Forex Prediction Engine: Framework, Modelling Techniques and Implementations

Abstract

This paper presents foreign exchange (Forex) prediction engine that included framework, modelling techniques and implementations, to support the needs of financial organizations or individual investors. In the financial sector, Forex prediction is considered to be a complex field, based on the noise of exchange rates as the major challenge. However, there are some financial instruments that are available to guide the individual for future investments. We propose a promising Forex prediction engine using historical Forex data, to extract a pattern movement over a period of time series using Linear Regression Line (LRL) technique and the proposed segmentation algorithm. Subsequently, Artificial Neural Network (ANN) algorithm is applied to classify unique groups of uptrend and downtrend patterns. Dynamic Time Warping (DTW) algorithm is implemented through brute force to identify the similarity trend and the matched result that is used to predict the trend movement for next day. This research reveals the structure of Forex prediction engine by the description of an integrated framework using technical analysis method and machine learning algorithm. The experimental results of $AUD - USD$ and $EUR - USD$ currencies demonstrated 71% of accuracy and its predictions are reliable.

Keywords: Forex Prediction Engine, Linear Regression, Artificial Neural Network, Dynamic Time Warping

1 Introduction

The Foreign Exchange (Forex) market is the largest and most liquid financial market in the world, with several trillion US Dollar (*USD*) traded every day (PFGBEST, n.d.). Exchange rates rarely represent the fair value in Forex market, due to variations in supply and demand from investors, and the influence of many correlated factors including economic, political and even psychological ones (King et al., 2011). The exchange rates in the Forex market are unrestricted – for example, a floating foreign exchange rate system was implemented by industrialised nations in 1973.

Forex prediction is a very complex and non-linear dynamic field in which the non-stationary and noise of exchange rates are the major challenges (Koedijk et al., 2006). This is because the exchange rate between one country and its partner is affected by export and import trading, the valuation of assets, interest rate differentials and speculation about future events. Furthermore, the exchange rates do not obtain complete information to predict the future trend since the future trading flows are unknown (Virginia, 2009). There are many factors that influence the progress of Forex market, such as general economic conditions, supply and demand of the exchange transactions in market. Therefore, the level of difficulty in predicting the exchange rates is high.

Nonetheless, there are some risks involved when investing in Forex trading due to the uncertain movement of exchange rates. Many investors have requested intelligent techniques in order to get better prediction on the exchange rates so that they would have higher chances in gaining profit from the investment.

Since the late 1970s, researchers and experts from different sectors have started to develop various kinds of prediction models such as technical analysis methods and Artificial Intelligence (AI) techniques. The common matter in Forex prediction is that researchers are unaware of the behaviour and trend patterns of price within the financial sectors that have utilised different technique analysis methods such as Candlestick Pattern (Nison, 1991), Linear Regression Line (LRL) (Rockefeller, 2011), Exponential Moving Average (EMA) (Dzikevičius and Šaranda, 2010) and Moving Average Convergence Divergence (MACD) (Tanaka-Yamawaki and Tokuoka, 2007) to predict the direction of financial market growth. In addition, those methods require a certain amount of data to identify the trend for prediction model. As a result, the historical data becomes vital information in the process of prediction, due to it being able to provide hidden information in representing the different kinds of financial trends. Thus, the

implementation of trend patterns has to be improved for the existing prediction model. This leads to the discovery of a unique cluster of trend patterns from the historical data.

The rapid growth of computer technology has created an ever constant evolution of AI techniques. These techniques are implemented for classification and prediction in the financial sector. Since the 1990s, researchers from the computer science sector have applied AI techniques such as Artificial Neural Network (ANN) (Lawrence, 1997; Naeini et al., 2010), Expert Systems (ES) (Lee and Jo, 1999), Support Vector Machine (SVM) (Kim, 2003), Hidden Markov Model (HMM) (Hassan and Nath, 2005) and Genetic Algorithms (GA) (Hassan et al., 2007) to learn trend patterns for prediction, and have proven the usefulness of these techniques in learning the trends patterns. In this paper, we combine Dynamic Time Warping (DTW) with other methods to predict the trend of currency exchange rates using one technical indicator Linear Regression Line (LRL), which is novel to the best of our knowledge.

The implementation of the LRL has been applied for analysing and forming general trend patterns of historical Forex data between the *USD* and other major currencies, particularly Australian Dollar (*AUD*) and Euro (*EUR*). ANN is then implemented to learn trend patterns that were analysed by the LRL. To predict future trends of the Forex market, DTW is used to measure the similarity of training trend patterns in the proposed prediction model. However, the result of experiments revealed that the selected features are very important in the learning process of ANN.

Section 2 briefly describes the background and related work of the existing prediction models. In section 3, the framework of the proposed prediction model is discussed. Section 4 shows the result of the LRL process in identifying the trend patterns, ANN training process and the DTW prediction result of the proposed model. A discussion and future work of the proposed prediction model conclude this paper.

2 Background and Related Work

2.1 Other Approaches for Prediction

Many researchers examine Forex prediction (Kamruzzaman and Sarker, 2004; Yao and Tan, 2000; Maknickienė and Maknickas, 2012). The common focus is the use of technical indicators as the feature vector to represent the trend (Dzikevičius and Šaranda, 2010; Janssen et al., 2012; Kuepper, 2012; Caginalp and Laurent, 1998). Common technical analysis methods used include Candlestick Pattern (Person, 2002; Shmatov, 2012), EMA (Edwards and Magee, 2007), MACD (Fernández-Blanco et al., 2008) and OHLC (Open, High, Low and Close) bar chart (Ge, 1998). These methods analyse trend signals based on the historical data (Achelis, 1995; Sewell, 2007; Grunder et al., 2012).

Several technical analysis methods have been used as feature extraction by researchers from the fields of computer science to identify financial trend patterns. Lee and Jo (1999) proved that candlestick patterns with financial trading rules was useful to develop an ES in predicting the financial trend movement. The limitation of their ES is that it lacked automated learning. Another approach was applied by Kardos and Cwiok (2011), utilising EMA and MACD as input features to train ANN. The experiment showed that the ANN had successfully recognised the signals for prediction. Nobakht *et al.* (2010) applied OHLC as input features for HMM to train, and concluded that OHLC was able to represent the financial trends for HMM to predict the future of trends.

According to Kamijo and Tanigawa (1990) and Goswamil *et al.* (2009), using the candlestick patterns as input features enable the enhancement of ANN in recognising the patterns for prediction. Yao and Tan (2000) proposed a combination of ANN with technical analysis methods in predicting exchange rates. They suggested that trading strategies could be considered as extra features to enhance the prediction model.

Zhai *et al.* (2007) introduced the combination of seven technical indicators with related news to analyse the stock trends for SVM to predict the direction of stock price. As the result of this experiment, the proposed prediction model was enhanced and provided 70% accuracy rate. Based on the experiments that were provided by

Rao and Hong (2010), it was shown that by using MACD and EMA to analyse data as input pattern for SVM to learn, the model can represent more accurate prediction results.

In the past 10 years, technical analysis has been combined with AI techniques in financial time series problems to improve multivariate prediction ability, which statistical methods lack (Klassen, 2005). In this paper, we combine DTW with one indicator to predict the trend of Forex market, which is novel to the best of our knowledge.

