

A Novel Hybrid Ensemble Model to Predict FTSE100 Index by Combining Neural Network and EEMD

Bashar Al-hnaity, Maysam Abbod

Abstract— Prediction stock price is considered the most challenging and important financial topic. Thus, its complexity, nonlinearity and much other characteristic, single method could not optimize a good result. Hence, this paper proposes a hybrid ensemble model based on BP neural network and EEMD to predict FTSE100 closing price. In this paper there are five hybrid prediction models, EEMD-NN, EEMD-Bagging-NN, EEMD-Cross validation-NN, EEMD-CV-Bagging-NN and EEMD-NN-Proposed method. Experimental result shows that EEMD-Bagging-NN, EEMD-Cross validation-NN and EEMD-CV-Bagging-NN models performance are a notch above EEMD-NN and significantly higher than the single-NN model. In addition, EEMD-NN-Proposed method prediction performance superiority is demonstrated comparing with the all presented model in this paper, and was feasible and effective in prediction FTSE100 closing price. As a result of the significant performance of the proposed method, the method can be utilized to predict other financial time series data.

I. INTRODUCTION

In recent years the ability of prediction accuracy is crucial to many decision makers in different processes such as planning, organizing, scheduling, selling and buying, strategy formulation, policy making and supply chain management. Hence the importance of reaching high accuracy in the prediction area has received high attention from researchers and extensive effort has been invested. Despite the huge interest and number of studies which has been conducted on financial time series prediction, this field remains important and is an active topic of research at the present time and the future [1]. Due to security issues in the stock market important financial bridges were developed between investors and listed companies in which the dynamic changes in the stock market directly reflect on stock market index. Many benefits can be obtained from studying stock indices such as providing a guidance reference for investors in order to make financial decisions, which additionally aids governments develop the correct rules to reduce the negative effects of stock market volatility and promote the stable development of stock markets where stock indices closely develop the national economy [2].

In econometrics, forecasting, modelling and trading financial indices (financial time series) are considered among of the most challenging and difficult topics. That can be a result of the financial time series characteristics: nonlinear, non-stationary, noisy, high degree of uncertainty and hidden relationships. These characteristics are a reason of the information unavailability, affecting the behaviour of financial markets between past and future captured prices. According to [3] traditional linear methods and the majority of sophisticated non-linear machine learning models have failed to capture the

complexity and the nonlinearities that exists in financial time series, particularly during uncertainty periods such the credit crisis time in 2008 [4].

Prediction financial time series are of considerable practical interest. Fundamental and technical are the most common approaches of stock price prediction. Where technical approach is based on historical financial time series data, the fundamental approach is based on financial status and performance of the company [5]. The ability to mine valuable information from a mass of historical data is the main inherent advantage of artificial neural networks and it's efficient used in financial areas. According to [6], [7], [8], [9], [10], [11], [12] ANN applications have been widely used and are lately popular in forecasting/predicting financial time series. Moreover the ANN can prevent problem domains of stock market prediction by its ability of learning and generalizing the nonlinear data trends. An ANN application achieves higher accuracy than traditional methods because of its ability to adapt to the data patterns and relationships between the input and output [13], [14].

Failure to utilize single artificial techniques to capture the non-stationary property and accurately describe the moving tendencies that exist in financial time series, which are due to its fluctuation and the dynamic changes in the relationship between independent and dependent variables. Such fluctuation and structural changes, which are often caused by political events, economic conditions, traders' expectations and other environmental factors, are important characteristics of financial time series. This study adopts a novel hybrid intelligent prediction ensemble model based on neural network back propagation and ensemble empirical mode decomposition. EEMD performs the EMD over an ensemble of the signal plus Gaussian white noise in order to prevent any mixing problem by populating the whole time-frequency space and get the advantage of the dyadic filter bank behaviour of the EMD [14],[15],[16]. Thus EMD was developed by [17] to analyze non-stationary time series data, which adaptively represents its local characteristic.

The proposed EEMD will adaptively decompose the complicated raw data into a residue and a verity number of intrinsic mode functions (IMFs), high frequency and low-frequency sequences and trend series. ANN back propagation model is chosen to predict each IMF and hence the final prediction would be the sum of each equal weighted prediction values. To verify the effectiveness of the proposed model, FTSE100 index closing price is selected to be the experimental example.

This paper is structured as follows: In the methodology section 2 the theory of ANN, AR and the proposed EEMD are briefly present. In section 3 experimental results are discussed. Experimental results on FTSE100 predictions are described in section 4. The last section concludes this paper.

Bashar AL-hnaity is with Electronic and Computer Engineering Department, Brunel University, UB8 3PH, London, UK (Tel:+447794439165; e-mail: bashar.al-hnaity@brunel.ac.uk).

Maysam Abbod is with Electronic and Computer Engineering Department, Brunel University, UB8 3PH, London, UK (e-mail: maysam.abbod@brunel.ac.uk).

II. METHODOLOGY

A. Artificial Neural Network

Artificial neural networks (ANN) are a nonlinear statistical technique. The ability of generalizing result from unseen data for actively system on actual time basis is the main advantage of utilizing ANN. ANN models are comprised of densely interconnected adaptive processing units. Pattern recognition and classification are the most well-known tasks neural network perform in. The features behind ANN capability of forecasting problems are as follows:

- First feature, the few prior assumptions are needed, which are the important traits of NN data-driven, where NN by learning from the training data can capture the logic or relationship, even the unknown ones.
- Second feature, the ability of generalization after the training process to update its parameters (weights and bias), thus predictions on unknown data can be made.
- Third feature, according to [8], two layer NN with enough hidden neurons can approximate any continuous function with satisfactory accuracy.
- Fourth feature, the majority of real system and prediction models are nonlinear. As NN is nonlinear it can be advantageous in forecasting.

