Time Series Analysis and Possible Applications

Mirjana Ivanović*, Vladimir Kurbalija*
* University of Novi Sad, Faculty of Sciences, Novi Sad, Serbia
e-mail: {mira, kurba}@dmi.uns.ac.rs

Abstract - Data presented in form of time series as its analysis and applications recently have become increasingly important in different areas and domains. In this paper brief overview of some recently important standard problems, activities and models necessary for time series analysis and applications are presented. Paper also discusses some specific practical applications.

I. INTRODUCTION

Recently data presented in form of time series as its analysis and applications have become increasingly important in different domains. A time series is a set of ordered observations on a quantitative characteristic of a phenomenon at equally spaced time points. One of the main goals of time series analysis is to forecast future values based on existing series. There are a lot of possible domains and various fields of research based on time series like: business, engineering, environometrics, economics, medicine, politics, social sciences, and a lot of others. Initial and significant book in the area *Time Series Analysis: Forecasting and Control*, has been published by Box and Jenkins (1970, 1976) bringing comprehensible methodologies and providing a modernized approach to time series analysis and forecasting. After that a numerous books and enormous number of research, theoretical and practical, papers have been published changing attitude and increasing significance of such kind of data.

One of the main activities in this domain is time series modeling which supports to carefully collect and rigorously study the past observations of a time series in order to develop an appropriate model. Such model serves to describe the inherent structure of the series and to produce adequate framework that then has been used to generate future values for the series, and make useful forecasts.

The use of computers and appropriate software tools is necessary in modern quantitative analysis. It is also extremely important in time series analysis where different complex algorithms and extensive computations are unavoidable. In many research areas and situations adoption of different methodologies and usage of adequate software tools are necessary to provide simplified and effective processes for complex subjects in time series analysis and forecasting.

The rest of the paper is organized as follows. In second section a brief overview of contemporary techniques for time series modeling and processing are presented. Section III is devoted to some of standard activities for better representation, pre-processing and similarity analysis of time series. Section IV briefly discusses some possible practical applications of time series in several domains. Section V introduces a framework for analysis and prediction in time series processing. Last section concludes the paper.

II. SOME IMPORTANT TIME SERIES MODELS

Time series analysis, modeling and forecasting have fundamental importance to various practical domains. To improve the efficiency and accuracy of time series modeling and forecasting in last two decades a lot of innovative, important models have been proposed. Some authors recognize [1] three important classes of time series models: the stochastic, neural networks and SVM based models.

Among the frequently used and evidently most popular stochastic time series models is the *ARIMA - Autoregressive Integrated Moving Average* [2, 3] model. The essential assumption in implementation of this model is: time series are linear and follows a particular known statistical distribution, like the normal distribution. ARIMA model is further explored for establishing other subclasses of models, like the *Autoregressive (AR)*, *Moving Average (MA)* and *Autoregressive Moving Average (ARMA)*. Box and Jenkins [4] had proposed very successful variation of ARIMA model for known phenomena of seasonal time series forecasting, the *SARIMA - Seasonal ARIMA.*

In spite the fact that ARIMA model is flexible, simple for optimal model building process it suffers of severe limitation. One of important limitations is pre-assumed linear form of the associated time series which becomes inadequate in different practical situations. Consequently various non-linear stochastic models have been developed in order to overcome this drawback. On the other hand these models are not so straight-forward and simple from implementation point of view as the ARIMA models.

Recently, general revitalization of artificial neural networks (ANNs) also has influenced its growing attentions in the fields of time series forecasting [2, 5]. ANNs especially have been successfully applied for forecasting and classification purposes in wade range of areas [5]. The excellent feature that influenced use of ANNs in forecasting problems is their inherent capability of non-linear modeling, without any presumption about the statistical distribution followed by the adequate observations [5].

ANNs possess two important features, first of all they are data-driven (as appropriate model is adaptively formed based on the given data) and on the other hand they are self-adaptive [6]. During the past several years a considerable amount of research have been carried out towards the application of neural networks for time series modeling and forecasting.

