NEURAL NETWORK BASED HYBRID PREDICTION MODELS FOR HEALTHCARE APPLICATIONS

ABSTRACT

Prediction models based on different concepts have been proposed in recent years. Improving the accuracy of prediction models has remained as a challenging task for researchers. The prediction accuracy depends not only on the model but also on the complexity of the data. Hence, it is important to choose the best model based on the complexity of data in the prediction.

The time series prediction model determines future trends based on past values and corresponding errors. There are many models used for time series prediction, such as statistical techniques including linear regression, moving average, exponential smoothing, autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) and soft computing methods, such as artificial neural networks and adaptive neuro-fuzzy inference system.

In the development of time series analysis, it is well known that many phenomena are nonlinear, namely the relationship between the past and current events is nonlinear. Thus, the linear time series models are not sufficient and appropriate for these cases. As a consequence, nonlinear time series models have gained importance in time series prediction. The accuracy rates obtained using linear time series models may not high as they have limitations in handling the non-linear relationships among the data. Models based on artificial neural networks model are considered to be better in handling such non-linear relationships.

In the real-world, the time series data consist of complex linear and nonlinear patterns and it may be difficult to obtain high prediction accuracy rates using only linear or neural network model. Hybrid model which combines both linear and neural
network models provides a better solution for time series prediction. The hybrid model can be expected to yield high prediction accuracy results.

As the accuracy of prediction model depends on the pattern of the data, it is essential to determine the appropriate models for the linear and non-linear parts of the data when developing a hybrid model. The hybrid models proposed so far by different authors do not take into account the pattern of data and they merely combine some known linear and non-linear models to design the hybrid models. In this study, the enhanced hybrid models are proposed which into account the pattern of data in selecting the best linear model and also in optimizing the configuration of the neural network. Simulation results show that these models yield better results compared to known hybrid models for univariate time series prediction.

An improved adaptive neuro-fuzzy inference system (ANFIS) is also proposed for time series prediction. This model combines neural network and fuzzy inference system. Unlike the known methods where the number of input lags is arbitrarily fixed or decide using trial and error experiments, the proposed method make use of a systematic approach for determining the optimum number of input lags for obtaining the best accuracy results.

One of major problems in designing predictive models for multivariate data (i.e data containing many attributes) is the selection of appropriate attributes that will contribute to the final output. It is possible that some attributes can even obscure the results. The selection of independent variables for prediction is an important aspect in data analysis. It is possible to get better accuracy results with reduced number of attributes. In this study, an enhanced method to improve the accuracy for the healthcare prediction is proposed using attribute selection optimization algorithm and machine learning methods.

Our primary research aim is to develop hybrid methods for univariate time series prediction and to propose multivariate prediction method which integrates
attribute selection algorithm with machine learning for improving the prediction accuracy. The objectives of this dissertation are as follows:

(i) To identify and evaluate the performance of the existing statistical, soft computing techniques and hybrid models and to propose an enhanced hybrid model by combining linear and neural network models to improve the prediction accuracy for healthcare univariate time series prediction, (ii) To propose a strategy for the hybrid model to determine whether a linear-nonlinear combination or a nonlinear-linear combination is best suited for the input data, (iii) To enhance the existing adaptive neuro-fuzzy inference system by incorporating a novel technique to determine the optimum input lags for obtaining best accuracy results, (iv) To design a dual hybrid model which can be used as a decision support system in healthcare applications, (v) To integrate attribute selection algorithm and machine learning methods for improving the prediction accuracy and (vi) To compare the performance of the proposed methods with the known methods.

The experimental results show that all the proposed methods perform better than the other models. For measuring the performance accuracy, two indicators, namely, the root mean square error (RMSE) and the mean absolute percentage error (MAPE) for prediction are used.