Financial time series can exhibit "time warping," i.e., nonuniform time compression or expansion on a time scale (Lebaron, 2006). Relatively few research programs explicitly address time warping. Coelho (2011) introduced a procedure where the DTW was used as a measure of similarity between two periods. Applying the DTW to volume, volatility and returns, Wong and Yeung (2008) provided a forecast to particular stocks and then computed its correlation with the real stock price. (The basic theory and details of both procedures were, however, not described in detail and no comparative data were given.)

Applying the DTW combined with LRL and ANN to technical indicators has, to our knowledge, never been attempted and we believe that using technical indicators has the advantage over raw data because relative trends are "clearer" with technical analysis. Using novel machine learning based method of combined algorithms, our Forex prediction engine uses the historical data to automate learning and model the financial trends movement for future prediction.

2.2 Linear Regression Line (LRL)

LRL is a statistical tool that uses the slope value of the regression line as identifier. A straight line is drawn through the time series data in order to identify the distances between the prices of different time stamps (Rockefeller, 2011). The slope value is used to identify two major trend patterns; a positive slope is defined as an uptrend whilst a negative slope is defined as a downtrend. The following equation defines a straight line which is used to identify the trends over the historical data:

$$y = mx + b, \qquad (1)$$

where y is defined as price, m is defined as slope, x is defined as the number of time frames and b defines as the y intercept.

Eq. (1) calculates a diagonal line across the time series data for identifying the financial trends in a certain time stamp. Financial experts concluded that LRL was able to identify a buying signal when the price is under the diagonal line, and when the price is above the line, it is considered a selling signal.

Traditionally, LRL has been applied to many real world situations and the model is easy to develop and implement. According to Zhang (2001), Ahangar *et al.* (2010) and Naeini *et al.* (2010), they have found that LRL is only able to identify the direction of current and past trend changes rather than being able to predict future trends directly. Few studies have addressed that LRL works well in identifying and describing trend patterns for prediction purposes (Rinehart, 2003; Olaniyi et al., 2011).

Our study utilised LRL to analyse the overall trends from the historical data, and proposed our segmentation algorithm to extract different types of trend patterns. Then, we proposed six features that are described in the later section to represent the trend patterns, instead of using vector values to identify the selling and buying signals, or to describe the direction of trend patterns as features. This is because the vector values do not provide clear information for representing trend patterns.

2.3 Artificial Neural Network (ANN)

ANN is a mathematical model that inspired by the neural network of human brain. Human brains consist of few hundred billions of neurons, with each neuron acting as an independent biological information processing unit (Gershenson, 2003). This concept contributes a strong inspiration for building an intelligent neural model which

known as ANN. This model functions by simulating the functioning of brain for problem solving on a computer system (Alvarez, 2006; Jain et al., 1996).

Generally, ANN acts as a black box to classify an output based on patterns recognised from a given input (Dongare et al., 2012). Fig. 1 shows the basic structure of an ANN. It is consists of an input layer, hidden layer and output layer. All the nodes of each layer are connected to each node in the next layer. During the learning and training process, the training inputs are presented at the input layer, with connections in the hidden layer used to acquire classification, while the desired output is presented at output layer. Moreover, ANN learns according to the given input features.
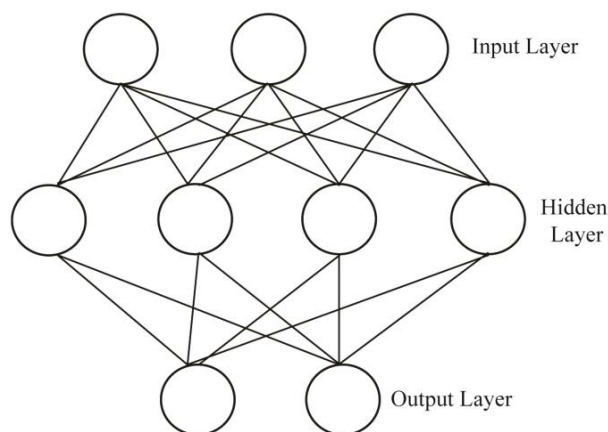
Fig. 1 The Structure of ANN

Many studies have examined that features representation is a critical factor in the performance of ANN because it contains important hidden information in order to train the ANN consistently for predicting the future trends (Charkha, 2008; Desai et al., 2013; Enke et al., 2011; Goswami et al., 2009; Kamijo and Tanigawa, 1990; Tiong et al., 2013). Few experiments were conducted by Lawrence (1997), Yao and Tan (2000), Kordos and Cwiok (2011), which have utilised technical analysis methods to analyse the trend signals and to train the ANN for predicting the direction of trends. To enhance the prediction model, they suggested that trading strategy need to be considered as extra features for prediction. No studies, however, have fully utilised the trend patterns as features to train the ANN for prediction. It is difficult to determine the usefulness of trend patterns from the Forex data.

In this study, we proposed six features that were mentioned as input parameters to train the ANN for classifying the trend patterns. This work, we used WEKA Multi-Layer Perceptron (MLP) classifier with its default settings for training. Furthermore, we also implemented new fusion technique – DTW as opposed to use ANN as a prediction technique that was reviewed for predicting the future trend patterns.

2.4 Dynamic Time Warping (DTW)

DTW is a pattern matching algorithm that measures the similarity between two time series data by warping the distance. It aims to map between two sequences of time series data by warping the axis iteratively until a match is found between themselves (Salvador and Chan, 2007). Then, a distance measurement is needed to determine the similarity between two sequences of time series data. Euclidean measurement is an efficient measurement which calculates the sum of squared distance from each point of one time series to other point of time series data, to find the distance warp path.

Fig. 2 shows an illustration of a warping distance between two time series. Each vertical line connects a point in one time series to its correspondingly similar point to the other time series, while the bold solid line and bold

dotted line represents a distinct time series. To find a warping path from two sets of time series data, the following equation shows the idea:

$$Dist(W) = \sum_{k=1}^{k=K} Dist(w_{ki}, w_{kj}),\qquad(2)$$

where *Dist (W)* is the distance of warp path for *W* and *Dist (w_{ki}, w_{kj})* is the distance between two set of time series data point (w_{ki} from bold solid line and w_{kj} from bold dotted line shown in Fig. 2). The *k* elements represent the warp path between both data point. The calculation of *Dist (w_{ki}, w_{kj})* applies a Euclidean distance measurement to calculate the distance between each vector from w_{ki} and w_{kj}. Each vector value from w_{ki} and w_{kj} has been normalised between -1 and 1.
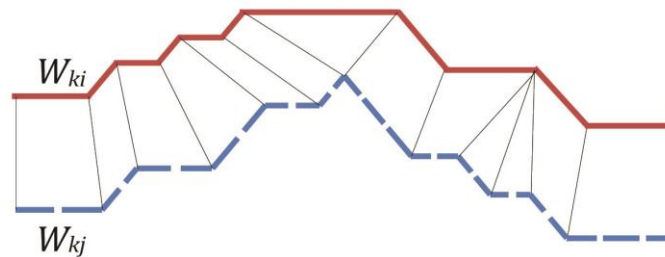

