Expert Systems With Applications 55 (2016) 194–211

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Computational Intelligence and Financial Markets: A Survey and Future Directions

Rodolfo C. Cavalcante[a,b,*], Rodrigo C. Brasileiro[b], Victor L.F. Souza[b], Jarley P. Nobrega[b], Adriano L.I. Oliveira[b]

[a] Núcleo de Ciências Exatas, Universidade Federal de Alagoas, Arapiraca 57309005, AL, Brasil
[b] Centro de Informática, Universidade Federal de Pernambuco, Recife, PE, Brasil

ARTICLE INFO

Keywords:
Financial markets
Computational intelligence
Trading systems
Deep learning
Online learning

ABSTRACT

Financial markets play an important role on the economical and social organization of modern society. In these kinds of markets, information is an invaluable asset. However, with the modernization of the financial transactions and the information systems, the large amount of information available for a trader can make prohibitive the analysis of a financial asset. In the last decades, many researchers have attempted to develop computational intelligent methods and algorithms to support the decision-making in different financial market segments. In the literature, there is a huge number of scientific papers that investigate the use of computational intelligence techniques to solve financial market problems. However, only few studies have focused on review the literature of this topic. Most of the existing review articles have a limited scope, either by focusing on a specific financial market application or by focusing on a family of machine learning algorithms. This paper presents a review of the application of several computational intelligent methods in several financial applications. This paper gives an overview of the most important primary studies published from 2009 to 2015, which cover techniques for preprocessing and clustering of financial data, for forecasting future market movements, for mining financial text information, among others. The main contributions of this paper are: (i) a comprehensive review of the literature of this field, (ii) the definition of a systematic procedure for guiding the task of building an intelligent trading system and (iii) a discussion about the main challenges and open problems in this scientific field.

1. Introduction

Financial markets are very important for the economical and social organization of modern society. Financial activities play an important role in the world economy, since they influence the economic development of several countries worldwide (Lin, Chiu, & Lin, 2012). In these kinds of markets, the success of an investor depends on the quality of the information he uses to support the decision-making, and on how fast he is able to take decisions. Due to its practical importance, the analysis of financial market movements has been widely studied in the fields of finance, engineering and mathematics in the last decades (Yoo, Kim, & Jan, 2005). Recently, some statistical and soft computing mechanisms have been proposed to provide support to the decisions of investors in different financial market segments.

Two approaches commonly used to analyze and predict financial market behaviors are (i) the fundamental analysis and (ii) the technical analysis. The former approach studies the economic factors that may influence market movements, and it is best suited for a longer term prediction spectrum. The technicians, on the other hand, believe that the price already includes all the fundamentals that affect it. In this sense, technicians usually model the historical behavior of a financial asset as a time series, believing that the history tends to repeat itself (Murphy, 1999). This modeling approach avoids the analysis of those subjective economic factors.

Financial time series prediction can be considered one of the main challenges in the time series and machine learning literature (Tay & Cao, 2001). In the last decades, several approaches have been proposed to predict financial time series and to provide decision-making support systems (Teixeira & Oliveira, 2010). Two major classes of work which attempt to forecast financial time series are the statistical models and machine learning approaches (Wang, Wang, Zhang, & Guo, 2011). Traditional statistical methods

* Corresponding author at: Núcleo de Ciências Exatas, Universidade Federal de Alagoas, Arapiraca 57309005, AL, Brasil. Tel.: +558234821829.
E-mail addresses: rodolfo.cavalcante@arapiraca.ufal.br, rcc5@cin.ufpe.br (R.C. Cavalcante), rcb5@cin.ufpe.br (R.C. Brasileiro), vlfs@cin.ufpe.br (V.L.F. Souza), jpn@cin.ufpe.br (J.P. Nobrega), alio@cin.ufpe.br (A.L.I. Oliveira).

http://dx.doi.org/10.1016/j.eswa.2016.02.006
0957-4174/

generally assume that the time series under study are generated from a linear process (Kumar & Murugan, 2013), and try to model the underlying time series generation process in order to make predictions about the future values of the series. However, financial time series are essentially complex, highly noisy, dynamic, nonlinear, nonparametric, and chaotic in nature (Si & Yin, 2013).

Machine learning techniques have been applied with relative success in modeling and predicting financial time series (Lee, 2009). Many machine learning techniques are able to capture nonlinear relationship between relevant factors with no prior knowledge about the input data (Atsalakis & Valavanis, 2009). Among these techniques, artificial neural networks (ANN) have been widely used in forecasting time series, since they are data-driven, self-adaptive methods able to capture nonlinear behaviors of time series without any statistical assumptions about the data (Lu, Lee, & Chiu, 2009; Tay & Cao, 2001). Due to these advantages, several types of ANNs and hybrid mechanisms have been used in forecasting financial time series (Atsalakis & Valavanis, 2009).