There are various ANN forecasting methods and algorithms like the multi-layer perceptrons (MLPs), which are characterized by a single hidden layer *Feed Forward Network (FNN)* [6] but also widely used variation of FNN like *Time Lagged Neural Network (TLNN)*. A very interesting and new approach is oriented towards ANN model for seasonal time series forecasting *Seasonal Artificial Neural Network (SANN)*. It had been presented in more details in [7].

This model is unexpectedly simple and has been experimentally verified to be quite successful and efficient in forecasting seasonal time series.

A significant step forward happened in the area of time series forecasting with the development of concept of support vector machine (SVM) [8, 9]. The initial aim of SVM was to solve pattern classification problems but later they have been started to be applied in many other fields including time series prediction problems [9]. Apart from fact that it is useful for classification tasks it is also intended for a better generalization of the training data. For this reason the SVM methodology getting one of the frequently used techniques for time series forecasting problems. The objective of SVM is to use the *structural risk minimization (SRM)* principle to find a decision rule with good generalization capacity. Important feature of SVM is that here the training is equivalent to solving a linearly constrained quadratic optimization problem. Unlike the other traditional stochastic or neural network methods solution achieved by applying SVM method is always unique and globally optimal. Above all the most significant property of SVM is that the quality and complexity of the solution can be independently controlled, irrespective of the dimension of the input space [10]. Usually in SVM applications, the input points are mapped to a high dimensional feature space using special functions - *support vector kernels*. In last several years numerous SVM forecasting models have been developed for time series forecasting among the most frequently used are: *Least-square SVM (LS-SVM)* and *Dynamic Least-square SVM (LSSVM).*

III. ANALYSIS OF TIME SERIES

Time series embody one form of sequential data where the value of some numeric parameter in an observed process is recorded periodically in time. Each element of a time series describes the phenomenon under examination at a specific point in time. The parameters of interest can be various: the recorded values of some scientific sensor, the daily number of hits of some website, the value of shares on a stock market, the number of rainy days per year, etc. The observation rate can vary from one millisecond to several years, and in the general case the observation rate need not to even be uniform. However, the common characteristic of time series is that they usually consist of a large number of points (recorded values), which can make the handling of time series difficult.

Time series are used for storage, display and analysis of data across a wide range of different domains, including various areas of science, medicine, economics, ecology, telecommunications and meteorology [11–13]. In almost every scientific field, measurements are performed over time and the collected data can be organized in the form of time series with the aim of extracting meaningful knowledge.

Time-series analysis applies different methods, mainly from statistics, data mining and machine learning, in order to model and understand the process which generated time series or to make forecasts. Accordingly, several task types are important for time-series analysis, most notably indexing, classification, clustering, prediction, forecasting, segmentation, and anomaly detection [14–17]. In addition, there are three important concepts which need to be considered when dealing with time-series data, and which will be described in more detail: pre-processing transformation, time-series representation and similarity/ distance measure.

A. Pre-processing Transformations

Very commonly "raw" time series comes with different kinds of distortions. For most real-world applications these distortions are not meaningful, and they should be removed. This is the task of pre-processing phase. Generally, there are four groups of pre-processing transformations: **Offset translation, Amplitude scaling, Removing linear trend, Removing noise.**

Offset translation: Frequently, the values of various time series does not belong to the same interval, although their overall shape is very similar. In those situations, the values of time series should be translated into the same value-interval. The simplest solution for this problem is to subtract the mean value of each time series.

Amplitude scaling: In some situations, two or more time series have different amplitudes. If this property is undesirable, the amplitude scaling transformation should be applied. The most common approach is to subtract the mean value from time series and to divide this difference by standard deviation.

Removing linear trend: In some cases, time series contain some constant linear trend (growing or decreasing). In these situations, the most common approach is to fit a straight line through the time series, and then to subtract that line from the time series.

Removing noise: Very often, raw time series contains some kind of noise. There are many techniques for noise removal, but most of them are based on moving averages methods. The essence of these methods is to average each data point of time series with two or more of its neighbors.

Finally, it is worth mentioning that sometimes the distortions are the most interesting thing about the data, and in these cases pre-processing is not performed. The selection of pre-processing tasks depends on data and on task types which have to be performed.