Fig. 2 A warping path between two time series data

Basically, DTW generates a grid table between two sequences of time series data for warping the shortest distance. Fig. 3 shows the example of the distance matrix table. As can be seen, both sequences start on the bottom right of the grid. Each cell can be placed with a distance measure comparing the corresponding elements of them until it reaches the top left. To find the best match between two sequences, Eq. (2) is applied to find a path through the grid which minimises the total distance between them. A warping path result is represented as the black grid in Fig. 3, which determines the similarity of trend patterns between the time series data for prediction.
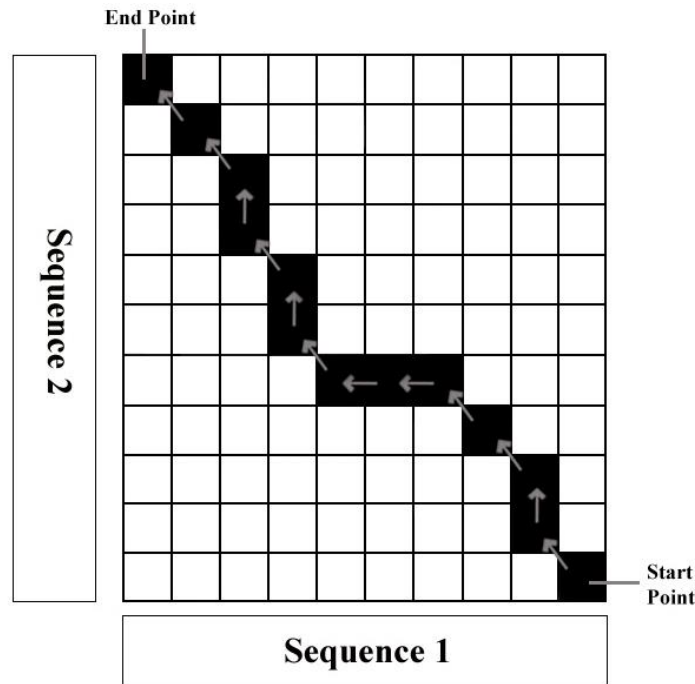

**Fig. 3 Sample of DTW Distance Matrix Table**

DTW is a robust technique that often used to identify two different waveforms that represent the same patterns (Niels, 2004). Additionally, DTW has also been applied in other fields, such as data mining, handwriting recognition and face recognition, where the similarity of shapes is desired. Berndt and Clifford (1994) implemented DTW algorithm to find the patterns in time series data. According to their result, they concluded that this technique can be encouraged for further investigation. Yu *et al.* (2011) also concluded that DTW can strongly be selected in different matching position.
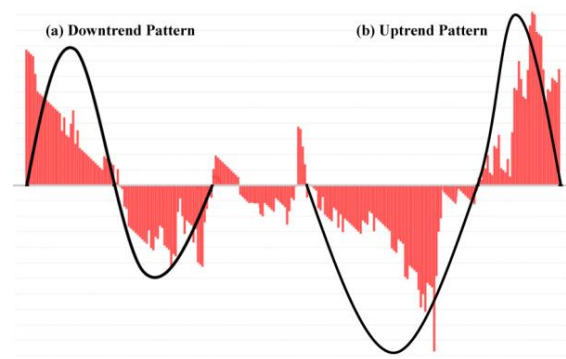
3 Prediction Model Framework

This paper proposes a prediction model using ANN, supported by LRL and DTW. The model is separated into three stages: Data Analysis, ANN Training and DTW Prediction. In the first stage, LRL is applied to analyse and form the general Forex trend patterns. In the next stage, ANN is implemented to learn and train with trend patterns for classification. To predict future trend patterns, DTW is used to determine the degree of similarity based on the trained model at the last stage.

3.1 Data Analysis

This section describes the general process of financial time series data analysis. LRL is implemented to analyse historical Forex data and representing the trends in the form of patterns. In this step, we focused on the closing price from the Forex exchange rates data in a one-year interval. Fig. 4 shows an approach of using Eq. (1) to derive a straight line through the Forex exchange rates of different time stamp, which identifies the overall financial trend from the different time stamps. Fig. 5 illustrates the trend patterns in a histogram chart from Fig. 4 using Eq. (1) to identify the currency exchange rates trend between *EUR* and *USD* on the first week of January 2012. LRL determines the distribution and use for rate projection to form uptrends and downtrends.
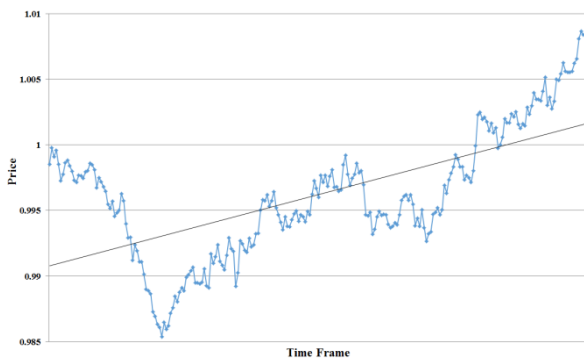
Fig. 4 Forex Trend Analysed by LRL   **Fig. 5 LRL Analysis Result in Histogram Format**

Generally, trend patterns can be analysed as "Uptrend" and "Downtrend" shown in Fig. 5. To identify these main archetypes, we proposed an algorithm – Segmentation Algorithm to segment the trends, fitting them into these archetypes, and storing the values of trend patterns vector for recognising the patterns.

Based on a given sample of time series data shown in Fig. 5, the proposed Segmentation Algorithm (see Fig. 6) begins from the starting point of the time series data to extract the different types of trend patterns from Fig. 5. The algorithm searches for the first turning point to identify as the first half of the pattern, and it continues to find the second turning point. Once the second turning point is detected, the algorithm sets it as the ending point of a trend pattern, and segments the pattern. The algorithm will then begin from the previously ended point as the new starting point, and then continues to extract a new trend pattern from the overall financial trend shown in Fig. 5. The minimum threshold area size for the first and second half of the patterns is set as 5. If the detected patterns do not

meet this requirement, the algorithm will continue to identify the patterns. Those trend patterns are represented in vector format with normalised values.

```
 1:  T ← {X₁, X₂, ..., Xᵢ} //time series data
 2:  P ← { } //storing the trend patterns
 3:  negativeCount = 0 //Initialise as counter
 4:  positiveCount = 0 //Initialise as counter
 5:  for k ← 0 to T do
 6:      if Xₖ < 0 then
 7:          P(k) ← Xₖ
 8:          negativeCount++
 9:          continue
10:      else if Xₖ > 0 then
11:          P(k) ← Xₖ
12:          positiveCount++
13:          continue
14:      else
15:          if negativeCount < minArea then
16:              P(k) ← Xₖ
17:              negativeCount = 0
18:              positiveCount++
19:          else if negativeCount >= minArea then
20:              P(k) ← Xₖ
21:              negativeCount = 0
22:              positiveCount++
23:          else if positiveCount < minArea then
24:              P(k) ← Xₖ
25:              positiveCount = 0
26:              negativeCount++
27:          else if positiveCount >= minArea then
28:              P(k) ← Xₖ
29:              positiveCount = 0
30:              negativeCount++
31:          end if
32:      end if
33: end for
```

Fig. 6 Segmentation Algorithm

Furthermore, we proposed new features to represent trend patterns that were shown in Fig. 7. This is because the proposed features able to represent the different types of trend patterns more efficiently for clustering and ANN training. This consists of a start point, an end point, and an area of trend patterns. K-mean clustering algorithm (Faber, 1994) is applied to cluster all the trend patterns into two main clusters – "Uptrend" and "Downtrend" with the proposed features. The implementation of K-mean algorithm provides clear and well separated information by grouping the extracted trend patterns into two main clusters.
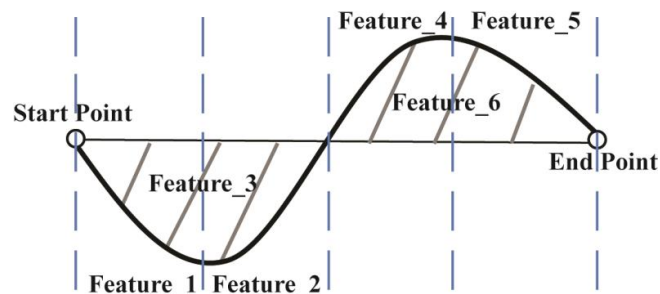

Fig. 7 Selected Features

3.2 ANN Training

After the data analysis stage, we implemented a classifier to train with trend patterns to determine an "Uptrend" or a "Downtrend" by using ANN. In order to calculate the minimum error of training process, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are used in the proposed model.

