Research on Data Mining and Investment Recommendation of Individual Users Based on Financial Time Series Analysis

Shiya Wang, Drexel University, Philadelphia, USA

ABSTRACT

With the continuous development of financial information technology, traditional data mining technology cannot effectively deal with large-scale user data sets, nor is it suitable to actively discover various potential rules from a large number of data and predict future trends. Time series are the specific values of statistical indicators on different time scales. Data sequences arranged in chronological order exist in our lives and scientific research. Financial time series is a special kind of time series, which has the commonness of time series, chaos, non-stationary and non-linear characteristics. Financial time series analysis judges the future trend of change through the analysis of historical time series. Through in-depth analysis of massive financial data, mining its potential valuable information, it can be used for individual or financial institutions in various financial activities, such as investment decision-making, market forecasting, risk management, customer requirement analysis provides scientific evidence.

KEYWORDS
Arima Model, Data Mining, Data Prediction, Financial Time Series, Smoothing

INTRODUCTION

The computerization of society has significantly enhanced the ability of people to generate and collect data. Huge amounts of data are pouring out from every corner of our lives. The explosive growth of stored or transient data has spurred the need for new technologies and automated tools to help us intelligently convert massive amounts of data into useful information and knowledge, leading to the emergence of a frontier discipline in computer science called data mining. Data mining refers to the automatic extraction of unknown but valuable representative knowledge from massive, uncertain, incomplete, fuzzy and noisy data sets. These patterns are hidden in large databases, data warehouses, other large amounts of information base or data stream. Data mining can be regarded as the process of discovering the existing relationships and rules, association patterns and trend prediction from massive data sets (Qua et al., 2013). Data mining can discover potential patterns and find out individual user information ignored by enterprise decision makers, so that it can be easily understood and clearly reflected to decision makers and provide reference for enterprise decision-making. Data mining is the mining of data with potential value of information, and this information has implicit, previously unknown, non-trivial, meaningful features. Data mining is a high-level process for identifying...
effective, novel, potentially useful and ultimately understandable patterns from data sets. It includes data cleaning, data integration, data selection and transformation, data mining, pattern evaluation and knowledge representation, and applies various methods to discover implicit rules and patterns from data sequences (Durante et al., 2014).

In the financial market, information continuously affects the market price changes. It is important and difficult to make the price trend judgment quickly. In practical application, it is often necessary to find the possible correlations between different financial time series. This kind of relationship has important reference value for people to understand the interaction of various financial time series more thoroughly and make reasonable decisions accordingly (Adhikari & Agrawal, 2014). Time series analysis has become an indispensable part of financial market research and is one of the important methods of financial quantitative analysis. Many research results in financial markets are based on time series analysis (Gui et al., 2015). Up to now, the importance of time series analysis has been widely recognized in the world. Financial market is a huge system, which is affected by many factors, and has a very complex movement law. Time series data is its comprehensive external manifestation. Because the data in the financial market is mainly time series, financial market analysis is often called financial time series analysis. Financial time series is the real measurement and record of the development data of financial variables in reality, and then through the quantitative analysis and summary of the corresponding mathematical statistics methods, we can find the inherent laws and behavioral characteristics of the financial market, thus providing important judgment basis for investment decision-makers. It is the theoretical basis of capital asset pricing and financial risk prevention.

Financial time series analysis is guided by financial theory, based on the principles and methods of time series analysis, taking financial data as information carrier, explaining the law of development and change of financial time series, explaining the internal structure of financial markets and their mutual influence, and then studying the risk prevention and avoidance in financial markets. Financial time series analysis considers the theory and practice of asset value evolution with time. It is a highly empirical subject. This is an important characteristic of financial time series analysis and general time series analysis. Financial time theory and its corresponding empirical time series contain an important factor in reality: uncertainty, for example, volatility of asset returns has a variety of definitions, volatility is not directly observed and so on. It is precisely because of these uncertainties in financial time series that statistical theory and methods play an important role in the analysis of financial time series. Financial time series is a special kind of time series, which has the commonness of time series, chaos, non-stationary and non-linear characteristics (Marszalek & Burczyński, 2015). There are many related researches, techniques and methods on time series analysis, including feature analysis, signal decomposition, clustering, pattern mining, similarity measurement and mathematical modeling. To sum up, the application of time series analysis methods and techniques in the process of financial information generated by a large number of financial time series in-depth analysis, fully mining the hidden information of the financial market, and market forecasting is the current stage, but also a realistic demand for a long time in the future and has important economic significance.

The rest of this paper is organized as follows. Section 2 discusses the flow of data mining technology is discussed., followed by time series analysis method in Section 3. The simulation test of data mining technique and time series analysis method is discussed. in Section 4. Section 5 concludes the paper with summary and future research directions.

**DATA MINING OVERVIEW**

**Data Mining Process**

The first step of data mining task is to define the purpose of mining, and then preconditioning the original data. When the cleaning work is completed, the corresponding mining algorithm can be
Contextualized Clustering in Exploratory Web Search
www.igi-global.com/chapter/contextualized-clustering-exploratory-web-search/10182?camid=4v1a

Query Recommendations for OLAP Discovery-Driven Analysis
Arnaud Giacometti, Patrick Marcel, Elsa Negre and Arnaud Soulet (2013). *Developments in Data Extraction, Management, and Analysis* (pp. 66-90).
www.igi-global.com/chapter/query-recommendations-olap-discovery-driven/70793?camid=4v1a