In this paper, BP feed forward neural network has been utilized. Back propagation is the most neural network structures commonly utilized initially proposed by Runelhart and McClelland in 1985. BPNN is a multilevel feed forward neural network with strong mapping capabilities. BPNN comprises of an input layer, one or several hidden layers and an output layer [19]. In BPNN the process of learning and training is constituted by forward propagation of input result and back propagation of error result. In order to improve the overall error to a required precision the study should be repeated where $E < \varepsilon$.

ε is the target accuracy, E is sample error where $E = \frac{1}{2} \sum_{k=1}^m (y^k - c^k)^2$, y^k is the desired output of NN, c^k is the actual output of NN, m is the length of sample. The below figure (1) illustrate the BPNN architecture.

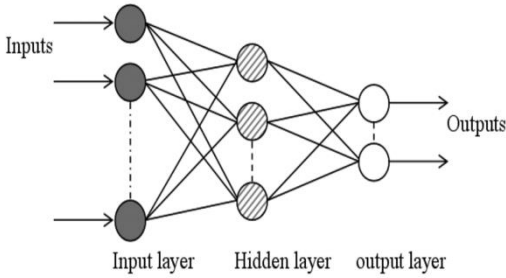


Figure 1. A three layer BPNN

B. Simple Auto-regressive model

Autoregressive model (AR) in this study is used as a benchmark model to evaluate the prediction power between the utilized models based on the relative improvements on root mean square error. The bellow function illustrate AR model used.

$$y(t) = a_1 y(t-1) + a_2 y(t-2) + \dots + a_n y(t-n) \quad (1)$$

In the above equation $y(t)$ is the predicted stock price based on the past close daily price, $y(t)-y(t-n)$ and the coefficients of AR model are a_1-a_n . 5 lagged daily price is the order of which is used in the AR model was varied and found to give better prediction result. The model coefficients were determined by using the Implemented “regress” function in the MATLAB.

C. Empirical Mode Decomposition

Empirical mode decomposition (EMD) is a method, which is characterized by the ability of analyzing non-linear and non-stationary signals. EEMD perform the EMD over an ensemble of the signal plus Gaussian white noise in order to prevent any mixing problem by populating the whole time-frequency space and get the advantage of the dyadic filter bank behaviour of the EMD [19].

EMD transformation mechanism of the signal $\chi(t)$ is defined by the decomposition of the signal into a small number of intrinsic mode functions (IMFs) or modes. There are two conditions the signal must satisfy in order to be considered as an IMF: (i) the number of extreme value and the number of zero crossing in the whole data set must either equal or differ at most by one. (ii) Everywhere the mean value of the upper and lower envelope must be zero. As the above point defines the IMFs, the decomposition processes of the signal are explained by the following steps [17]:

- First step for any signal $\chi(t)$ the local maxima and minima will be identified. Where all maxima and minima will be connected to produce upper and lower envelop by a cubic spline curve. m Is defined as the mean value of upper and lower envelop and the difference between $\chi(t)$ and m defined as h :
$$h = \chi(t) - m \quad (2)$$
- h Will be taken as a new original signal $\chi(t)$ and then the operation will be repeated in step (a) κ times until h is an IMF. To judge whether h is an IMF or not the below function is considered as a termination criterion:

$$\mathcal{D}_k = \frac{\sum_{t=0}^T |h_{(k-1)}(t) - h_k(t)|^2}{\sum_{t=0}^T |h_{(k-1)}(t)|^2} \quad (3)$$

In the above function \mathcal{D}_k is smaller than a predetermined value, h_k can be viewed as an IMF. Where, the first IMF will designate as $c_1 = h_k$.

- When c_1 is determined, the residue r_1 can be obtained by separating c_1 from the rest of the data as follows:

$$r_1 = \chi(t) - c_1 \quad (4)$$

After obtaining r_1 , the operation will be repeated in step (a) and (b) as r_1 is the new signal $\chi(t)$, until obtaining the second IMF c_2 . Moreover to the above, time j will be considered until r_j is smaller than a predetermined value or r_j becomes a monotone function in order to get all IMFs. A series of IMFs and a residue r in the end will be obtained after applying the above steps.

EMD is facing many drawbacks; mode mixing is one of the main drawbacks which usually cause intermittency of analyzing signal. This drawback can imply either a single IMF consisting of signals of dramatically disparate scales or a signal of the same scale appearing in different IMF components. To prevent such problems a new noise-assisted data analysis method EEMD was proposed. In EEMD, the ensemble of trails mean is the true IMF component. A decomposition result of the signal

plus a white noise of finite amplitude are consisting of each trail [15].

There are various white noise studies, which showed that the EMD method is an effective self-adaptive dyadic filter bank when applied to white noise, demonstrating the benefits on EEMD method [16],[17]. In EEMD IMFs components are defined as the mean of corresponding IMFs obtained via EMD over an ensemble of trials, which are generated by adding different realization of white noise of finite variance to the original signal $x[n]$. The below point describe the EEMD algorithm:

- a) White noise series add to the original signal. In other words generate $x^i[n]=x[n] + w^i[n]$, where $w^i[n](i = 1, \dots, I)$ are different realization of which Gaussian noise,
- b) Decompose the signal with added white noise into IMFs by EMD. Each $x^i[n](i = 1, \dots, I)$ is fully decomposed by EMD getting their modes $IMF_k^i[n]$, where $k=1 \dots K$ indicates the modes.
- c) Repeating steps (a) and (b) for certain number of times with different white noise each time and obtain corresponding IMF components of the decomposition.
- d) Calculate the mean of all the corresponding IMF components and take the mean as the final result for each IMF. In another words assign IMF_k as the k -th mode of $x[n]$, obtained as the average of the corresponding IMF_k^i : $IMF_k[n] = \frac{1}{I} \sum_{i=1}^I IMF_k^i[n]$.