B. *Time-series Representations*

As already mentioned, the dimensionality of time series, depending on domain, can be huge. Therefore, dimensionality of time series must often be reduced before applying certain data mining tasks upon it. Main purpose of time series representations is to decrease data dimensionality while keeping the important characteristics of the original time series which usually implies the overall behavior of time series and the shape. This process, of course, includes some information loss, so the main objective is to keep crucial information. By keeping important information, representation preserves time series features, behaving just like original time series in data mining algorithms. Among many proposed representation techniques, only the major ones will be described here.

Discrete Fourier Transform (DFT). DFT is actually used in many fields (such as digital signal processing, image processing, solving of partial differential equations, etc.), and has also found its application in time series data mining. Basic idea of DFT is that any signal can be represented by the super position of finite number of sine/cosine waves. After applying DFT, dimensionality of time series stays unchanged, so to reduce data dimensionality, we need to dismiss some Fourier coefficients. It is observed that only the first few coefficients appear to be dominant and therefore the rest can be omitted without great information loss [18]. In that way data series dimensionality can be efficiently decreased.

Discrete Wavelet Transform (DWT) [19]. Wavelets are functions that represent time series in terms of the sum and difference of prototype function, called mother wavelet. Unlike DFT, in which case each coefficient carries out only global information, DWT coefficients hold local information (each wavelet holds information of some time series segment). First coefficients are more global, while the each next group of coefficients refines global picture by adding more information to certain segments of time series. That is why they are very suitable for decreasing time series dimensionality. By taking first few DWT coefficients, we are actually taking global time series characteristics, and that is exactly what time series representation should hold.

Piecewise Aggregate Approximation (PAA) [20]. Underlying idea is very simple. Original time series is divided in N equal segments. For each segment mean value is calculated and stored. In that way dimensionality of original time series is decreased to N. Although very simple, this method proved to be reasonably good when used in data mining algorithms.

Piecewise Linear Approximation (PLA) [21]. Just like in PAA, time series is divided in N segments (here not necessarily of equal size). Each segment is then represented by a line. There are two ways of defining that line: it can be linear interpolation between start and end point of corresponding segment, or it can be linear regression that takes into account all points that are contained in the segment. Each segment is then presented by two numbers (a and b coefficients of line equation $y = ax + b$), so if dimensionality of final representation needs to be N, time series is divided into N/2 segments. As noted at the beginning, time series segments by definition do not have to be of equal size. Segments lengths are determined by different kind of algorithms whose main aim is to create representation of minimal reconstruction error. The problem with that approach is that distance measure upon this kind of representation does not support lower bounding, which is necessary in order to use representation in indexing algorithms.

Indexable Piecewise Linear Approximation (IPLA) [22]. IPLA solves the indexing problem of PLA in two aspects. According to first modification, IPLA does not allow segments of different lengths. Second modification introduces reset of x (time) component of each time series segment.

Symbolic Aggregate Approximation (SAX) [23]. This representation is built upon PAA. Idea of SAX is to convert data into a discrete format which is based on predefined alphabet. Before discretization, time series is transformed via PAA algorithm. In order to convert PAA coefficients to alphabet symbols, we must first define breakpoints that divide the distribution space into S equiprobable regions, where S is the size of alphabet. Each defined region is mapped into one alphabet symbol. Transformation of PAA representation is now trivial each coefficient is mapped and transformed into one alphabet symbol.

