A New Application of Hidden Markov Model in Exchange Rate Forecasting

Azunda Zahari¹, Jafreezal Jaafar²

¹Computer & Information Sciences Universiti Teknologi PETRONAS, Malaysia
²Computer & Information Sciences Universiti Teknologi PETRONAS, Malaysia
¹azundaazlina@yahoo.com; ²jafreezal@utp.edu.my

Abstract—This paper presents a new application of Hidden Markov Model (HMM) as a forecasting tool for the prediction of the currency exchange rate between the US dollar and the euro. The results obtained show that the difference between price gaps which consists open, high, and low price can be selected to produce the best model parameter of Hidden Markov Model. Three model parameters based on Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Log likelihood value has been calculated. The empirical results show that the projection from the daily opening price to the highest and lowest price of the day produce lowest AIC, BIC and Log likelihood values respectively. The price pattern based on the daily closing value has been ignored because of poor model performance compared to the other two price gaps. As it turns out, the proposed application outperforms the model which uses the closing price as an input variable in terms of model parameter based on AIC, BIC, and Log likelihood values.

Keywords—Hidden Markov Model, Log likelihood, AIC, BIC, time series forecasting, exchange rates

I. INTRODUCTION

There are a too many machine learning techniques focusing on modelling the financial time series, intending to improve forecasting accuracy. The importance of financial markets for the development of a new model reflected on the increase number of time series studies began in (Johnson, 1969). According to (Johnson, 1969), foreign exchange forces by demand and supply, the freedom of rates to move in response to market forces does not imply that they will in fact move significantly or erratically; they will do so only if the underlying forces governing demand and supply are themselves erratic.

Hidden Markov Model has been studied in many areas including earthquake hazard(Can, Ergun, & Gokceoglu, 2014), marine environment monitoring (Kevin, Emilie, Alain, & Denis, 2015), corporate default counts(Li & Cheng, 2015), stock market forecasting (Arumugam & R, 2013) including exchange rate prediction ((Lee, 2013) and (Komariah & Sin, 2015)). Nevertheless their drawbacks and conflicting experimental results of their forecasting ability are generating uncertainty about their application among researchers.
The contribution of this paper is to introduce in exchange rate market a new application of Hidden Markov Model, which try to overcome some of previous research limitations. In particular our proposed application reduces the numbers of HMM parameters that the traders need to practice while on the other hand it increases the pips accumulation or profit gains in each trading period. The proposed technique is easy to implement in comparison to the application of trend forecasting methods researched by ((Park, Lee, & Lee, 2011) and (Nguyen & Nguyen, 2015))

Hidden Markov Model that has been already proposed in the literature eliminates the risk of getting in both market directions, whether in bull or bear market. The main reason behind our decision to use EUR versus USD is that it is the most tradable pairs by traders. For that reason, it is a tradable quantity which makes our trading simulation more practical. The study examines the exchange rates from their first trading day since the end of January 2012.

II. LITERATURE REVIEW

Researching easy yet effective prediction techniques is a very difficult and complex task for foreign exchange rate practitioners. The conventional statistical methods seem to fail to capture discontinuities, the nonlinearities and the high complexity of datasets such as financial time series. Soft computing techniques like Hidden Markov Model (HMM) provide enough learning capacity and are more likely to capture the complex non-linear models which are dominant in the financial markets but their initial parameter is difficult to set. ((Hassan, 2009) and (Badge & Srivastava, 2010)). Despite of that, previous techniques let the trader have to identify the trend ((Bakhch, Tsang, & Ng, 2015) and make decision whether to buy or sell position (Oladimeji, 2016).

In contrast to previous studies, which focusing more on predicting the closing price, this study apply opening price to predict the price projection which is able to find the trader’s profitable entry point. The aim of this paper is to introduce a new application of Hidden Markov Model which is able to overcome the difficulties in setting the parameters of the model. For this purpose among the various Hidden Markov Model applications, we use the 3 input features consist of open, high, and low price value. These input values has been applied in (Hassan, 2009) which has proven experimentally to outperform the traditional Hidden Markov Model.