Table 1 lists the selected features representation of trend patterns that is extracted from "Uptrend" and "Downtrend", as mentioned in Fig. 5. These selected features are used as input features for the training process of ANN. These features are better than using vector format to train the ANN because trend patterns can be represent in different length of vectors, which influence the learning ability of ANN.

Table 1 Detail of Features Selection

Attributes	Description	Value
Feature_1	Distance between starting point to turning point.	1 – Up 2 – Normal 3 – Down
Feature_2	Distance between turning points to exchange point.	1 – Up 2 – Normal 3 – Down
Feature_3	The area between starting point to exchange point.	1 – 10
Feature_4	Distance between exchange point to turning point.	1 – Up 2 – Normal 3 – Down
Feature_5	Distance between turning point to ending point.	1 – Up 2 – Normal 3 – Down
Feature_6	The area between exchange point and ending point.	1 – 10

ANN is implemented to train with the selected features for classifying the unknown trend patterns as either "Uptrend" or "Downtrend" effectively for prediction. Then, DTW is utilised to identify for the pattern that is matched against the referential patterns in the group based on the classification results that belongs to either "Uptrend" or "Downtrend".

3.3 DTW Prediction

After the unknown trend patterns had been classified as one of the trained group, DTW is applied in order to compare the test patterns through brute force against the referential patterns from the database to identify the shortest warping path. DTW is used to identify the most similar patterns of the unknown trend patterns from train patterns. Eq. (2) is utilised to find the shortest path distance of the train pattern.

Once the warping distance of the train pattern is found, the distance matrix table based on DTW is generated. The matrix table shows the matched sequence elements from both unknown and train patterns. The result from the matrix table indicates that in order to predict future trends, DTW based identification of the unknown patterns using the train patterns is carried out, and the remaining sequence elements between the two trend patterns are used for prediction.

Based on the remaining part of sequence elements from the distance matrix table, 2×2 Structure Elements algorithm is applied to identify the diagonal. Fig. 8 shows the rules to identify the diagonal as the last matched

point between the two trend patterns. 1 is considered as starting point, 2 is considered as going to left direction, 3 is consider as up direction and 4 is considered as diagonal.

```
1: M //Two dimensional of Pattern Representative
2: for k ← 0 to M do
3:    for j ← 0 to M do
4:        if k == 0 and j == 0 then
5:            set as "1"
6:        else if k == 0 and j == 1 then
7:            set as "2"
8:        else if k == 1 and j == 0 then
9:            set as "3"
10:       else if k == 1 and j == 1 then
11:           set as "4"
12:       end if
13:   end for
14: end for
```

Fig. 8 2 ×2 Structure Elements Algorithm

Fig. 9 shows the implementation of 2×2 Structure Elements algorithm. The result obtained in distance warping path table that shown in Fig. 9, is traced by 2×2 Structure Elements in order to achieve a pattern representation table. The 2×2 Structure Elements algorithm begins by tracing from the starting point (value "1" as starting point) based on the warping path matrix table (see Fig. 9). Once the last point of the unknown pattern is found (based on the rules of 2×2 Structure Elements algorithm), the remaining sequence elements is used to predict the future of financial trends.
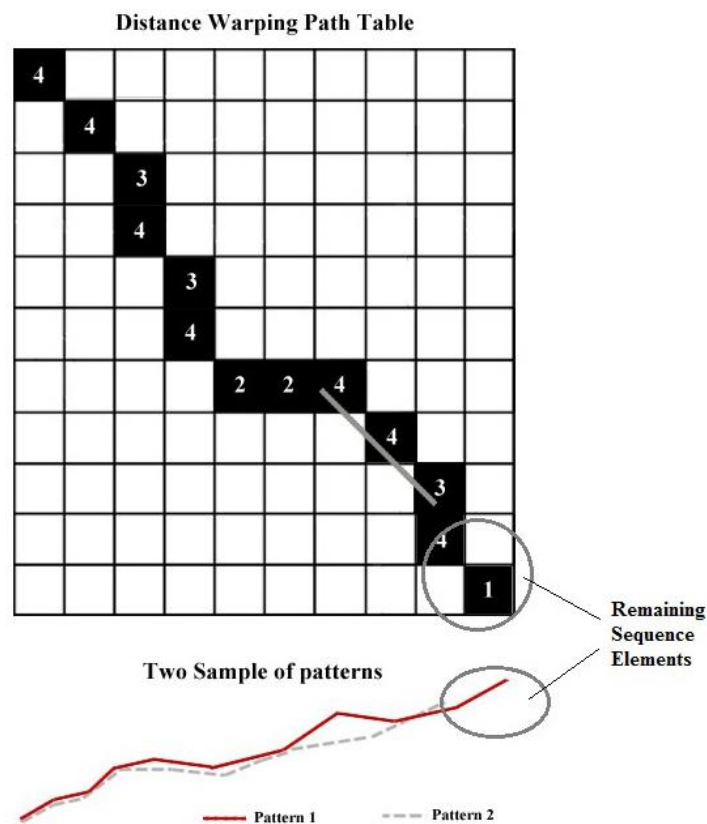

Fig. 9 Implementation of 2 ×2 Structure Elements Algorithm

4 Experimental Result and Discussion

The Forex data of *AUD – USD* and *EUR – USD* employed in the experiments consists of half-hour closing prices. The dataset encompasses the time range from 2nd January 2011 until 31st December 2011. The data is collected from Free Forex Historical Data website (HistData.com, 2013).