C. *Similarity/Distance Measures*

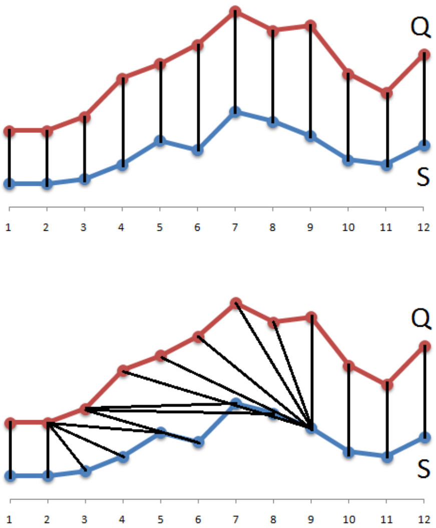
One of the most important aspects of time-series analysis is the choice of appropriate similarity/distance measure – the measure which tells to what extent two time series are similar. The choice of an appropriate time-series similarity measure is a critical point when dealing with many tasks in mining temporal data. While working with traditional databases we are interested in data that exactly match the given query, in the case of similarity-based retrieval of time series, we are looking for sequences that most resemble a given series. As similarity-based retrieval is explicitly or implicitly used in all main tasks of time series analysis (including classification, clustering, forecasting, etc.), it is important to carefully define the similarity measure between time series in order to reflect the underlying (dis)similarity of the specific data they represent. There is a large number of (dis)similarity measures for time series data proposed in the literature, but here we will present only the several major ones.

L_p-norm is one of the most popular classes of similarity measures that are used in data mining algorithms. Its definition is given in formula (1) where Q and C are time series of length n, q_i is the value of Q at time point i, and C_i is value of C at time point i. Variable p denotes the norm that is used. For example, p = 1 corresponds to the Manhattan distance, p = 2 is Euclidean distance, and p = ∞ is Chebyshev distance.

$$D(Q,C) = \sqrt[p]{\sum_{i=1}^{n}\left(q_i - c_i\right)^p} \qquad (1)$$

The advantage of Euclidean distance is that it is very easy to compute and to understand. There are, however, some disadvantages, too: the sequences must have the same number of points (can be avoided by interpolation to equal length), it is sensitive to shifting and scaling along the y-axis (can be precluded by normalizing the series), and it is also sensitive to distortions and shifting along the time axis. The last problem is successfully solved with elastic measures; all measures that are presented in further text are elastic distance measures.

Dynamic Time Warping (DTW). Euclidean distance is based on linear aligning of related points of time series (Fig. 1 (a)): the i-th point of the first series is paired with the i-th point of the second one. The assessment of the distance can be improved by warping the time axis of one or both of the sequences (Fig. 1 (b)). One of the most popular distance measures based on non-linear aligning is the Dynamic Time Warping [24].



(a) (b)

Figure 1. (a) Linear aligning of time series in case of the Euclidean distance, (b) Non-linear aligning of DTW

Longest Common Subsequence (LCS) [25]. This elastic similarity measure is one of the edit distances. The similarity of two time series is presented as the length of their longest common subsequence.

Edit Distance on Real Sequence (EDR) [26] is another edit distance that found its application in time series data mining. EDR value is actually minimal number of edit operations (insertions, substitutions and deletions) that need to be applied upon one time series in order to transform it to another time series. This similarity measure proved to be very robust to noise.

Edit distance with Real Penalty (ERP) [27]. Main problem with all mentioned elastic measures is that they do not satisfy the triangle inequality, so they cannot be used in indexing algorithms. ERP similarity measure combines L1-norm with DTW and EDR, dismissing segments of those measures that are causing them not to satisfy triangle inequality.

All of these elastic similarity measures are based on dynamic programming. It is well known that the computational complexity of dynamic programming algorithms is quadratic, which is often not suitable for larger real-world problems. However, the usage of global constraints such as Sakoe-Chiba band and Itakura parallelogram can significantly speed up the calculation of similarities. Furthermore, it is also reported [28] that the usage of global constraints leads to a qualitatively different measures which can have even improved accuracy of classification compared to unconstrained similarity measures. The accuracy of classification is commonly used as a qualitative assessment of a similarity measure.

IV. Some Possible Real Life Applications of Time Series Data

Significance and importance of time series analysis and practical use is getting more and more important in almost all areas of human life like science, engineering, technical, economical and sociological systems. So time series modeling is extremely important and this model is then used to generate future values for the series, i.e. to make appropriate forecasts. Time series forecasting thus can be termed as the act of predicting the future by understanding the behaviors detected in the past. Due to the indispensable importance of time series forecasting they are significant scientific trend in numerous practical fields such as business, economics, finance, science, medicine and engineering.

