

An Automated Stock Investment System Using Machine Learning Techniques: An Application in Australia

Authors : Carol Anne Hargreaves

Abstract : A key issue in stock investment is how to select representative features for stock selection. The objective of this paper is to firstly determine whether an automated stock investment system, using machine learning techniques, may be used to identify a portfolio of growth stocks that are highly likely to provide returns better than the stock market index. The second objective is to identify the technical features that best characterize whether a stock's price is likely to go up and to identify the most important factors and their contribution to predicting the likelihood of the stock price going up. Unsupervised machine learning techniques, such as cluster analysis, were applied to the stock data to identify a cluster of stocks that was likely to go up in price – portfolio 1. Next, the principal component analysis technique was used to select stocks that were rated high on component one and component two – portfolio 2. Thirdly, a supervised machine learning technique, the logistic regression method, was used to select stocks with a high probability of their price going up – portfolio 3. The predictive models were validated with metrics such as, sensitivity (recall), specificity and overall accuracy for all models. All accuracy measures were above 70%. All portfolios outperformed the market by more than eight times. The top three stocks were selected for each of the three stock portfolios and traded in the market for one month. After one month the return for each stock portfolio was computed and compared with the stock market index returns. The returns for all three stock portfolios was 23.87% for the principal component analysis stock portfolio, 11.65% for the logistic regression portfolio and 8.88% for the K-means cluster portfolio while the stock market performance was 0.38%. This study confirms that an automated stock investment system using machine learning techniques can identify top performing stock portfolios that outperform the stock market.

Keywords : machine learning, stock market trading, logistic regression, cluster analysis, factor analysis, decision trees, neural networks, automated stock investment system

Conference Title : ICMCSSE 2018 : International Conference on Mathematical, Computational and Statistical Sciences and Engineering

Conference Location : Sydney, Australia

Conference Dates : March 29-30, 2018