The mentioned three input variables has been applied in several studies including (Nath, 2005), (Kentucky & Emam, 2008), (Tiong, Ngo, Lee, & Bar, 2013). This methodology was applied for the prediction of the price of financial time-series. Despite the high prediction accuracy of the derived model, this methodology does not provide any method for analyzing the price projection from open price value. Moreover, the Hidden Markov Model to the task of forecasting financial series with ambiguous empirical evidence. (Ahani, Abass, & Okunoye, 2010) use the Hidden Markov Model to develop a simulation model of the Nigerian Stock Exchange using real life data collected from the daily stock market prices of the Nigerian Stock Exchange between the years 2005 and 2008 which assumed that the hidden states are bull, bear and Even. These hidden states, along with the observable sequences of large rise, small rise, no change, large drop and small drop, were used to develop the Hidden Markov Model. The visualization reveals that the stock exchange will only be steady periodically irrespective of initial conditions. The non-stationary and non-linearity of time series lead to an extension of the Hidden Markov Model to include a novel exponentially weighted Expectation-Maximization (EM) algorithm in (Zhang & Science, 2004) to handle these two challenges. This extension allows the HMM algorithm are proved using techniques from the EM Theorem. Experimental results show that their proposed model consistently beat the S&P 500 Index over five 400-day testing periods from 1994 to 2002, including both bull and bear markets.

Referring to ((Hu, 1987) and (Christensen & Pedersen, 2011)), this study present criteria used to estimate the application of HMM parameters based on likelihood, AIC and BIC value. The Bayesian information criterion (BIC) and Akaike Information Criterion (AIC) values were used for model selection.
The lowest BIC and AIC value is preferred. On the other hand the highest Log likelihood value is chosen for the best model.

Modelling high accuracy techniques for predicting time series is a very crucial problem for scientists and decision makers. The traditional statistical methods seem to fail to capture the discontinuities, the nonlineairties and the high complexity of datasets such as financial time series. Soft computing techniques like Hidden Markov Model provide enough learning capacity and are more likely to capture the complex non-linear models which are dominant in the financial markets but their initial parameter is difficult to set ((Hassan, 2009), and (Badge & Srivastava, 2010)). (Bakhach et al., 2015) apply the Directional Change (DC) approach aims to capture directions price movements – whether to take a long or short position.

The main objective of this paper is to introduce a new application of Hidden Markov Model, which try to overcome some of these limitations. The proposed technique eliminates the number of parameters that the practitioner needs to experiment while on the other hand it increases the forecasting ability of the model.

Hidden Markov Model is employed in (Li & Cheng, 2015) to fit global, U.S. and European annual corporate default counts. All parameters are calibrated by using the Expectation-Maximization algorithm while the standard errors of the estimated parameters are conducted by Monte Carlo method. Little forecasting research has been done under the Directional Change (DC) framework motivates (Bakhach et al., 2015) to formulate a forecasting problem to answer the question of whether the current trend (up or down) will continue for a particular percentage (which is decided by the investor) before the trend ends. The trends movement also considered in (Komariah & Sin, 2015) for predicting the trends of USD Dollar/IDR Rupiah movement using Hidden Markov Model using the hidden variables correspond to people’s sentiment manifested by tweets every day. The study of financial time series trends continued by (Sandoval, 2015), which presents Hierarchical Hidden Markov Model used to capture the USD/COP market sentiment dynamics choosing from uptrend or downtrend latent regimes based on observed feature vector realizations calculated from transaction prices and wavelet-transformed order book volume dynamics. The new classification method for identifying up, down, and sideways trends in Forex market foreign exchange rates proposed by (Talebi, Hoang, & Gavrilo, 2014).

The performance of state-of-the-art machine learning techniques are investigated by (Theofilatos & Karathanasopoulos, 2012) in trading with the EUR/USD exchange rate using five supervised learning classification techniques (K-Nearest Neighbors algorithm, Naive Bayesian Classifier, Artificial Neural Networks, Support Vector Machines and Random Forests). These techniques were applied in one day movement prediction of the EUR/USD exchange rate using autoregressive terms as inputs. The performance of the selected techniques was benchmarked with Naive Strategy and moving average convergence/divergence model. The result shows that trading strategies produced by the machine learning techniques using Support Vector Machines and Random Forests outperformed all other strategies in terms of annualized return and sharp ratio. To the best of their knowledge, they came out with the first application of Random Forests in the problem of trading with the EUR/USD exchange rate providing extremely satisfactory results.