4.1 Data Analysis Result

Based on the historical Forex data, we have successfully analysed the overall trend shown in Fig. 10 using LRL. Once the overall trends have been analysed, the patterns can be identified as either "Uptrend" or "Downtrend", as mentioned in Fig. 5.
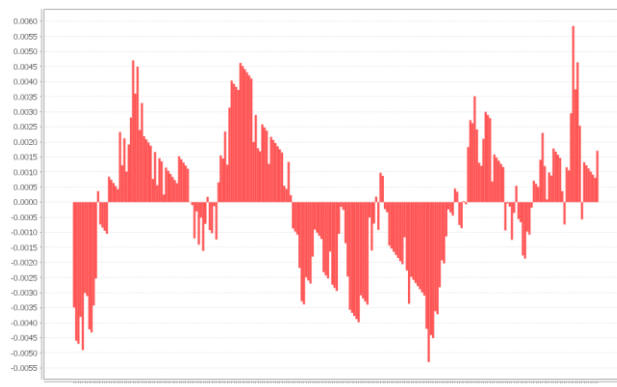

Fig. 10 Forex Trend in Histogram Format

Figs. Fig. 11 and Fig. 12 show the samples result of segmentation for trend patterns. The Segmentation Algorithm proposed in this paper has identified different types of trend patterns. According to the result, we demonstrated that LRL has successfully distinguished the patterns of uptrend and downtrend based on the vector values of time series data and the rules from the proposed Segmentation Algorithm.
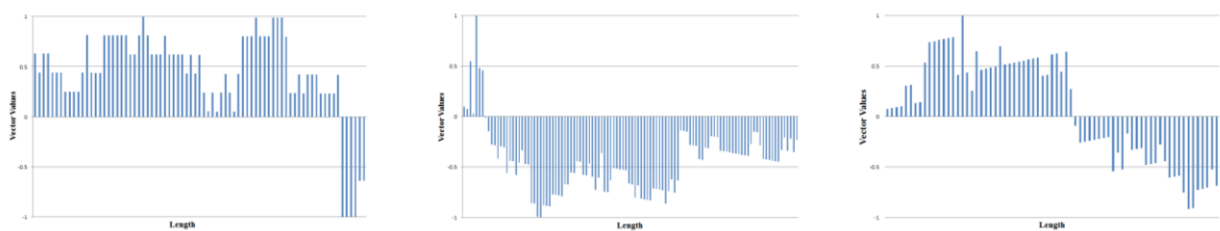

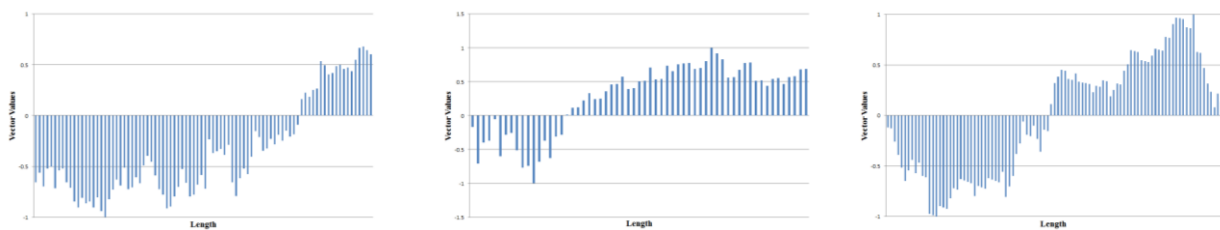**Fig. 11 Samples of different Downtrend Patterns**

Fig. 12 Samples of different Uptrend Patterns

4.2 ANN Training Process Result

ANN – Multilayer Perceptron is implemented from the WEKA (Waikato, n.d.) open source library to train with trend patterns to classify trend patterns as "Uptrend" or "Downtrend". In the experiments, the dataset has been separated into two types – training set and testing set, in a 7:3 ratio. A new set of Forex data are tested with the proposed model for validation. The result concludes that ANN could be used to classify trend patterns.

According to

Table 2, the result shows that ANN has classified trend patterns properly according to the Forex data. The result of the model for *AUD – USD* shows 0.0051 for MAE and RMSE, followed by *EUR – USD* shows 0.0044 for MAE and RMSE.

Table 2 Training Result of Classification for "Uptrend" and "Downtrend" using ANN

Data	Number of Dataset		Result	
	Training	*Testing*	*MAE*	*RMSE*
AUD – USD 2011	118	51	0.0051	0.0051
EUR – USD 2011	139	60	0.0044	0.0044

35 datasets, randomly selected from each of the historical Forex data: *AUD – USD* and *EUR – USD* that encompassed from 2012 are used to evaluate the proposed selection features are useful in training process of ANN. The evaluation results shown in Table 3 proved that ANN had classified trend patterns with 100% correctness.

Table 3 Evaluation Result for Classification for "Uptrend" and "Downtrend" using ANN

Sampling Data	Number of Evaluation Dataset	Classification Result
AUD – USD 2012	35	100%
EUR – USD 2012	35	100%

Table 4 showed the comparison result of classification using ANN with different type selection features. The result had shown that the proposed method could classify the patterns better than other approaches. In conclusion, the proposed selection features shown in Table 1 as input parameters had successfully represented trend patterns for ANN to train. Additionally, the result shown in Table 4 also proved that feature selection is one of the most important parameters before the training process of ANN on trend patterns. This is because without the suitable features to represent trend patterns, ANN would not be able to train efficiently for classifying the patterns.

Table 4 Comparison of Different Selection Features in Classification

| Method	Author	Classification Result
ANN with Candlestick Pattern	Kamijo and Tanigawa (Kamijo & Tanigawa 1990)	94%
ANN with ARIMA	Kamruzzaman and Sarker (Kamruzzaman & Sarker 2004)	99%
ANN with ARIMA	Yao and Tan (Yao & Tan 2000)	85%
ANN with LRL and Proposed Features	Tiong *et al.*	100%

4.3 Proposed Prediction Model Result

After successfully classified 40 and 35 datasets, which are randomly selected from each of the historical Forex data: *AUD – USD* and *EUR – USD* in the previous stage, this experiment considers those patterns as partial unknown patterns for the purpose of evaluating DTW in the application of predicting future trends.

In this experiment, Fig. 13 is assigned as a fully uptrend pattern and Fig. 14 is assigned as a partial uptrend pattern. DTW was implemented to identify the warping path for the fully known patterns to predict the partial unknown patterns. DTW generated a cost matrix table which is shown in Table 5.
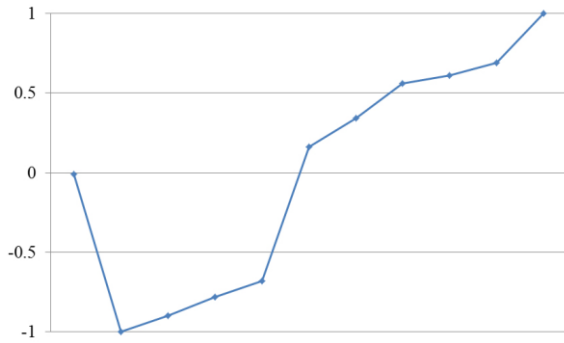

Fig. 13 Full Uptrend Pattern
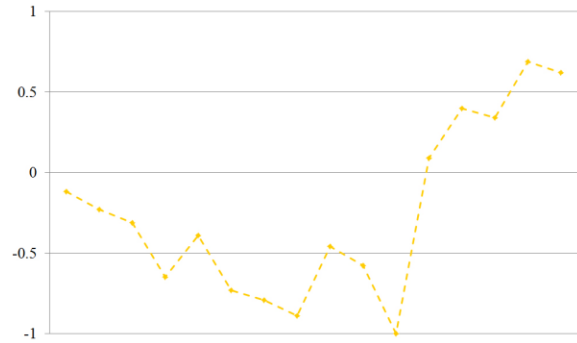
Fig. 14 Partial Uptrend Pattern

According to the proposed model, DTW is implemented to determine the minimum distance path along with two sequences of trend patterns that shown in Figs. Fig. 13 and Fig. 14. Table 5 shows the resultant distance matrix of the warp path between partial (Fig. 14) and full (Fig. 13) trend patterns from the experiment. Based on the result, DTW had successfully calculated and traced the minimum value from the starting point (right bottom in Table 5) to the ending point (top left in Table 5) by using Eq. (2).