The choice of the best and most appropriate forecasting models depends on the forecasting horizon and usually there are three types of time series forecasts.

Point Forecast: a single number or a "best guess." It does not provide information on the level of uncertainty around the point estimate/forecast.

Interval Forecast: relative to a point forecast, this is a range of forecasted values which is expected to include the actual observed value with some probability.

Density Forecast: this type of forecast provides information on the overall probability distribution of the future values of the time series of interest.

However, the point forecast is the most commonly used type of forecast by businesses managers and policymakers.

A. Time Series in Finanial and Business Domains

Policymakers and business managers on a regular basis use forecasts of financial variables to help make important decisions about production, purchases, market conditions, and other choices about the best allocation of resources [29].

Time series analysis is an integral part of financial analysis with applications to the prediction of interest rates, foreign currency risk, stock market volatility, and a lot of other similar tasks and activities.

In comparison to other areas and domains, the modeling of economic time series is characterized by several problems [29]:

- The empirical sample sizes in economics (usually ranges between 25 - 100 observations) are generally small, especially compared to other areas like physics or biology with more than several thousands of observations.

- Economic time series are highly dependent and they are correlated with other economic time series. In the economic science, problems are almost never concerned with univariate series. An example can be a function of income, and at the same time, consumption also affects income directly and through various other variables.
- Economic time series are also often dependent over time. Many series display high autocorrelation, as well as cross autocorrelation with other variables over time.
- Economic time series are generally non-stationary. Their means and variances change over time, implying that estimated parameters might follow unknown distributions as an alternative of standard tabulated distributions like the normal distribution. Non-stationarity arises from productivity growth and price inflation. In some situations, for example, inference in econometrics become quite complicated, and requires the development of novel statistical techniques for handling stochastic trends. Examples of such new techniques include the concepts of cointegration and common trends, and asymptotic theory for integrated variables.
- Economic time series cannot be assumed to be drawn from samples like in classical statistics. Since the sampling process can be controlled the variables which make up the sample can be seen as random variables. Hypothesis are then formulated and tested conditionally on the assumption that the random variables have a specific distribution. Economic time series are seldom random variables drawn from some underlying population in the classical statistical sense.
- Different variables like, money, prices, GDP, and dividends are given from history. To get a different sample it is necessary to re-run history, which is obviously impossible. The way statistic theory deals with this situation is to reverse the approach taken in classical statistic analysis, and build a model that describes the behavior of the observed data.
- Finally, from the view of economics, the subject of statistics deals mainly only based on the estimation and inference of covariances. On the other hand the econometrician must also give estimated parameters an economic interpretation. When it comes to time series, economic theory is an integrated part of the modeling process. Many econometric studies fail because researchers assume that their estimates can be given an economic interpretation without considering the statistical properties of the model, or the simple fact there is in general not a one to one correspondence with observed variables and the concepts defined in economic theory.

B. Application of Time Series in Medical Domains

Medical time series usually are combined with other methods like CBR-Case-based reasoning and data mining. These synergies may be of the form of pre-processing for feature mining from time series [30] or for retrieval of cases involving temporal features. Recent trends in applying CBR in medicine involve the representation of time series data for the purpose of reason about patient progress over time [31]. In medical domains important is to determine the transition of behavior over time rather than draw conclusions based on the absolute values in time series.

A characteristic example of combination of case-based decision-support system and time series data has been applied for diagnosing respiratory sinus arrhythmia (RSA) based on sensor readings [32]. The data mining facilities of this system allow for time-series classification and pattern identification. Following the case-based reasoning paradigm, classification of time series is performed by observing similar time series. In particular, the classifier examines sequences of patterns and their co-occurrence with class labels. The method is experimentally evaluated with a clinical case library, showing that it improves the classification accuracy.

Similar approach [30] the challenge of analyzing time series generated from lengthy and complex nature of signals is examined. In this approach a set of qualitative, interpretable features from original and usually real-valued time series data has been extracted. Features are generated by transforming the time series of real numbers into a symbolic series by temporal abstraction or symbolic approximation. Classification is then performed on such symbolic sequences and authors examine four alternative ways to index time series cases using discovered key sequences.