D. Ensemble Empirical Mode Decomposition with Adaptive Noise

As above point explained how observing operation in the EMD, each is $x^i[n]$ decomposed independently from other realizations and so for each one a residue $r_k^i[n] = r_{k-1}^i[n] - IMF_k^i[n]$ is obtained. This proposed method by [17] presented, the decomposition mode will be noted as \widehat{IMF}_k and unique first residue propose as:

$$r_1[n] = x[n] - \widehat{IMF}_1[n], \quad (5)$$

In the above equation $\widehat{IMF}_1[n]$ is obtained in the same way as in EEMD. After that, using first EMD over an ensemble of $r_1[n]$ plus different realization of a given noise to obtain \widehat{IMF}_2 by averaging. Then, next residue is defined as: $r_2 = r_1[n] - \widehat{IMF}_2[n]$. Until the stopping criterion is reached the previous procedure steps with rest of the modes will be continued. This method can be described as following algorithm if $x[n]$ is the targeted data.

- 1) To obtain the first modes and compute, decompose by EMD I realization $x[n] + \epsilon_0 w^i[n]$.
 $\widehat{IMF}_1[n] = \frac{1}{I} \sum_{i=1}^I IMF_1^i[n] = \overline{IMF}_1[n]$.
- 2) In the first stage ($k = 1$) calculate the first residue as in Eq. (5):
 $r_1[n] = x[n] - \widehat{IMF}_1[n]$.
- 3) Decompose realizations $r_1[n] + \epsilon_1 E_1(w^i[n])$, $i = 1 \dots I$ until their first EMD mode and define the second mode:

$$\widehat{IMF}_2[n] = \frac{1}{I} \sum_{i=1}^I E_1(r_1[n] + \epsilon_1 E_1(w^i[n]))$$

- 4) For $k = 2 \dots K$ calculate the $k - th$ residue:

$$r_k[n] = r_{(k-1)}[n] - \widehat{IMF}_k[n]. \quad (6)$$

- 5) Decompose realizations $r_k[n] + \epsilon_k E_k(w^i[n])$, $i = 1 \dots I$, Until their first EMD mode and define the $(k+1)$ -The mode as:

$$\widehat{IMF}_{(k+1)}[n] = \frac{1}{I} \sum_{i=1}^I E_1(r_k[n] + \epsilon_k E_k(w^i[n])) \quad (7)$$

- 6) Then for next k go to step 4.

Steps 4 to 6 are performed until the obtained residue is no longer feasible to be decomposed. The final residue satisfies:

$$R[n] = x[n] - \sum_{k=1}^K \widehat{IMF}_k, \quad (8)$$

With K the total number of modes is complete. Therefore, the given signal $x[n]$ can be expressed as:

$$x[n] = \sum_{k=1}^K \widehat{IMF}_k + R[n]. \quad (9)$$

Eq. (9) makes the proposed decomposition complete and provides an exact reconstruction of the original data. At each stage an observable selection of SNR is allowed by ϵ_i coefficients. It has been suggested by Wu and Hung [14] that regarding the amplitude of the added noise, small amplitude values for data dominated should be used by high-frequency signals and vice versa [19].

E. Ensemble learning models

The research on finding the best methodologies and systems to combine multiple predictive models has become very active. These systems or methodologies are called ensemble methods by machine learn community [20]. Ensemble learning of classifiers is a paradigm of classifiers set whose individual decisions are combined in a different number of ways to classify new examples. Comparing ensemble methods with single a classifier the main goal of utilizing ensemble methods is to generate from the original data set a diversified learn and train data sets, which performs better than an individual classifier. Accurate and diverse classifiers must be the in order of ensemble learning classifiers to be more accurately compared to individual classifier [21]. In this case being accurate means better than random guessing and diverse means the classifiers make uncorrelated errors. According to many studies [22] utilizing ensemble learning methods have often shown a higher accuracy in practice than the classifiers they are composed off. In financial markets 50% is the usual prediction accuracy, thus using ensemble methods can be very useful in prediction the financial market, since 1% improvement can bring large pecuniary benefits. Moreover to the improvement of the performance, and having a final stable system which combines several different classifiers results can maintain functionality even if one underlying classifier fails. for successful investment decision system and in financial markets, the need of consistency system is very important.

There are two models, first models, which have been derived from different executions of the same learning algorithm, are called homogeneous. Bagging and boosting are the main techniques for constructing ensembles of homogeneous and they are considered as the most famous [23], [24]. Second models, which have been derived from running different learning algorithms on the same data set, are called heterogeneous. The most effective methods, which combine the result of heterogeneous classifiers, are voting and stacking [24].

In this study and in order to get single output, majority weighted vote, weighted average and average mean have been adopted to combine the result of ensemble models. Several ensemble strategies have been used to ensure diversity of data training among classifiers such cross-validation, bagging and CV-bagging.

i) *Cross-validation (CV)*

It is the simplest and widely used method among sample variation methods of training data [24]. Dividing the data into several k-folds partitions and in each partition using random data generated from the original data set is the main objective of utilizing cross-validation. As the below figure shows in every partition of cross-validation, the original data set is divided in training and a validation set where each k-fold acts like an independent validation set for the model trained by k-folds. It has been noticed data dependency is reduced and the consistency of results improved by utilizing cross-validation model [25].

