

A Performance Comparison of Statistical and Machine Learning Techniques in Learning Time Series Data

Abstract

The task of analyzing and forecasting time-series data is very crucial task as this Time Series Analysis (TSA) task is used for many applications such as Economic Forecasting, Sales Forecasting, Budgetary Analysis, Stock Market Analysis, Yield Projections, Process and Quality Control, Inventory Studies, Workload Projections, Utility Studies, Census Analysis, Network Monitoring and Analysis and many more. The techniques used can be classified into two categories, namely statistical and machine learning techniques. As a result, the selection of several prediction methods will continue to be an alternative for researchers to obtain more accurate prediction results. This paper outlines and presents the comparison of predictive performance of statistical and machine learning techniques, namely ARIMA, Back-Propagation Neural Network (BPNN), and genetic algorithms (GA) for analyzing and predicting short-term time series network traffic activity datasets. In other words, this paper examines the forecasting performance of ARIMA, BPNN and GA models for the time series data related to network traffic activity data which is obtained from the ICT Universitas Mulawarman. The performances of these techniques are compared based on the errors measured, namely Mean Squared Error (MSE). Based on the results obtained, BPNN is found to be very efficient in learning a time series data. This paper is concluded by recommending some future works that can be applied in order to improve the prediction accuracy.