In [33] authors extend mentioned approaches by developing a hybrid case-based reasoning system for the application of stress diagnosis, which is based on time-series measurements in the form of signals and textual medical meta-data. To address the noise in the signals system uses a fuzzy similarity searching procedure. Fuzzy similarity aims at more accurate similarity estimates compared to traditional distance-based similarity measures for time series like Discrete Time Warping. More details on approaches in this direction can be found in [33].

C. Large-Scale Astronomical Time Series Databases

One of contemporary and modern areas where time series analysis can play important role are different areas of astronomy and astrophysics [34]. Time-domain astronomy (TDA) is rapidly changing faced by the exponential growth of the sample size, data complexity and data generation rates of new astronomical sky surveys. For example, the Large Synoptic Survey Telescope (LSST), which will begin operations in northern Chile in 2022, is planned to generate more than 150 Petabyte imaging dataset of the southern hemisphere sky. The LSST is intended to improve understanding of time-varying astrophysical objects, and also to reveal a plethora of yet unknown faint and fast-varying phenomena [34]. So this new field of data-driven astronomy introduced new challenging areas: astroinformatics and astrostatistics. These new data-oriented paradigms encompass significant disciplines like statistics, data mining and machine learning, knowledge discovery, and computational intelligence. The role of time series analysis here is of great significance for using robust methods needed for the rapid detection and classification of known astrophysical objects as well as the unsupervised characterization of novel phenomena.

Photometry as a part of astronomy is aimed to the precise measurement of visible electromagnetic radiation from astronomical objects. Several techniques and methods are applied to transform the raw astronomical data into standard units of flux or intensity. The basic tool in the analysis of astronomical brightness variations is the light curve. A light curve is a plot of the magnitude of an object's electromagnetic radiation expressed as a function of time. This light curve analysis is challenging, because of the sheer size of the databases, and also due to the characteristics of the data itself. Astronomical time series are unevenly sampled due to constraints in the observation schedules, telescope allocations and other limitations. The sampling is randomized because observations for each object happen at different times every night under different weather conditions. Discontinuities in light curves can also be caused by different technical factors. Astronomical time series are also affected by several noise sources.

In general, errors in astronomical time series are non-Gaussian and heterocesdastic, i.e., the variance of the error is not constant, and changes along the magnitude axis. Other common problematic situations arising in TDA are the sample-selection bias and the lack of balance between classes. Generally the astrophysical phenomena of interest represents a small fraction of the observable sky, hence the vast majority of the data belongs to the "background class". This is especially noticeable when the objective is to find unknown phenomena, a task known as novelty detection. Sufficient coverage and exhaustive labeling are required in order to have a good representation of the sample, and to assure capturing the rare objects of interests and highly promising application of time series analysis.

V. FAP – Framework for Analysis and Prediction

All these concepts, when introduced, are usually separately implemented and presented in different publications. Every newly-introduced representation method or distance measure has claimed a particular superiority. However, this was usually based on comparison with only a few other representatives of the proposed concept. On the other hand, to the best of our knowledge there is no freely available software system for time-series analysis and mining which supports all important concepts, with the exception of the work proposed in [17]. Being motivated by these observations, we have designed a multipurpose, multifunctional system FAP – Framework for Analysis and Prediction [35]. FAP supports all important concepts and activities connecter to time series data like: different representations, applications of wide range of similarity measures and pre-processing tasks; with the possibility to easily change some existing or to add new concrete implementation of any other concept.

FAP is mainly designed for time-series researchers in order to test and compare existing and newly introduced time-series concepts, methods and models. However, it can also be useful to professionals from different domains not familiar with time-series analysis as an assistance tool for choosing appropriate methods for their own time series data sets. FAP is freely available and can be downloaded from http://perun.dmi.rs/fap/.

The Framework for Analysis and Prediction has been already successfully employed within various research domains including: investigation of the influence of global constraints on distance measures [28, 36], developing a distributed distance matrix generator based on agents [37], mining time series in the psychological domain [38, 39], time-series analysis in the medical domain [40], and in financial forecasting [41].