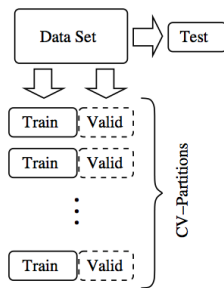


Figure 2. For every partition of the cross-validation, the data is divided in training and Validation set

ii) *Bagging*

Many scholars define it as a shortcut for bootstrap aggregation and it is the first ensemble learning algorithms [26]. The basic role of implementing such a model and the perfect performance result are considered the main advantages of using bagging model. In order to achieve trained data diversity, creating a number of bags in which each bag is randomly filled by data generated from the original data set without replication. In another words each bag has a randomly selected training data set and all bags trained are one classifier. The below methods have been proposed to ensure that the most diversification of data is trained to the best of our knowledge. This might be the first time these methods are introduced and utilized in financial time series predictions (stocks / indices predictions).

iii) *CV-Bagging*

This method is a result of combing the cross-validation model with the bagging model. The following points illustrate the way of how this proposed methods works:

- Dividing training data into k partitions where each partition generates data randomly from the training data set as CV techniques mention above.
- From each partition a data set will be selected randomly into different number of bags according to bagging techniques.
- After the above point NN model will be applied for training the proposed method. For instance if there are 15 partitions and 100 bags in each partition, in

total there will be 1500 neural networks trained. This is explained later explained later in this paper

F. *Combining classifiers*

The employed strategy in combining classifiers is one of the key components of any ensemble system. There are two groups of combination rules, first group: trainable vs. non-trainable combination rules. Second group: combination rules that apply to class labels vs. to class-specific continuous outputs. In combination rules, there are different ways of determining their parameters. For instance parameters in trainable combination rules are usually called weights and they are determined by separating the training algorithm. EM algorithm is one of the key algorithm paradigms, which is used by mixture-of-experts model. The determined combination parameters which are created by trainable rules are usually instance specific. As a reason of that they are also called dynamic combination rules. In non-trainable rules, separate training is not involved conversely beyond that used for generating the ensembles.

MLP and RBF networks are classifiers models providing continuous-valued outputs. Despite their sufficiently dense training data requirement which is rarely met in practice, they are often interpreted as posterior. In this study and due to space restriction, the author only reviews the adopted methods [27].

i) *Combining continuous outputs*

Each given classifier class provides a continuous output, which is interpreted as the degree of support given for that class additionally, this continuous output is usually accepted as an estimate of the posterior probability for that class. Algebraic combiners are in general non-trainable combiners of continuous outputs. They are a function of the support received by individual classifiers; total support for each class will be obtained [28].

ii) *Algebraic combiners*

There are different types of combination function in algebraic combiners such as: mean rule, weighted average, trimmed mean, minimum/maximum/median rule, and product rule and generalized mean. This study adopts the weighted average combination function [27]. This kind of function can be defined as a hybrid of the mean and the weighted majority voting, in this case the only different than weighted majority voting is weights are not applied to class label; they are applied to the actual continuous outputs. Depending on how the weights are obtained this combination rule can qualify either as a trainable or non-trainable. This all depends on the way the weights are obtained. During ensemble generation as it happened in Adaboost, if weights are obtained as part of the regular training, it is a non-trainable combination rule. On the other hand it is a trainable combination rule when the weights are obtained by using separate training. It is very usual that every classifier has its weight, or in another cases for each class and for each classifier [26].

III. HYBRID ENSEMBLE PREDICTION MODEL

Utilizing past values and another variables, time series predicts the future. Financial time series has a certain non-stationary, nonlinear characteristics. These characteristics imply that the statistical distributions of time series can change over time. The cause of these changes may be caused by economic

fluctuations, or political and environmental events. As a result, it is difficult to use any single artificial technique to capture financial time series characteristics property and accurately describe its moving tendency, which lead to different and not accurate financial time series prediction result. In order to address these issues, a novel hybrid ensemble model based on neural network and complete ensemble empirical mode decomposition with adaptive noise is proposed as follow:

- At this stage the proposed EEMD has been used to decompose the original stock market index to several intrinsic mode functions.
- In the second stage NN has been used to predict IMFs respectively and also NN ensemble model is applied to predict IMFs. Combining method was additionally applied to combine different NN ensemble models for each IMF.
- Final prediction value can be obtained by the sum of above final predicted result. Weighted average combination function was utilized to combine the different NN ensemble methods. An optimiser Genetic algorithm, has been used to determine the weight of the combiner. The below fig.3 illustrated the overall process of the proposed approach for stock index prediction.

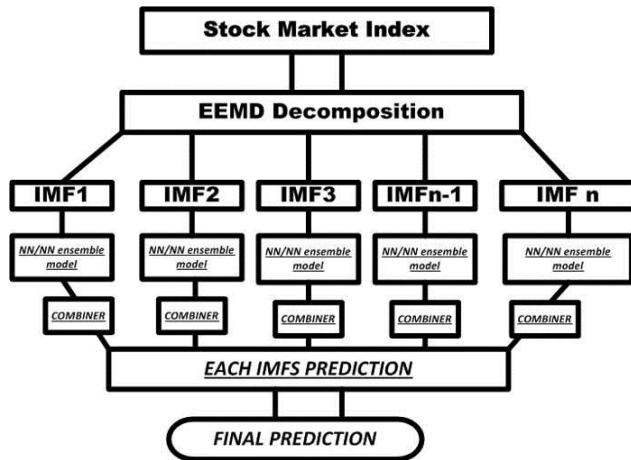


Figure 3. The process of NN_EEMD model

Applying the proposed method can more fully capture the local fluctuations of the original data. Therefore, proposed method decomposed the original data into different IMFs where each IMF has simpler frequency components and stronger regularity comparing with the original data. Furthermore EEMD reduce the complexity and improve the efficiency and accuracy of the prediction model.