The USDCHF trading strategy devise by (Lee, 2013) using the dynamics of gold as a filter. The strategy is about modeling both USDCHF and gold using a coupled hidden Markov model (CHMM). The observations were done using the indicators, RSI and CCI, which will be used in producing the trading signals. The study intend to take advantage of intermarket analysis and the Markov property implicit in the hope that these most probable values will produce profitable trading results. The performance of ARIMA and artificial neural networks model was examined by (Adebiyi, Adewumi, & Ayo, 2014) for forecasting stock data obtained from New York Stock Exchange. The study results show the superiority of neural networks model over ARIMA model. (Xhaja, Ktona, & Brahushi, 2014) deals with the application of Support Vector Machine (SVM) and Artificial Neural Network (ANN) in one step forecasting of
EUR/ALL exchange rate based on the patterns that previously occurred. EUR/ALL exchange rate on weekdays for a period of four years is used under study. Several experiments are conducted applying different parameters to both techniques. Root Mean Square Error (RMSE) is used as a measure to evaluate and compare the predictive power of SVM and ANN.

The EURJPY exchange direction is examined by (Haeri, Hafezi, & Rezaie, 2015) which attempt to forecast the price direction whether increase or decrease using five major indicators including exponential moving average (EMA), stochastic oscillator (KD), moving average convergence divergence (MACD), relative strength index (RSI) and Williams %R (WMS %R). Proposed approach is using hidden Markov models and CART classification algorithms is used for forecasting direction (increase or decrease) of Euro-Yen exchange rates in a day. Then the method is compared with CART and neural network. Comparison shows that the forecasting with proposed method has higher accuracy.

Generalized Autoregressive Conditional Heteroscedastic approach is considered in (Milano, 2013) to model the Eurodollar spot and futures exchange rate’s volatilities using daily and weekly observations, which apply different currency futures hedging strategies which minimize the foreign exchange risk in USD. (Ahmed & Straetmans, 2015) attempts to predict the cyclical behavior of exchange rates by using five risk factors, viz., violations of uncovered interest rate parity (UIP), relative purchasing power parity (RPPP) and pseudo-parity for equity returns, relative (cross-country) TED spreads and relative term spreads. These factors are found to forecast periods of depreciation or appreciation and subsequent reversals to form trading strategies and hedge their positions as well as re-balance their carry trade positions.

Compared to previous study, this paper avoids hedging strategy which is more risky to the trader’s portfolio. The proposed application also reduces the factors involved in making trading decisions and eradicate complex financial market trend calculations.

III. PROPOSED METHOD IN MODELLING THE EURO (EURO)/US DOLLAR (USD) EXCHANGE RATES

In this paper, the HMM has been applied to investigates the similar data pattern in exchange rates data that occur on a daily basis. Most time series analysis technique involves some form of identifying market trends. Hence, the traders need to decide whether they should buy, sell or hold the position. On the other hand, this study suggest the valuable risk based on daily price projection which the traders need to fill the buying and selling orders in each trading period. The HMM is proven as a powerful tool that work well for sequential data modelling. Hidden Markov Model consists of a finite set of states, each of which is associated with multidimensional probability distribution. Transitions among the states are governed by a set of probabilities called transition probabilities. In a particular state an observation can be generated, according to the associated probability distribution.

HMM contains following elements:
1) \( N \) is Number of hidden states
2) \( Q \) is Set of states \( Q = \{1,2,\ldots,N\} \)
3) \( M \) is Number of symbols
4) \( V \) is Set of observation symbols \( V = \{1,2,\ldots,M\} \)
5) \( A \) is State Transition Probability Matrix

In order to implement HMM, the following three fundamental questions should be resolved
1. Given the model \( \lambda = (A, B, \pi) \) how do we compute \( P(O|\lambda) \), the probability of occurrence of the observation sequence \( O = O_1, O_2, \ldots, O_T \).
2. Given the observation sequence \( O \) and a model \( \lambda \), how to choose a state sequence \( q_1, q_2, \ldots, q_T \) that best explains the observations.
3. Given the observation sequence O and a space of models found by varying the model parameters A, B and \( \pi \), how do we find the model that best explains the observed data.