VI. Conlusion

Time series are characteristic representatives of complex data. Analysis and applications of time series have become progressively more important in a variety of fields of research and real life problems, including business, economics, engineering, medicine, social sciences, environometrics, politics, and others. Also time series analysis, modeling and forecasting has fundamental importance in these practical domains. In the paper we briefly presented several contemporary, innovative, and important models and discussed their possible applications.

References

[1] Adhikari R, Agrawal R. K. (2013) An Introductory Study on Time Series Modeling and Forecasting, arXiv:1302.6613v1.

[2] Zhang G.P. (2003) Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 50: 159–175

[3] Cochrane J.H. (1997) Time Series for Macroeconomics and Finance. Graduate School of Business, University of Chicago.

[4] Box G.E.P., Jenkins G. (1970) Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco, CA.

[5] Kamruzzaman J, Begg R, Sarker R. (2006) Artificial Neural Networks in Finance and Manufacturing, Idea Group Publishing, USA.

[6] Zhang G., Patuwo B.E., Hu M.Y. (1998) Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting 14:35-62.

[7] Hamzacebi C. (2008) Improving artificial neural networks' performance in seasonal time series forecasting. Information Sciences 178: 4550-4559.

[8] Farooq T., Guergachi and Krishnan S. (2007) Chaotic time series prediction using knowledge based Green's Kernel and least-squares support vector machines. Systems, Man and Cybernetics. ISIC. 7-10 Oct. 2007: 373-378.

[9] Raicharoen T., Lursinsap C., Sanguanbhoki P. (2003) Application of critical support vector machine to time series prediction. Circuits and Systems. ISCAS '03.Proceedings of the 2003 International Symposium on Volume 5, 25-28 May, pages: V-741-V-744.

[10] J.A.K. Suykens and Vandewalle J. (2000) Recurrent least squares support vector machines. IEEE Trans. Circuits Systems-I, 47(7):1109 – 1114.

[11] Hand DJ, Mannila H, Smyth P (2001) Principles of Data Mining. MIT Press, Cambridge, MA, USA

[12] Laxman S, Sastry PS (2006) A survey of temporal data mining. Sadhana 31:173–198

[13] Esling P, Agon C (2012) Time-series data mining. ACM Computing Surveys 45:12:1–12:34.

[14] Faloutsos C, Ranganathan M, Manolopoulos Y. (1994) Fast subsequence matching in time-series databases. In: Proceedings of the ACM SIGMOD International Conference on Management of Data; Minneapolis, Minnesota, USA: 419–29.

[15] Han J, Kamber M, Jei P. (2012) Data mining: Concepts and techniques. 3rd ed. Morgan Kaufmann Publishers.

[16] Keogh, EJ. (2006) A decade of progress in indexing and mining large time series databases. In: Proceedings of the 32nd International Conference on Very Large Data Bases (VLDB); Seoul, Korea. p. 1268.

[17] Ding H, Trajcevski G, Scheuermann P, Wang X, Keogh EJ. (2008) Querying and mining of time series data: Experimental comparison of representations and distance measures. Proceedings of the VLDB Endowment. 1:1542–52.

[18] R. Agrawal, C. Faloutsos, and A. Swami. (1993) "Efficient similarity search in sequence databases," in Proceedings of the 4th International Conference on Foundations of Data Organization and Algorithms, pp. 69–84.

[19] K.-p. Chan and A. Wai-chee Fu. (1999) "Efficient time series matching by wavelets," in Proceedings of the 15th International Conference on Data Engineering, pp. 126–133.

[20] E. Keogh, K. Chakrabarti, M. Pazzani, and S. Mehrotra. (2001) Dimensionality reduction for fast similarity search in large time series databases. Knowledge and Information Systems, vol. 3, no. 3, pp. 263–286.

[21] E. Keogh and M. Pazzani, (1998) "An enhanced representation of time series which allows fast and accurate classification, clustering and relevance feedback," in Fourth International Conference on Knowledge Discovery and Data Mining (KDD'98), pp. 239–241.

[22] Q. Chen, L. Chen, X. Lian, Y. Liu, and J. Xu Yu, (2007) "Indexable pla for efficient similarity search," in Proceedings of the 33rd International Conference on Very Large Data Bases, pp. 435–446.