IV. EMPIRICAL ANALYSIS

This study examined FTSE100 index. The data were collected from London stock Exchange market. Closing daily price was used. Data set covers a period of 02/04/1984 to 28/02/2014 (including 7557 points). In this study the time series is sufficient for testing the issue of non-stationary properties of financial data.

A. Time series prediction and Data preproesing

Utilizing the single NN, EEMD-NN, EEMD-NN-Bagging, EEMD-NN-CV and EEMD-NN-CV-Bagging to process the time series, will demonstrate the superiority of the proposed model. In this study data set is divided into two sets, training

and testing. The first 7307 data is used as a training data set and the last 250 is used as a testing data set. In order to prevent problems associated with cumulative errors from previous period for out-of sample prediction, this study considers one-step-ahead prediction.

Selection and pre-processing the data are crucial step in any modelling effort, particularly for generalizing the new predictive model. Data sets are divided into two sets: training and testing, which are took a bellow forms: For training data: $Tr = [x_{(n)} \dots \dots x_{(n+5)}] [x_{(n+6)}]$, where m is a random number permutation $1 < m < N_{Tr}$, N_{Tr} is the data size. For testing data: $Ts = [x_{(n)} \dots \dots x_{(n+5)}] [x_{(n+6)}]$, where n is $1: N_{Ts}$, N_{Ts} is the testing data set size. Figure 4 below illustrate the architecture of the data prepossessing.

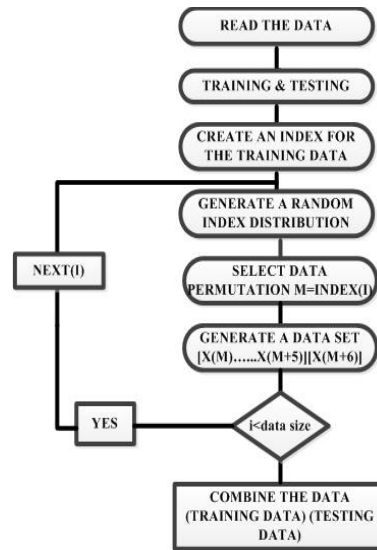


Figure 4. The data pre-processing frame

B. Decomposing FTSE100 index by EEMD

Using the EEMD method, which was firstly proposed by [26], in this study to decompose the FTSE100 index closing price to partition to 14 IMFs, as Fig. 5 is depicted the decomposition obtained by the utilized the introduced method. It is noticed that the difference of frequency between IMFs. The frequencies of IMF 1, IMF 2, IMF 3, IMF 4, IMF 5 and IMF 6 are much higher than the rest and they are reflecting the non-stationary and the randomness information of the original FTSE 100 closing price.

The periodic trends of IMF 7, IMF 8 and IMF 9 are obvious and are called the periodic components of the original FTSE 100 closing prices. From IMF 10 to IMF 14 they are called the trend components.

C. The empirical Result analysis and discussion

After FTSE100 closing price decomposed by EEMD into different numbers of IMFs, the pervious explained methods, in sections II part A and E, will be utilized to predict each IMF respectively as figuer3 illustrate. Firstly single BPNN is used to predict each IMF and the residue r . Each IMF was divided into training and testing data set as explained in section III part A.

It is very important to determine the architecture of the network topology. One input layer, two hidden layers and one output neuron are the selected topology test in this study. A range of 20-5 is the number of the neurons in both two hidden layer were determine after a trials and error test were conducted. After reaching the pre-determine number of epochs, NN will stop the learning (training) process. Based on lowest Mean Square Error the optimal topology has been selected.

Secondly as it explained on the methodology part section E, different ensemble techniques were used to predict each IMF. The figures 8, 9 and 10 show the result from utilizing the ensemble techniques: cross validation, Bagging, and CV-Bagging. Finally and after the final prediction result obtained by the sum of the each processed IMF by EEMD-NN-Bagging, EEMD-NN-CV and EEMD-NN-CV-Bagging, a Weighted average combination function was applied to combine them on one final result as shown in figure 11.