The exchange rate data is published consists of the opening, closing, highest and lowest price for each time horizon including daily and hourly data. Based on mathematical expression has been researched by (Badge, 2012), this study apply exchange rate prices on opening \( \{o_i\} \), highest \( \{h_i\} \), and lowest \( \{l_i\} \) for each time horizon. The historical data has been used to predict the price range from the opening price to the highest price and from the opening price to the lowest price. The problem can be expressed as follows:

\[
p_{t+1} = \prod_{i=1}^{n} f(o_i, h_i, l_i)
\]

where;

\( i \) = the stage in model on daily basis. In order to apply HMM, the historical data is divided into training and testing data set. Before applying Hidden Markov Model the main thing is to choose number of states and observation symbols. Given an observation sequence and the set of hidden states, we need to learn the Transition probability matrix \( \{A\} \) and the Observation probability matrix \( \{B\} \), which are determined as follows:

\[
A = \{a_{ij}\} \text{ s.t. } a_{ij} = \frac{p[q_{t+1} = s_j / q_t = s_i]}{i \leq 1, j \leq N}
\]

\[
B = \{b_j(K)\} \text{ s.t. } b_j(k) = \frac{p[v_k = t / q_t = s_j]}{1 \leq j \leq N, 1 \leq k \leq M}
\]

Where \( q_t \) is the state at time \( t \). Baum-Welch algorithm is used as a learning algorithm. The Baum-Welch algorithm is a generalized expectation-maximization algorithm which to calculate maximum likelihood estimates and posterior mode estimates for the parameters of an HMM. The Baum-Welch algorithm searching for

\[
\lambda = \max p(O/ \lambda)
\]

which is the HMM \( \lambda \) that maximizes the probability of the observation \( O \).

The application of the algorithm can be expressed as follows:
The forward probability and the backward probability has been calculated for each HMM state. The next stage is to find forward probability that can be expressed as \( \alpha_i(t) = p(O_1= o_1, \ldots, O_t= o_t, Q_t = i) \lambda \), that is the probability of considering the partial sequence \( o_1, \ldots, o_t \) and ending up in state \( i \) at time \( t \). The calculation of \( \alpha_i(t) \) as follows:

\[
\alpha_i(t) = \pi_i b_i(O_t)
\]

\[
\alpha_i(t + 1) = b_i(O_{t+1}) \sum_{i=1}^{N} \alpha_i(t).a_{ij}
\]

The backward probability is defined as the probability of the ending partial sequence \( O_{t+1} \ldots O_T \) given that started at state \( i \), at time \( t \). The calculation of \( \beta_i(t) \) as follows:

\[
\beta_i(T) = 1
\]

\[
\beta_i(t) = \sum_{i=1}^{N} \beta_j(t + 1)a_{ij}b_j(O_{t+1})
\]
The Baum-Welch algorithm is used to find the similar patterns with the current day from the HMM data pattern which assume that the produced likelihood value on day $n$ is $l_p$. Based on the historical data set using the HMM those observation sequences are located which would produce the same or nearly same value of $l_p$. The HMM found the $y$th day’s observation sequences which would produce the same likelihood value $l_{op}$. The price range variable on day ‘$y$’ and ‘$y+1$’ is determined.

$$p_{y+1} = (p_{y+1} - p_y) + o_i \quad (9)$$

According to the above equation, $p_y$ is the projection price of matched data pattern, $p_{y+1}$ equal to next day projection price of matched data pattern and $p_i$ representing the projection price of today’s data.

The HMM is trained to calculate the likelihood values of price range from the open price (Fig. 1) between two data sequences, in which past day price range behavior is first located which is almost similar to that of the current day. The range calculated produces the likelihood values for the current day, and then observation sequences are located which give nearly equal values which are formed from the past data sets. Then price difference of that day’s price and next to that day’s price are taken and hence next day’s price range is forecasted by adding the above difference to the current day’s opening price. In order to train HMM based tool is developed for time series forecasting, based on historical records of daily opening price utilization as depicted in Fig.1. The prior probability $\pi$, a random number has been chosen and normalized:

$$\sum_{i=1}^{N} \pi_i = 1.$$  

In this study, there are three input variables has been used in calculating the exchange rate price range. It consists of the opening, high and low price values of the day. The price projection calculated from the daily open price range which depicted in Fig.2 is the new method introduces in this study compares to previous researchers that concern more on the closing price of the day. The daily price range value as shown in Fig.3 which produced from the projection can be applied in risk and reward calculation in each trading period. In order to find the similar data patterns, the well known Baum-Welch algorithm is applied. The first step in training the parameters of an HMM is to see what states and transitions the model thinks are likely on each trading day. Those likely states and transitions can be used to reestimate the probabilities using the ‘forward-backward” or Baum-Welch algorithm, increasing the likelihood of the training data.