2×2 structure elements are applied, as mentioned in Fig. 8 to find the remaining parts for predicting the future trends. Figs. Fig. 15 to Fig. 17 demonstrate the performance of our proposed prediction model in predicting the future direction of financial trends with satisfactory results. In these experiments, DTW and 2×2 structure elements algorithms have successfully found the matched vector elements ("Matched" points shown in Figs. Fig. 15 to Fig. 17) from the patterns. Then, the proposed 2×2 Structure Elements algorithm has effectively identified the remaining parts for predicting the future of financial trends.

**Table 5 DTW Distance Matrix Table for Minimum Warp Path**

0.031	**0.044**	**0.102**	**0.746**	**0.892**	1.625	2.401	3.222	3.545	3.912	4.865	4.940	5.359	5.747	6.480	7.183
1.015	0.971	0.94	0.411	1.218	**1.113**	1.29	1.422	2.053	2.639	2.639	3.668	5.040	6.382	7.434	8.136
1.922	1.834	1.759	0.643	1.142	1.257	**1.213**	1.268	1.822	2.332	2.409	3.36	4.655	5.920	7.530	9.012
2.752	2.62	2.501	0.799	1.297	1.208	1.231	**1.234**	1.711	2.144	2.298	3.172	4.390	5.578	7.111	8.613
3.506	3.329	3.166	0.877	1.375	1.218	1.263	1.330	**1.634**	**1.990**	**2.221**	3.018	4.160	5.271	6.726	8.151
3.643	3.510	3.391	1.689	1.191	2.092	2.163	2.252	1.82	2.169	3.111	**2.314**	2.565	2.785	3.350	3.884
3.808	3.720	3.645	2.529	1.533	2.119	3.064	3.180	2.339	2.383	3.318	2.435	**2.537**	2.730	3.267	3.774
4.371	4.326	4.296	3.767	2.272	2.859	3.489	4.478	3.255	3.299	3.93	2.954	2.609	**2.742**	2.869	2.978
4.961	4.961	4.975	5.032	3.039	3.626	4.257	4.931	4.199	4.243	4.873	3.500	2.812	2.842	**2.854**	2.935
5.579	5.623	5.668	6.268	3.834	4.421	5.051	5.726	5.170	5.214	5.845	4.074	3.041	3.072	2.926	**2.907**
6.595	6.639	6.728	7.358	5.026	5.613	6.244	6.919	6.539	6.583	7.214	5.046	3.669	3.699	3.240	**3.252**
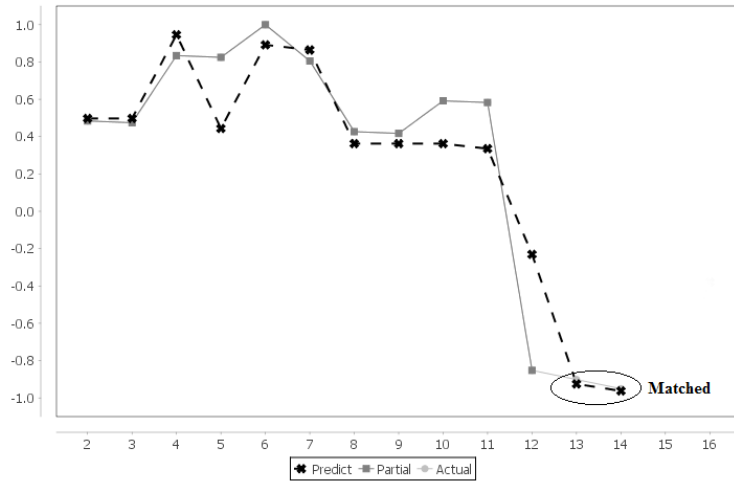
Fig. 15 Prediction Result of *AUD – USD* Exchange Rates in the First Week of February, 2012
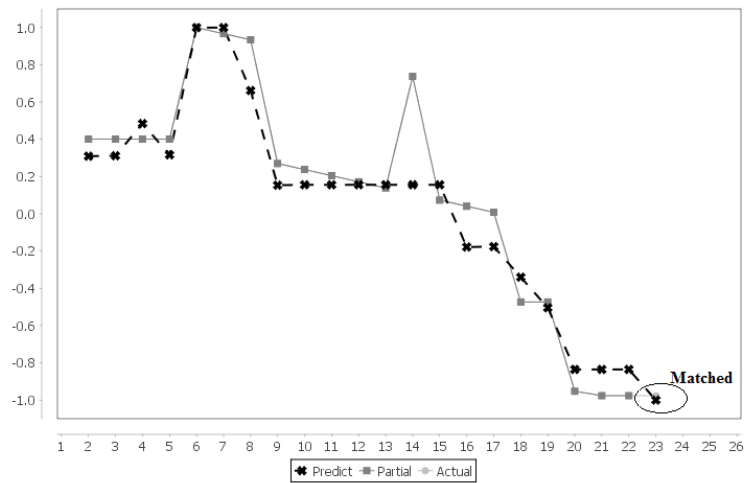
Fig. 16 Prediction Result 2 of *AUD – USD* Exchange Rates in the Second Week of May, 2012
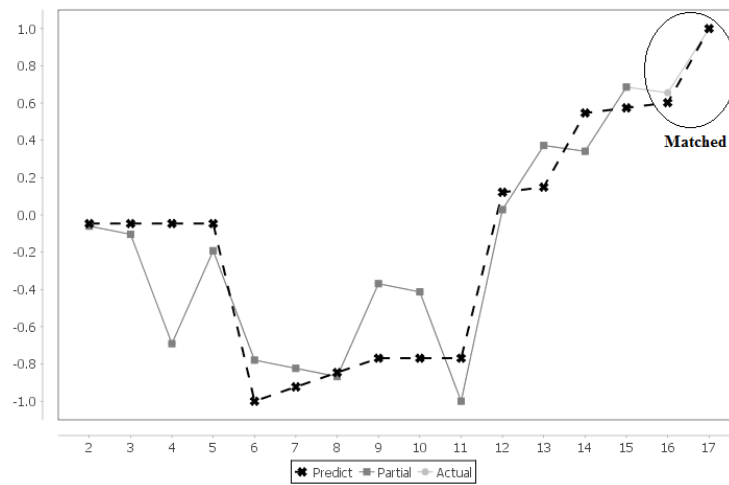
Fig. 17 Prediction Result 3 of *EUR – USD* Exchange Rates in the First Week of June, 2012

According to the prediction results that were shown in Figs. Fig. 15 – Fig. 17, DTW is able to predict the trend patterns which able to identify the direction of future trends. We concluded that DTW has successfully identified the patterns and has proved to be able to predict future trends. The use of DTW is explored to predict Forex movement for $AUD - USD$ and $EUR - USD$. For comparison purposes, the values predicted by the DTW models are compared with two of the most popular strategies for supervised machine learning and classification, ANN and SVM, both trained using the WEKA open source library.

As shown in Table 6, the comparison indicates that our DTW predictor obviously outperforms the ANN and SVM models in prediction accuracies on the datasets with accuracy rates between $70.5 - 72.5\%$. This means that the proposed approach is capable of predicting the future trends better than other approaches like ANN and SVM. The results also show that our proposed features have successfully represented the patterned information and eliminated the effect of noise in the data.