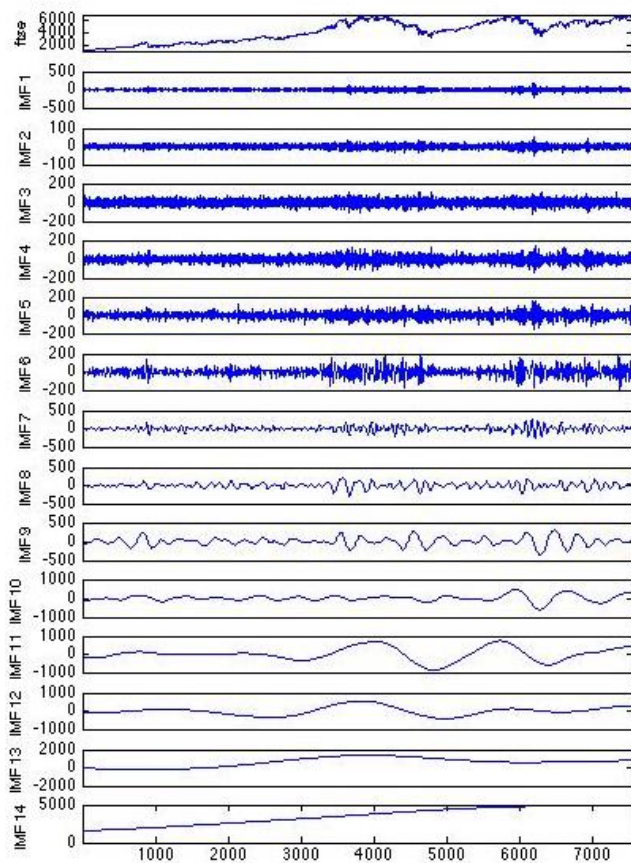
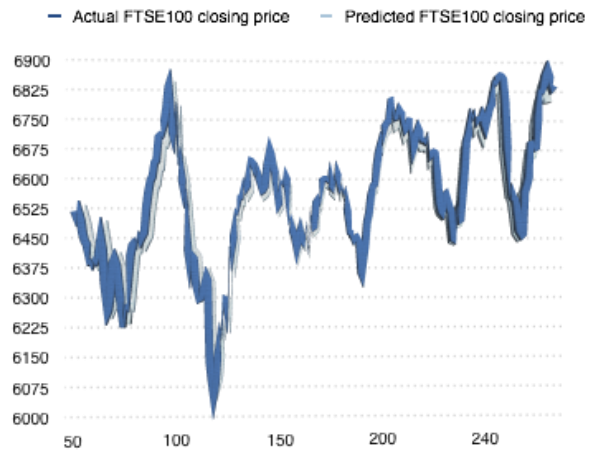
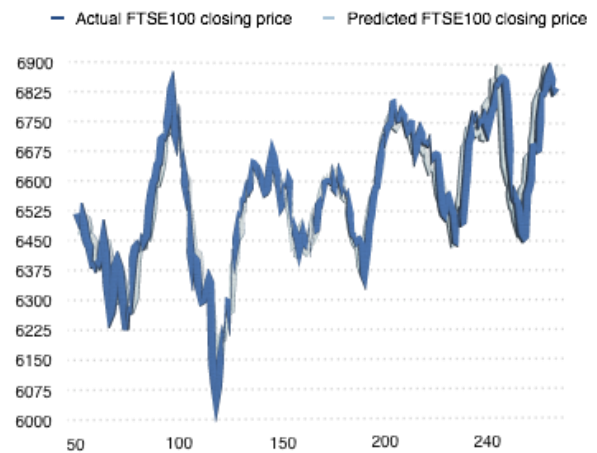


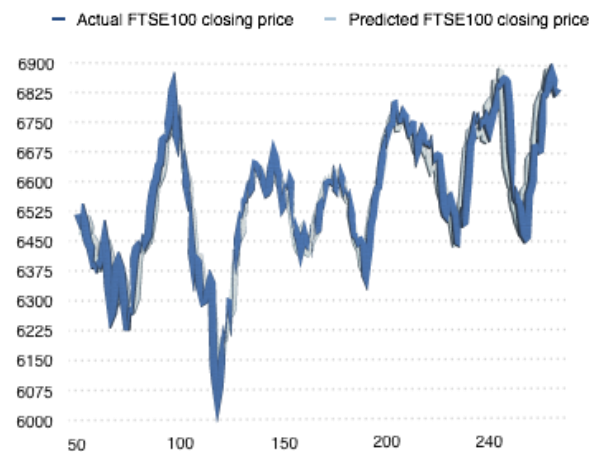
Figure 5. IMFs of FTSE100 index through EEMD