![Fig. 1 Historical daily opening data of Euro per Dollar.](image-url)
Fig. 2  Price Projection in two directions: upward and downward

Fig. 3  Price range between high and low of the day
IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

In order to train the HMM, the dataset is divided into two sets, one training set start from 4/1/2013 (3032 observations) and one test set begin at 29/7/2013 to 5/5/2016 (3574 observations). For example, the study compute the statistical parameters an HMM using the exchange rate of EUR/USD for the period 2/1/2012 to predict the range (TABLE 1) of price target 15/1/2012. The trained HMM and the past data, produced consistent 10 pips daily price range movement has been shown in Fig. 4 as pips gain, for the period 2 January 2012 to 15 January 2012. It can be stated that total pips accumulation for each trading period is greater than 0 which indicates that the probability of profitability is positive value. The Table 1 shows the similarities between two datasets (range of price movement 12 October 2011 and 20 May 2013 is 10 pips). It seems that 28 May 2013 price pattern repeated the price that of 2 January 2012.

TABLE 1
DAILY DATA PATTERN FOR THE SPECIFIED TRADING PERIOD

<table>
<thead>
<tr>
<th>Date</th>
<th>Price</th>
<th>Range from the opening price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Open</td>
<td>High</td>
</tr>
<tr>
<td>12/10/2011</td>
<td>1.29234</td>
<td>1.29304</td>
</tr>
<tr>
<td>20/5/2013</td>
<td>1.29805</td>
<td>1.29915</td>
</tr>
<tr>
<td>Data pattern on 21/5/2013 that match data pattern on 12/10/2011 and 20/5/2013</td>
<td>1.25962</td>
<td>1.26242</td>
</tr>
</tbody>
</table>

Fig. 4 Pips accumulation for overall trading period based on Positive (P) and Negative (N) transitions
The Table 2 shows that Model 1 and Model 2 produce the lowest AIC and BIC values compared to the Model 3. The Model 3 produces the highest AIC and BIC values which cannot be considered as a preferred model. In term of Log likelihood feature value as depicted in Table 3, Model 1 and Model 2 still outperform the Model 3 with 9474.581. Model 3 fail to beat almost entirely the model parameter comparisons, the trace left behind by the training set points, as a result making it the model with the poor parameter estimation gain. The third model produces the lowest AIC and BIC with -18909.16 and -18811.27 values compared to Model 1 and Model 3. It can be seen that in the AIC and BIC parameter values, the first model produce the lowest value. Finally, it can be concluded that the Model 1 has proven has the best parameter estimation compared to Model 2 and Model 3.

<table>
<thead>
<tr>
<th>Model</th>
<th>Difference</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Low and Open</td>
<td>-19633.48</td>
<td>-19535.59</td>
</tr>
<tr>
<td>Model 2</td>
<td>High and Open</td>
<td>-19678.55</td>
<td>-19580.65</td>
</tr>
<tr>
<td>Model 3</td>
<td>Close and Open</td>
<td>-18909.16</td>
<td>-18811.27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>Difference</th>
<th>Log likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Low and Open</td>
<td>9836.742</td>
</tr>
<tr>
<td>Model 2</td>
<td>High and Open</td>
<td>9859.273</td>
</tr>
<tr>
<td>Model 3</td>
<td>Close and Open</td>
<td>9474.581</td>
</tr>
</tbody>
</table>

V. CONCLUSIONS AND FUTURE WORK

This paper introduces the new approach of Hidden Markov model for financial forecasting purposes. The study applies the proposed technique to the task of forecasting the EUR/USD daily exchange rate which compared its result with 3 different models development. In terms of results, the selected technique outperforms all its benchmarks in terms of AIC, BIC, and Log likelihood value and trading efficiency for both in- and out-of-sample periods. The proposed technique produce consistent pips accumulation over the long run. Further improvements needs to be made on the proper selection of variables as inputs, more variables including technical and fundamentals factors. Future research could explore the possibility of coming up with a risk adjustment to produce easy and profitable trading strategy without burdening the
traders with high risk and reward ratio. Recent research done on this field has shown the prospective return and applications that could be obtained for exchange rate forecasting by employing this easy predictive application. Future work could be to examine the impact of fundamental and technical variables, instead of price gaps, on the exchange rate price movement.

ACKNOWLEDGEMENT

The authors like to thank to all reviewers at Computer & Information Sciences, Universiti Teknologi PETRONAS, Perak, Malaysia who greatly helped us.

REFERENCES


