of the self-organizing map algorithm in the field of corporate finance. Not only the SOM method is able to improve the classical method for bankruptcy prediction but it also questions the scoring models.

References

[1] A. Sorjamaa, J. Hao, N. Reyhani, Y. Ji, A. Lendasse, Methodology for long-term prediction of time series, Neurocomputing 70 (16–18) (2007) 2861–2869.
[2] A. Weigend, N. Gershenfeld (Eds.), Times Series Prediction: Forecasting the Future and Understanding the Past, Addison-Wesley Publishing Company, 1994.
[3] A. Lendasse, E. Oja, O. Simula, M. Verleysen, Time series prediction competition: the CATS benchmark, Neurocomputing 70 (13–15) (2007) 2325–2329 , doi:10.1016/j.neucom.2007.02.013.
[4] A. Lendasse (Ed.), in: ESTSP 2007: Proceedings, Multiprint Oy / Otamedia, 2007.

[5] A. Lendasse (Ed.), in: ESTSP 2008: Proceedings, Multiprint Oy/Otamedia, 2008.
[6] T. Honkela, M.-S. Paukkeri, M. Pöllä, O. Simula (Eds.), in: Proceedings of AKRR'08, International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning, Helsinki University of Technology, Espoo, Finland, 2008.
[7] R. Rosen, Anticipatory Systems, Pergamon Press, Oxford, UK, 1985.
[8] S.F. Crone, N. Kourentzes, Feature selection for time series prediction—a combined filter and wrapper approach for neural networks, Neurocomputing (2010), this issue, doi: 10:1016/j.neucom.2010.01.017.
[9] S.B. Taieb, A. Sorjamaa, G. Bontempi, Multiple-output modelling for multi-step-ahead time series forecasting, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.11.030.
[10] M. Olteanu, Revisiting linear and non-linear methodologies for time series prediction—application to ESTSP'08 competition data, in: European Symposium on Time Series Prediction, ESTSP'08, 2008, pp. 139–148.
[11] F. Wyffels, B. Schrauwen, A comparative study of reservoir computing strategies for monthly time series prediction, 2010, this issue, doi:10:1016/j.neucom.2010.01.016.
[12] M. Espinoza, T. Falck, J.A.K. Suykens, B.D. Moor, Time series prediction using ls-svms, in: European Symposium on Time Series Prediction, ESTSP'08, 2008, pp. 159–168.
[13] P.J.L. Adeodato, A.L. Arnaud, G.C. Vasconcelos, R.C. Cunha, D.S. Monteiro, Exogenous data and ensembles of mlps for solving the estsp forecast competition tasks, in: European Symposium on Time Series Prediction, ESTSP'08, 2008, pp. 169–176.
[14] A. Guillen, L.J. Herrera, G. Rubio, H. Pomares, A. Lendasse, I. Rojas, New method for instance or prototype selection using mutual information in time series prediction, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.11.031.
[15] F. Montesino-Pouzols, A.B. Barros, Automatic clustering-based identification of autoregressive fuzzy inference models for time series, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.11.028.
[16] G. Rubio, L.J. Herrera, H. Pomares, I. Rojas, A. Guillen, Design of specific-to-problem kernels and use of kernel weighted k-nearest neighbours for time series modelling, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.11.029.
[17] D. Sovilj, A. Sorjamaa, Q. Yu, Y. Miche, E. Séverin, Op-elm and op-knn in long-term prediction of time series using projected input data, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.11.033.
[18] R. Nybo, Fault detection and other time series opportunities in the petroleum industry, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.10.020.
[19] L.G.M. Souza, G.A. Barreto, On building local models for inverse system identification with vector quantization algorithms, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.10.021.
[20] C. Lemke, B. Gabrys, Meta-learning for time series forecasting and forecast combination, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.09.020.
[21] F. Mateo, D. Sovilj, R. Gadea, Approximate *k*-nn delta test minimization method using genetic algorithms: application to time series, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.11.032.
[22] M. Korpela, H. Mäkinen, P. Nöjd, J. Hollmén, M. Sulkava, Automatic detection of onset and cessation of tree stem radius increase using dendrometer data, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.11.035.
[23] P. du Jardin, Predicting bankruptcy using neural networks and other classification methods: the influence of variable selection techniques on model accuracy, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.11.034.
[24] E. Séverin, Self organizing maps in corporate finance: quantitative and qualitative analysis of debt and leasing, Neurocomputing (2010), this issue, doi:10:1016/j.neucom.2009.12.024.

Amaury Lendasse *, Timo Honkela, Olli Simula
Department of Information and Computer Science, Aalto University School of Science and Technology, 00076 Aalto, Finland
E-mail addresses: Amaury.Lendasse@hut.fi, lendasse@hut.fi (A. Lendasse), Timo.Honkela@hut.fi (T. Honkela), Olli.Simula@hut.fi (O. Simula).

* Corresponding author.