Data point
Figure 6. The result of FTSE 100 based on single NN Model



Data point
Figure 7. The result of FTSE100 based on EEMD and NN Model



Data point
Figure 8. The result of FTSE100 based on EEMD and CV/NN Model

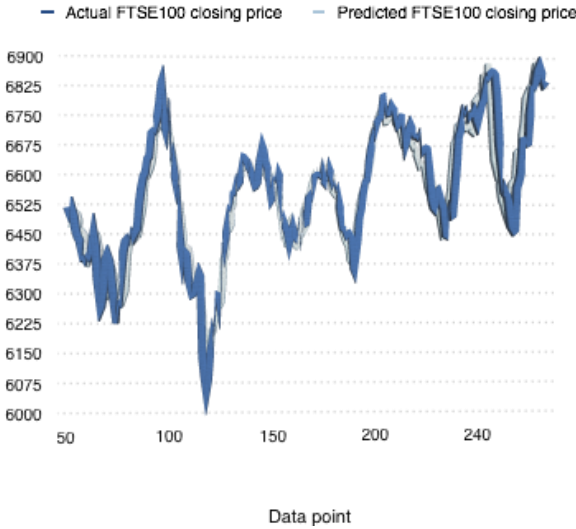


Figure 9. The result of FTSE100 based on EEMD and Bagging/NN Model

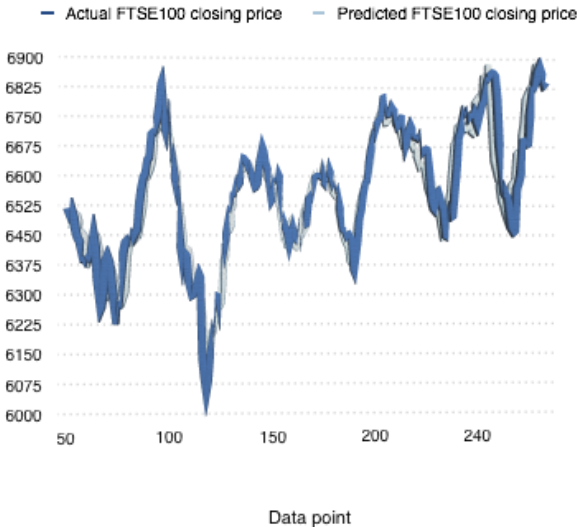


Figure 10. The result of FTSE100 based on EEMD and CV/Bagging/NN Model

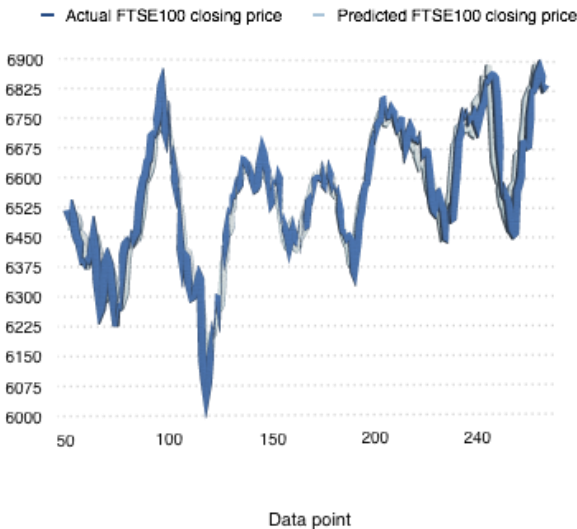


Figure 11. The result of FTSE100 based on proposed method EEMD/NN Model

D. Result Evaluation

There are many different performance measurement to evaluate the superiority of the single NN, EEMD-NN, EEMD-NN-Bagging, EEMD-NN-CV, EEMD-NN-CV-Bagging and EEMD-Proposed method models, mean squared error (MSE) and root mean squared error (RMSE) are utilized in this paper as a models performance measurement.

$$MSE = \frac{1}{N} \sum_{t=1}^N (x(t) - \hat{x}(t))^2 \quad (10)$$

$$RMSE = [\frac{1}{N} \sum_{t=1}^N (x(t) - \hat{x}(t))^2]^{1/2} \quad (11)$$

The experimental result indicate that EEMD-proposed method has the highest prediction accuracy than the single NN , EEMD-single-NN, EEMD-Bagging-NN, EEMD-Cross Validation-NN and EEMD-CV-Bagging as shown in Table I. That is because comparing ensemble methods with single classifier the main goal of utilizing ensemble methods is to generate from the original data set a diversified learn and train data sets, which performs better than an individual classifier techniques into one final output. After decomposing the original data by EEMD the unsteady and nonlinear characteristics of the data has been changed into components which have fixed frequency and periodicity. This is why the accuracy has been improved by utilizing the EEMD combined with ensemble methods as shown in Table I. The table presents the result of each model after they were ran ten times and the mean and standard deviation of these ten prediction errors is calculated for each model as shown bellow. Simple Autoregressive model (simple-AR) is used in this paper as a benchmark, where as the table I illustrate that all four hybrid models and the proposed model outperform the simple-AR model.

TABLE I. PREDICTION ERROR OF THE DIFFERENT MODELS

Model	Prediction Error RMSE	
	MEAN	STD
Single-NN	49.88	1.0974
Simple-AR	49.48	0.0
EEMD-NN	27.42	1.1826
EEMD-Bagging-NN	27.04	1.1788
EEMD-Cross Validation-NN	26.17	0.7169
EEMD-CV-Bagging-NN	26.85	0.7849
EEMD-NN-Proposed Method	25.31*	0.6265*

V. CONCLUSIONS

This paper studies the FTSE100 index closing price and successfully establishes a new hybrid ensemble prediction model. EEMD was introduced for the first time in modelling a combination strategy to predict stock market index (financial time series data) in this paper. In addition, combining ensemble methods with neural network and EEMD was proposed and applied in order to improve the prediction accuracy. Firstly, EEMD was used to decompose FTSE100 closing price into several IMFs. Secondly, in prediction part, we apply the different proposed methods as explained above in the methodology section and empirical analysis section in order to

obtain the final prediction result. In order to improve the prediction result, three ensemble methods Bagging, CV and CV-Bagging were adopted to combine EEMD-NN. In EEMD-Proposed method, weighted average has been used for combining EEMD-NN-Bagging, EEMD-NN-CV and EEMD-NN-CV-Bagging prediction result into one final prediction result, the GA is used to optimize the weight. In this study and as shown in table I using EEMD-NN, EEMD-NN-Bagging, EEMD-NN-CV, EEMD-NN-CV-Bagging and EEMD-NN-Proposed Method enhance the financial time series prediction performance significantly comparing with single NN. Thus, the prediction performance of the EEMD-NN-Proposed Method superiority is demonstrated by comparing it with Single NN, EEMD-NN, EEMD-NN-Bagging, EEMD-NN-CV and EEMD-NN-CV-Bagging precisely. Therefore, the EEMD-NN-Proposed Method in this paper is really an efficient model for FTSE100 closing price prediction.