In the last years, some researches, motivated by the fundamental analysis, have investigated how to predict future market movements by mining informations in textual format. In this branch of research, financial news (Groth & Muntermann, 2011; Schumaker, Zhang, Huang, & Chen, 2012), financial reports (Wang, Huang, & Wang, 2012), and even information in micro-blogs (Ruiz, Hristidis, Castillo, Gionis, & Jaimes, 2012) are considered relevant source of informations for predicting future market behavior. These approaches investigated how mining important features on textual data or how to identify author sentiments in the news to improve the forecasting of future financial values.

The main purpose of this article is to review the recent computational intelligence approaches designed to solve financial market problems. Despite the high number of scientific papers which propose intelligent and expert systems, techniques and algorithms applied to financial markets, just a few review articles have been published in relevant journals and conferences in the last years. Most of the existing review articles have focused on a specific financial market application or on a family of machine learning algorithms. In this article, we make a comprehensive survey of machine learning methods applied to the financial context published from 2009 to nowadays. The scope of this review includes articles dealing with preprocessing and clustering of financial data, forecasting of future market movements, mining of financial information, among others. This survey not only summarizes the main primary studies, but also it identifies some challenges of applying machine learning techniques to solve financial problems and discusses some open problems and promising future researches in this field.

The remainder of this paper is organized as follows. Section 2 discusses the research method used in this survey and some other surveys related to this work. Section 3 describes some basic concepts which represent a background for this research. In addition, a methodology for constructing an intelligent trading system is proposed in this section. Section 4 provides a description of the main primary studies selected for be reviewed. These primary studies were grouped according to their main objectives. Section 5 provides a general discussion about these main primary studies. Section 6 provides a discussion about the challenges and open problems of this research area. Section 7 provides the final considerations of this paper.

2. Research methodology and related work

The primary studies discussed in this survey were recovered from Science Direct, Google Scholar and IEEExplore digital libraries for the period from 2009 to 2015. This publication year criterion was adopted because we observe a lack of literature review articles with the same scope as the research presented in this paper. Three main research questions were formulated to guide this research: (i) What kind of financial market problems have been solved by computational intelligence algorithms?; (i) What kind of computational intelligence algorithms have been proposed to solve these financial problems?; (iii) What are the main challenges and research opportunities that remain open in this research field?. In this survey, we have used combinations of the following keywords: “financial markets”, “stock markets”, “portfolio” “financial time series”, “forecasting”, “machine learning”, “computational intelligence”, “neural networks”, “support vector machines”, “extreme learning machines”, “fuzzy systems”, “novelty detection”, “clustering”, “feature selection”, “text mining”, “survey”, “review”, “comparative study”. The criterion to include the primary studies in this survey was that the article should propose the use of any computational intelligence technique to solve a financial problem. Articles that propose an expert system for automatic negotiation had priority of inclusion.

It is important to highlight that the results of this review do not include all applications of computational intelligence methods to financial market problems. Papers not published in reputable journals and conferences or with poor description of the intelligent method used, of the learning algorithm, or with a weak evaluation method were not included in this survey.

Each selected article was organized and discussed according to its primary goal into three main categories and some subcategories:

1. Preprocessing: 1.1 feature selection and extraction; 1.2 denoising and outlier detection; 1.3 time series segmentation; 1.4 clustering.
2. Forecasting: 2.1 artificial neural networks; 2.2 support vector machines; 2.3 hybrid methods; 2.4 optimization methods; 2.5 ensemble methods; 2.6 others approaches.
3. Text mining.

We started the survey by examining review articles that cover applications of computational intelligence methods to financial market problems. In the literature, there are not many review articles that cover this topic. Most of the published surveys found in the literature are specialized in some computational intelligence technique, such as ANNs or genetic algorithms (GA). Vanstone and Finnie (2009) proposed an empirical method for developing automatic trading systems using ANN. In that study, the authors reviewed some work that use fundamental and technical analysis to construct intelligent trading systems. They also described some key steps in using ANN for this purpose, namely the selection of inputs and outputs, the determination of the neural architecture and how to process the output signals produced by an ANN to create rules to enter and exit trades.

Li and Ma (2010) surveyed the application of neural networks in several subareas of financial markets. They enumerated some primary studies that apply ANN to exchange rates forecasting, stock market forecasting, and prediction of banking and financial crisis. In that short paper, no details of which ANN architectures or learning strategies were used in the primary studies surveyed. The work proposed by Roshan, Gopura, and Jayasekara (2011) surveyed the use of ANN in forecasting financial time series. In that paper, the authors surveyed the literature under two perspectives: the preprocessing mechanisms applied to input data and the neural network design used in primary studies. Soni (2011) also surveyed the application of ANNs to stock market prediction. That paper provides a brief history of ANNs and basic concepts of stock market before enumerating some primary studies that use this approach to solve the stock market forecasting problem. Despite the fact that these reviews have an intersection with the period of our research, they have a limited scope, since they focus in just one family of computational intelligence algorithms.