Table 6 Prediction Result

Approach	Tested Forex Data	Time Stamp	Number of Tested Dataset	Accuracy (%)
ANN and DTW	$AUD - USD$ in 2012	Weekly	40	70.0
	$EUR - USD$ in 2012			72.5
	$AUD - USD$ in 2013	Weekly	35	70.5
	$EUR - USD$ in 2013			71.0
ANN	$AUD - USD$ in 2012	Weekly	40	47.5
	$EUR - USD$ in 2012			50.0
	$AUD - USD$ in 2013	Weekly	35	51.4
	$EUR - USD$ in 2013			54.2
SVM	$AUD - USD$ in 2012	Weekly	40	50.0
	$EUR - USD$ in 2012			54.2
	$AUD - USD$ in 2013	Weekly	35	45.7
	$EUR - USD$ in 2013			47.5

4.4 Comparison with Other Approaches

Table **7** presents a comparison of methods used by the proposed approach with three other state-of-the-art Forex prediction approaches. Both Caginalp and Laurent (1998) and Lee and Jo (1999) used candlestick patterns as input features to analyse trend patterns, and achieved a 67% and 72% accuracy rate respectively. Zhai *et al.* (2007) used seven technical analysis methods to represent trend patterns, and the proposed model achieved a 70% accuracy rate in prediction. In Tiong *et al.* (2013), candlestick patterns were used as input parameters to train the prediction model. As seen in the table, although the method of the research is simple, it did not provide a satisfying result for practical use. Since the testing conditions were different it is not possible to draw absolute conclusions from the comparisons.

In this paper, we propose six new features based on LRL, followed by K-mean and the proposed segmentation algorithms to cluster the trend patterns. Table 3 shows that our proposed features can be classified using ANN. Our proposed model provided a 71% accuracy rate in prediction using DTW. This experiment proves that the proposed features can represent the trend patterns in a more efficient way.

However, Lee and Jo (1999) proposed ES that was provided better prediction accuracy compared to our proposed prediction approach. Yet, the rule-based ES lacks automated learning that has to adjust itself according to new conditions and parameter values in different situations. It is evident that our proposed feature outperforms the various features representation for trend identification in all tests. Furthermore, our proposed approach has proven that a new fusion approach – ANN and DTW can achieve a better prediction result.

Table 7 Comparison between Proposed Approach and Previous Studies

Authors	Features	Proposed Approach	Accuracy (%)
Caginalp and Laurent (1998)	Three-day candlestick patterns	Statistic method	67
Lee and Jo (1999)	Candlestick patterns and financial strategy rules	ES	72
Zhai *et al.* (2007)	Seven indicators and news	SVM	70
Tiong *et al.* (2013)	Candlestick patterns (Real Body, Upper Shadow, Lower Shadow, Candlestick Position)	ANN	59
Proposed approach	Proposed six features	ANN and DTW	71

5 Conclusion and Future Work

The technical analysis method – LRL proposed in this paper has successfully analysed the trend patterns from Forex data. Additionally, the proposed segmentation algorithm has effectively identified different trend patterns.

The selected features mentioned in Table 1 are fully utilised to represent the patterns for ANN to train, and the result in Table 3 showed that ANN had effectively classified the partial patterns of "Uptrend" and "Downtrend" with the features. During the training process of ANN, the selected features are very important. This is because ANN is sensitive to the input parameters in the training process.

After ANN identifies trend patterns, DTW is implemented for predicting future trends. In the prediction experiment, the results shown in Table 6 proved that DTW had successfully predicted the nature of trends. In summary, the proposed method in this paper had successfully predicted the future trends with 70% – 71% accuracy rates.

For future work, we will use other historical data such as stock price as the primary information for prediction. Further exploration in the training process stage may use the supervised learning algorithms, such as HMM and SVM with the selected features to identify trend patterns and explore the differential using ANN in the training process. The performance of DTW in the prediction will be considered as an optimal process instead of using brute force computation to identify patterns.

References

Achelis, S.B. (1995) *Technical Analysis from A to Z*, Chicago.

Ahangar, R.G., Yahyazadehfar, M. and Pournaghshband, H. (2010) 'The Comparison of Methods Artificial Neural Network with Linear Regression Using Specific Variables for Prediction Stock Price in Tehran Stock Exchange', *International Journal of Computer Science and Information Security*, Vol. 7 No.2, pp.38–46.

Alvarez, E.G. (2006) *Artificial Neural Networks*. http://edugi.uni-muenster.de/eduGI.LA2/downloads/02/ArtificialNeuralNetworks240506.pdf. (Accessed 26 July 2012)

Berndt, D.J. and Clifford, J. (1994) *Using Dynamic Time Warping to Find Patterns in Time Series*,

Caginalp, G. and Laurent, H. (1998) 'The Predictive Power of Price Patterns', *Applied Mathematical Finance*, Vol. 5, pp.181–205.

Charkha, P.R. (2008) 'Stock Price Prediction and Trend Prediction Using Neural Networks' in *First International Conference on Emerging Trends in Engineering and Technology*, Maharashtra, India, pp. 592–594.

Coelho, M. (2011) *Patterns in Financial Markets: Dynamic Time Warping*. Nova School of Business and Economics, Lisboa, Portugal.

Desai, J., Trivedi, A. and Joshi, N.A. (2013) *Forecasting of Stock Market Indices Using Artificial Neural Network*, Ahmedabad.

Dongare, A.D., Kharde, R.R. and Kachare, A.D. (2012) 'Introduction to Artificial Neural Network', *International Journal of Engineering and Innovative Technology*, Vol. 2 No. 1, pp.189–194.

Dzikevičius, A. and Šaranda, S. (2010) 'EMA Versus SMA Usage to Forecast Stock Markets: The Case of S&P 500 and OMX Baltic Benchmark', *Business: Theory and Practice*, Vol. 11 No. 3, pp.248–255.

Edwards, R. and Magee, J. (2007) *Technical Analysis of Stock Trends,* 9th ed., W. H. C. Bassetti.

Enke, D., Grauer, M. and Mehdiyev, N. (2011) 'Stock Market Prediction with Multiple Regression, Fuzzy Type-2 Clustering and Neural Networks', *Procedia Computer Science*, Vol. 6, pp.201–206.

Faber, V. (1994) 'Clustering and the Continuous k -Means Algorithm', *Los Alamos Science*, Vol. 22, pp.138–144.

Fernández-Blanco, P., Bodas-Sagi, D., Soltero, F. and Hidalgo, J.I. (2008) 'Technical Market Indicators Optimization using Evolutionary Algorithms' in *The 10th Annual Conference on Genetic and Evolutionary Computation*, Atlanta, Georgia, pp. 1851–1858.

Ge, X. (1998) *Pattern Matching in Financial Time Series Data*.

Gershenson, C. (2003) *Artificial Neural Networks for Beginners*.

Goswami, M.M., Bhensdadia, C.K. and Ganatra, A.P. (2009) 'Candlestick Analysis Based Short Term Prediction of Stock Price Fluctuation using SOM-CBR' in *IEEE International Advance Computing Conference*, Patiala, India, pp. 1448 – 1452.

Grunder, B., Macdonald, M., Sutter, C. and Zingg, P. (2012) *Technical Analysis - Explained*.

Hassan, M.R. and Nath, B. (2005) 'Stock Market Forecasting Using Hidden Markov Model: A New Approach', in *5th International Conference on Intelligent Systems Design and Applications*, Wroclaw, Poland, pp.192–196.