REFERENCES

- [1] D. Zhang, Q. Jiang and X. Li, "Application of neural networks in financial data mining," *International Journal of Computational Intelligence*, vol. 1, pp. 116-119, 2004.
- [2] J. Yang and X. Dai, "Prediction of shanghai and shenzhen 300 index based on EMD-SVM model," in *Systems and Informatics (ICSAI), 2012 International Conference on*, 2012, pp. 1209-1212.
- [3] Y. Li and W. Ma, "Applications of artificial neural networks in financial economics: A survey," in *Computational Intelligence and Design (ISCID), 2010 International Symposium on*, 2010, pp. 211-214.
- [4] K. Theofilatos, A. Karathanasopoulos, P. Middleton, E. Georgopoulos and S. Likothanassis, "Modeling and trading FTSE100 index using a novel sliding window approach which combines adaptive differential evolution and support vector regression," in *Artificial Intelligence Applications and Innovations* Anonymous Springer, 2013, pp. 486-496.
- [5] P. Chang and C. Liu, "A TSK type fuzzy rule based system for stock price prediction," *Expert Syst. Appl.*, vol. 34, pp. 135-144, 2008.
- [6] K. A. Smith and J. N. Gupta, "Neural networks in business: techniques and applications for the operations researcher," *Comput. Oper. Res.*, vol. 27, pp. 1023-1044, 2000.
- [7] Y. S. Abu-Mostafa, A. Atiya, M. Magdon-Ismael, H. White and J. Racine, *Special Issue on Neural Networks in Financial Engineering*. IEEE, 2001.
- [8] A. Adebisi, C. Ayo, M. Adebisi and S. Otokiti, "An Improved Stock Price Prediction using Hybrid Market Indicators," *African Journal of Computing & ICT*, vol. 5, pp. 124-135, 2012.
- [9] M. Khashei, M. Bijari and G. A. Raissi Ardali, "Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks (ANNs)," *Neurocomputing*, vol. 72, pp. 956-967, 2009.
- [10] M. Khashei and M. Bijari, "An artificial neural network (p,d,q) model for time series forecasting," *Expert Syst. Appl.*, vol. 37, pp. 479-489, 2010.
- [11] E. Hajizadeh, H. D. Ardakani and J. Shahrabi, "Application of data mining techniques in stock markets: A survey," *Journal of Economics and International Finance*, vol. 2, pp. 109-118, 2010.
- [12] C. S. Vui, G. K. Soon, C. K. On, R. Alfred and P. Anthony, "A review of stock market prediction with artificial neural network (ANN)," in *Control System, Computing and Engineering (ICCSCE), 2013 IEEE International Conference on*, 2013, pp. 477-482.
- [13] Soni, "Applications of ANNs in stock market prediction: a survey," *International Journal of Computer Science & Engineering Technology*, vol. 2, pp. 71-83, 2011.
- [14] H. Yu and H. Liu, "Improved stock market prediction by combining support vector machine and empirical mode decomposition," in *Computational Intelligence and Design (ISCID), 2012 Fifth International Symposium on*, 2012, pp. 531-534.
- [15] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, pp. 1-41, 2009.
- [16] P. Flandrin, G. Rilling and P. Goncalves, "Empirical mode decomposition as a filter bank," *Signal Processing Letters, IEEE*, vol. 11, pp. 112-114, 2004.
- [17] S. Shi, W. Liu and M. Jin, "Stock price forecasting using a hybrid ARMA and BP neural network and markov model," in *Communication Technology (ICCT), 2012 IEEE 14th International Conference on*, 2012, pp. 981-985.
- [18] A. Chen, M. T. Leung and H. Daouk, "Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index," *Comput. Oper. Res.*, vol. 30, pp. 901-923, 2003.
- [19] Y. Zhang and L. Wu, "Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network," *Expert Syst. Appl.*, vol. 36, pp. 8849-8854, 2009.
- [20] Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 460, pp. 1597-1611, 2004.
- [21] M. E. Torres, M. A. Colominas, G. Schlotthauer and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, 2011, pp. 4144-4147.
- [22] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. Yen, C. C. Tung and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 903-995, 1998.
- [23] T. Ditterich, "Machine learning research: four current directions," *Artificial Intelligence Magazine*, vol. 4, pp. 97-136, 1997.
- [24] B. Qian and K. Rasheed, "Stock market prediction with multiple classifiers," *Appl. Intell.*, vol. 26, pp. 25-33, 2007.
- [25] R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*. John Wiley & Sons, 1999.
- [26] L. K. Hansen and P. Salamon, "Neural network ensembles," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, pp. 993-1001, 1990.
- [27] T. G. Dietterich, "Ensemble methods in machine learning," in *Multiple Classifier Systems* Springer, 2000, pp. 1-15.
- [28] L. Breiman, "Bagging predictors," *Mach. Learning*, vol. 24, pp. 123-140, 1996.
- [29] R. E. Schapire, Y. Freund, P. Bartlett and W. S. Lee, "Boosting the margin: A new explanation for the effectiveness of voting methods," *Annals of Statistics*, pp. 1651-1686, 1998.
- [30] M. Paliwal and U. A. Kumar, "Neural networks and statistical techniques: A review of applications," *Expert Syst. Appl.*, vol. 36, pp. 2-17, 2009.
- [31] S. L. Salzberg, "On comparing classifiers: Pitfalls to avoid and a recommended approach," *Data Mining and Knowledge Discovery*, vol. 1, pp. 317-328, 1997.
- [32] L. Breiman, "Bagging predictors," *Mach. Learning*, vol. 24, pp. 123-140, 1996.
- [33] R. Polikar, "Ensemble based systems in decision making," *Circuits and Systems Magazine, IEEE*, vol. 6, pp. 21-45, 2006.
- [34] L. I. Kuncheva, J. C. Bezdek and R. P. Duin, "Decision templates for multiple classifier fusion: an experimental comparison," *Pattern Recognit*, vol. 34, pp. 299-314, 2001.