Identifying Hidden Patterns Related to Indices in CSE through Data Mining

S.M.M.P. Sathkumara, S.H.D. Senanayake
Department of Computer Science and Technology, Uva Wellassa University, Badulla, Sri Lanka.

The linkage between stock market operations and the economic growth has been tested by most of the economists in various countries. The key of success in stock trading is to buy and sell stocks at the right time for the right price. "Buy Low, Sell High" sounds easy, but it is so difficult to carry out since the direction of stock market in the near future is almost unpredictable. With the advances in data mining, it has now become possible to predict the future market direction based on historical data. By this research, I have tested & evaluated the performance of the stock market operations in Sri Lankan context using more than ten years data from 2003 to 2015. The main purpose of this research is identifying hidden patterns related to price indices in Colombo Stock Exchange (CSE) through data mining. In here I have used following attributes to achieve the main objectives of the research, All Share Price Index(ASPI) and twenty sector indices in CSE. The Price Indexes were used which are composite representation for whole stock market operations. Those twenty sectors are Banks, Finance & Insurance, Beverage, Food & Tobacco, Construction & Engineering, Chemicals & Pharmaceuticals, Diversified holdings, Footwear & Textiles, Hotels & Travels, Health care, Investment trusts, Information Technology, Land & property, Manufacturing, Motors, Oil palms, Power & Energy, Plantations, Stores & Supplies, Services, Telecommunication, Trading. The WEKA software is used to analyse the data. The empirical findings of this research finally gave three highest percentages among twenty sectors with the twelve high percentages. I have tged 90 percent as the threshold value to find out those fifteen outputs. This study will be helpful to future researches and the potential stock market investors.

Keywords: Stock market, Data mining, Classification, Pattern recognition, Decision tree