Hassan, M.R., Nath, B. and Kirley, M. (2007) 'A Fusion Model of HMM, ANN and GA for Stock Market Forecasting', *Journal of Expert Systems with Applications*, Vol. 33, No. 1, pp.171–180.

HistData.com. (2013) [online] http://www.histdata.com/ (Accessed October 10, 2012).

Jain, A.K., Mao, J. and Mohiuddin, K.M. (1996) 'Artificial Neural Networks: A Tutorial', *IEEE Computational Science & Engineering*, Vol. 29 No. 3, pp.31–44.

Janssen, C., Langager, C. and Murphy, C. (2012) *Technical Analysis: What Is A Chart.* http://stockcharts.com/school/doku.php?id=chart_school:chart_analysis:what_are_charts (Accessed November 9, 2012).

Kamijo, K. and Tanigawa, T. (1990) 'Stock Price Pattern Recognition - A Recurrent Neural Network Approach', *International Joint Conference on Neural Networks*, Vol. 1 No. 1, pp.215–221.

Kamruzzaman, J. and Sarker, R.A. (2004) 'ANN-Based Forecasting of Foreign Currency Exchange Rates', *Journal of Neural Information Processing*, Vol. 3 No. 2, pp.49–58.

Kim, K. (2003) 'Financial Time Series Forecasting using Support Vector Machines. *Neurocomputing*, Vol. 55 No.2, pp.307–319.

Klassen, M. (2005) 'Investigation of Some Technical Indexes in Stock Forecasting Using Neural Networks' in *The Third World Enformatika Conference*, Istanbul, Turkey, pp. 75–79.

Koedijk, K.G., Lothian, J.R. and Van Dijk, M.A. (2006) 'Foreign Exchange Markets: Overview of The Special Issue', *Journal of International Money and Finance*, Vol. 25 No. 1, pp.1–6.

Kordos, M. and Cwiok, A. (2011) 'A New Approach to Neural Network Based Stock Trading Strategy' in *International Conference on Intelligent Data Engineering and Automated Learning*, Hefei, China, pp. 429–436.

Kuepper, J. (2012) *Introduction To Swing Charting.* http://www.investopedia.com/articles/technical/04/080404.asp#axzz2AYr0U5i7 (Accessed October 2, 2012).

Lawrence, R. (1997) *Using Neural Networks to Forecast Stock Market Prices.* http://www.cs.uiowa.edu/~rlawrenc/research/Papers/nn.pdf (Accessed October 2, 2012).

Lebaron, B. (2006) 'Time Scales, Agents, and Empirical Finance', *Medium Econometrishce*, Vol. 3, pp.20–25.

Lee, K.H. and Jo, G.S. (1999) 'Expert System for Predicting Stock Market Timing Using A Candlestick Chart', *Journal of Expert Systems with Applications*, Vol. 16 No. 4, pp.357–364.

Maknickienė, N. and Maknickas, A., 2012. 'Application of Neural Network for Forecasting of Exchange Rates and Forex Trading' in *The 7th International Scientific Conference Business and Management*. Vilnius Gediminas Technical University Publishing House Technika, Vilnius, Lithuania, pp. 122–127.

Naeini, M.P., Taremia, H. and Baradaran, H.H. (2010) 'Stock Market Value Prediction Using Neural Networks' in *International Conference on Computer Information Systems and Industrial Management Applications*, Krakow, Poland, pp. 132–136.

Niels, R. (2004) *Dynamic Time Warping: An Intuitive Way of Handwriting Recognition*. Radboud University Nijmegen.

Nison, S. (1991) 'Constructing the Candlesticks' in *Japanese Candlestick*, New York: New York Institute of Finance, pp. 21–26.

Nobakht, B., Joseph, C.E. and Loni, B. (2010) *Stock Market Analysis and Prediction using Hidden Markov Models*,

Olaniyi, S.A.S., Adewole, K.S. and Jimoh, R.G. (2011) 'Stock Trend Prediction Using Regression Analysis – A Data Mining Approach', *ARPN Journal of Systems and Software*, Vol. 1 No. 4, pp.154–157.

Person, J.L. (2002) *Swing Trading Using Candlestick charting with Pivot Point Analysis*.

PFGBEST, *Foreign Exchange Basics*,

http://www.pfgbest.com/services/forex/foreignexchangebasics.pdf (Accessed 4 November 2013).

R.King, M., Osler, C. and Rime, D. (2011) *Foreign Exchange Market Structure, Players and Evolution*,

Rao, S. & Hong, J. (2010) *Analysis of Hidden Markov Models and Support Vector Machines in Financial Applications*, University of California, UCB/EECS-2010-63.

Rinehart, M. (2003) *Overview of Regression Trend Channel (RTC)*.

Rockefeller, B., 2011. Drawing 'Trendlines', in *Technical Analysis For Dummies*. Indianapolis: Wiley Publishin, Inc., pp. 169–182.

Salvador, S. and Chan, P. (2007) 'FastDTW: Toward Accurate Dynamic Time Warping in Linear Time and Space', *Journal of Intelligent Data Analysis*, Vol. 11 No. 5, pp.561–580.

Sewell, M. (2007) *Technical Analysis*,
http://www.technicalanalysis.org.uk/ (Accessed 22 January 2013).

Shmatov, A. (2011) *Automated Recognition of Candlestick Patterns.*
http://stock-market-forecast.blogspot.com/2012/02/automated-recognition-of-candlestick.html (Accessed 8 September 2012).

Tanaka-Yamawaki, M. and Tokuoka, S. (2007) 'Adaptive Use of Technical Indicators for the Prediction of Intra-Day Stock Prices', *Physica A: Statistical Mechanics and its Applications*, Vol. 383 No. 1, pp.125–133.

Tiong, L.C.O., Ngo, D.C.L. and Lee, Y. (2013) 'Stock Price Prediction Model using Candlestick Pattern Feature', *International Journal Of Interactive Digital Media*, Vol. 1 No. 3, pp.58–64.

Virginia, 2009. *Foreign Exchange and Currency*.

Machine Learning Group at the Univeristy of Waikato, *WEKA Library*,
http://www.cs.waikato.ac.nz/ml/index.html (Accessed September 11, 2013).

Wong, P. and Yeung, C. (2008) *A Novel Combined Approach for Time Series Data Mining: Stock Trend Prediction Based on Matched Subsequences from Dynamic Time Warping in Wavelet Decomposed Time Series*, Hong Kong, China.

Yao, J. & Tan, C.L. (2000) 'A Case Study on Using Neural Networks to Perform Technical Forecasting of Forex', *Journal of Neurocomputing*, Vol. 34 No. 4, pp.79–98.

Yu, D., Yu, X., Hu, Q., Liu, J. and Wu, A. (2011) 'Dynamic Time Warping Constraint Learning for Large Margin Nearest Neighbor Classification', *Journal of Information Sciences*, Vol. 181 No. 13, pp.2787–2796.

Zhai, Y., Hsu, A. and Halgamuge, S.K. (2007) 'Combing News and Technical Indicators in Daily Stock Price Trends Prediction', *Advances in Neural Networks*, Vol. 4493, pp.1087–1096.

Zhang, G.P. (2001) 'An Investigation of Neural Networks for Linear Time-Series Forecasting', *Journal of Computers and Operations Research*, Vol. 28 No. 12, pp.1183–1202.