[23] J. Lin, E. Keogh, L. Wei, and S. Lonardi, (2007) "Experiencing sax: a novel symbolic representation of time series," Data Mining and Knowledge Discovery, vol. 15, no. 2, pp. 107–144.

[24] D. J. Berndt and J. Clifford, (1994) "Using dynamic time warping to find patterns in time series," in KDD Workshop, pp. 359–370.

[25] M. Vlachos, G. Kollios, and D. Gunopulos, (2002) "Discovering similar multidimensional trajectories," in Proceedings of 18th International Conference on Data Engineering, pp. 673–684.

[26] L. Chen, M. T. Ozsu, and V. Oria, (2005) "Robust and fast similarity search for moving object trajectories," in Proceedings of the 2005. ACM SIGMOD International Conference on Management of Data, pp. 491–502.

[27] L. Chen and R. Ng, (2004) "On the marriage of lp-norms and edit distance," in Proceedings of the 30th International Conference on Very Large Data bases, Volume 30, pp. 792–803.

[28] V. Kurbalija, M. Radovanovic, Z. Geler, and M. Ivanovic, (2014) "The influence of global constraints on similarity measures for time-series databases," Knowledge-Based Systems, vol. 56, pp. 49–67.

[29] Bo Sjö (2011), Lectures in Modern Economic Time Series Analysis, Editia a doua, Universitatea din Linköping, Suedia.

[30] Funk, P., Xiong, N. (2008) Extracting knowledge from sensor signals for case-based reasoning with longitudinal time series data. In: Perner, P. (ed.) Case-Based Reasoning in Signals and Images, pp. 247–284. Springer, Heidelberg

[31] Shahar Y., Miksch S., Johnson P. (1998) The Asgaard project: a taskspecific framework for the application and critiquing of time-oriented clinical guidelines. Artif Intell Med:14:29—51.

[32] Funk, P., Xiong, N. (2006) Case-Based Reasoning and Knowledge Discovery in Medical Applications with Time Series. Computational Intelligence: 238–253

[33] Begum S., Ahmed M. U., Funk P., Xiong N., Folke M.. (2011) Case-Based Reasoning Systems in the Health Sciences: A Survey of Recent Trends and Developments. IEEE Transactions on Systems, Man, and Cybernetics--Part C: Applications and Reviews

[34] Huijse P., Estevez P.A., Protopapas P., Prıncipe J.C., Zegers P. (2015) Computational Intelligence Challenges and Applications on Large-Scale Astronomical Time Series Databases. arXiv:1509.07823v1 [astro-ph.IM] 25 Sep.

[35] V. Kurbalija, M. Radovanovic, Z. Geler, and M. Ivanovic, (2010) "A framework for time-series analysis," in Proceedings of the 14th International Conference, AIMSA, pp. 42–51.

[36] Z. Geler, V. Kurbalija, M. Radovanovic, and M. Ivanovic, (2015) "Comparison of different weighting schemes for the knn classifier on time-series data," Knowledge and Information Systems, pp. 1–48.

[37] Mitrovic D, Ivanović M, Geler Z (2014) Agent-Based Distributed Computing for Dynamic Networks. Information Technology And Control 43:88–97.

[38] Kurbalija V, Ivanović M, von Bernstorff C, et al (2014) Matching Observed with Empirical Reality– What you see is what you get? Fundamenta Informaticae 129:133–147.

[39] Kurbalija V, von Bernstorff C, Burkhard H-D, et al (2012) Time-series mining in a psychological domain. In: Proceedings of the Fifth Balkan Conference in Informatics on - BCI '12. ACM Press, New York, New York, USA, pp 58–63.

[40] Kurbalija V, Radovanović M, Ivanović M, et al (2014) Time-series analysis in the medical domain: A study of Tacrolimus administration and influence on kidney graft function. Computers in Biology and Medicine 50:19–31

[41] Kurbalija V, Ivanović M, Budimac Z, (2009) Case-based curve behaviour prediction. Software: Practice and Experience 39:81–103.