

Chapter 10

Evaluation of Alternative Approaches in Classification Algorithms for Prediction of Stock Market Index: Case of Crobex

Silvija Vlah Jerić

 <https://orcid.org/0000-0003-4738-6337>

Faculty of Economics and Business, University of Zagreb, Croatia

ABSTRACT

This chapter tackles the problem of automatic recognition of favorable days for intra-day trading. The problem is modeled as a binary classification problem, and several approaches are tested for solving it. Croatian stock index CROBEX data is used and 22 technical indicators are calculated as predictor variables. Performance of five classifiers is evaluated and compared by using Cohen's kappa as evaluation metric: artificial neural network, support network machine, random forest, k-nearest neighbors, and naïve Bayes classifier. The results give insight to effectiveness of technical analysis in predicting the day favorability for CROBEX index and suggest that technical analysis makes sense and might work for this case.

INTRODUCTION

Stock market index is often used to describe the aggregate movements in a market as it represents a measurement of value of a section of the stock market and reflects the level of the free float value. Since its changes involve a stochastic component, it is very difficult to predict and forecasting stock market index value has always been one of the hottest topics in research. One of the ways of analyzing stock movement is technical analysis which consists of using past market information to predict the direction of prices.

DOI: 10.4018/978-1-7998-5083-0.ch010

The problem of automatic recognition of favorable days for intra-day trading can be modeled as a binary classification problem. Day trading consists in buying and selling financial instruments within the same trading day and favorable for intra-day trading means that the increase between the opening price and the closing price of the same day is large enough for obtaining a profit by buying at the opening price and selling at the closing price. Thereby “large enough” is defined by selecting a certain threshold. The underlying classification problem consists of attribution of one of two labels to daily records according to values of a selection of technical indicators. The class that is assigned to each daily record is 1 if the subsequent day is favorable for intra-day trading and 0 otherwise. The problem is to discover the relation between the selected indicators at a certain day and the market situation at the following day that would determine the class of day that is observed, according to technical analysis (Lo & Hasanhodzic, 2010). The objective, as in any classification problem, is to achieve the highest accuracy possible when comparing the prediction with the real class.

The goal of this chapter is to evaluate and compare different classification algorithms on automatic recognition of favorable days for intra-day trading. Classification problems has been intensively studied and there are many classification algorithms available today (Bruni & Bianchi, 2015; Hastie, Tibshirani, & Friedman, 2001). Different families of algorithms are tested: neural network, support vector machine, random forest, as well as k -nearest neighbors and naïve Bayes classifier as more typical classifiers. The idea is to test different families of algorithms but to give more attention to algorithms which are used more rarely than typical statistical methods which are still considered to be “mainstream” although is a lot of research in alternative approaches nowadays too. The algorithms will be applied to Croatian stock index CROBEX and it is the first attempt of such kind to the best of authors’ knowledge and should fill some gap in this area of research observed in literature (Šego & Škrinjarić, 2018). The results will give insight to effectiveness of technical analysis in predicting the index value, as well to observe the accuracy of different classification methods.

Background

There are many research efforts in the literature on different topics regarding stock market prediction. Some of them address regression problems by trying to predict stock market prices, while others treat them as classification problems, for example, by trying to predict the direction of changes in prices. Besides classical statistical approaches, there are many examples of using machine learning algorithms in stock market prediction problems, most often neural networks. One of the first efforts was by Kimoto, Asakawa, Yoda and Takeoka (1990) for predicting the index of Tokyo stock market by neural networks. Mizuno, Kosaka, Yajima and Komoda (1998) also used neural networks to predict the trade of stocks in Tokyo stock market. Also, Bayes classifiers (Shin & Kil, 1998; Tsaih, Hsu, & Lai, 1998; Pop, 2006) and SVM (Ince & Trafalis, 2007; Moreira, Jorge, Soares, & Sousa, 2006) have been used by researches for financial market prediction. Kara and Boyacioglu (2011) discussed stock price index movement using two models based on neural networks and support vector machine. Qi and Zhang (2008) investigated the best model of time series using neural network. Zemke (1999) compared neural network, k -nearest neighbor, naïve Bayesian classifier, and genetic algorithm, developing classification rules for their prediction accuracies on stock exchange index data. Qian and Rasheed (2007) investigated the predictability of the Dow Jones Industrial Average index using neural network, decision tree, and k -nearest neighbor.

In this chapter five algorithms are compared in terms of performance in predicting Croatian stock index CROBEX. Since the idea was to test different families of methods, as mentioned before, the following

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/evaluation-of-alternative-approaches-in-classification-algorithms-for-prediction-of-stock-market-index/260902

Related Content

Assessing Systemic Risk

Ian I. Mitroff and Abraham Silvers (2016). *International Journal of Risk and Contingency Management* (pp. 66-75).

www.irma-international.org/article/assessing-systemic-risk/152164

Recidivism Reduction Research

(2020). *Community Risk and Protective Factors for Probation and Parole Risk Assessment Tools: Emerging Research and Opportunities* (pp. 15-23).

www.irma-international.org/chapter/recidivism-reduction-research/242046

Artificial Intelligence, Terrorism, and Cyber Security: Challenges and Opportunities

Nika Chitadze (2025). *Machine Intelligence Applications in Cyber-Risk Management* (pp. 87-106).

www.irma-international.org/chapter/artificial-intelligence-terrorism-and-cyber-security/363124

Genetically-Modified Crop Risk Assessment and Biosafety: Integrative Meta Analysis and Communication Strategy

Mohammed Bashir Tijjani, Adamkolo Mohammed Ibrahim and Abubakar Tijjani Ibrahim (2026). *Mitigation Strategies and Risk Management for Food Contaminants* (pp. 243-290).

www.irma-international.org/chapter/genetically-modified-crop-risk-assessment-and-biosafety/389868

Assessing Systemic Risk

Ian I. Mitroff and Abraham Silvers (2016). *International Journal of Risk and Contingency Management* (pp. 66-75).

www.irma-international.org/article/assessing-systemic-risk/152164