View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Procedia Computer Science
Volume 29, 2014, Pages 2065–2075
ICCS 2014. 14th International Conference on Computational Science

Multi-Scale Foreign Exchange Rates Ensemble for Classification of Trends in Forex Market

Hossein Talebi[1], Winsor Hoang [2], and Marina L. Gavrilova[1]
[1] University of Calgary, Calgary, Canada
[2] CTS Forex Company, Calgary, Canada
htalebi@ucalgary.ca, winsorhoang@ctsforex.com, mgavrilo@ucalgary.ca

Abstract
Foreign exchange (Forex) market is the largest trading market in the world. Predicting the trend of the market and performing automated trading are important for investors. Recently, machine learning techniques have emerged as a powerful trend to predict foreign exchange (FX) rates. In this paper, we propose a new classification method for identifying up, down, and sideways trends in Forex market foreign exchange rates. A multi-scale feature extraction approached is used for training multiple classifiers for each trend. Bayesian voting is used to find the ensemble of classifiers for each trend. Performance of the system is validated using different metrics. The results show superiority of ensemble classifier over individual ones.

Keywords: Foreign Exchange, Multi-scale Features, Multivariate Gaussian Classifier, Bayesian Voting

1 Introduction

Foreign exchange market, also known as Forex is a currency trading market spread all around the globe. According to the Bank for International Settlements [1], average of daily exchange in foreign exchange markets is $5.3 trillion in April 2013. This huge amount of turnover makes this market the largest trading market in the world which is approximately 160 times larger than the New York Stock Exchange. This market is traditionally used by central banks, commercial banks, and hedge funds for currency trading. However, by the advent of the internet and its development, the market became available for small retailers. In Forex market, trading is done by selling and buying currency pairs, i.e. EUR/USD. There are several currency pairs, although the major ones in term of the amount of daily transactions are Euro vs US Dollar (EUR/USD), Australian Dollar vs US Dollar (AUD/USD), Great Britain Pound vs US Dollar (GBP/USD), US Dollar vs Canadian Dollar (USD/CAD), US Dollar vs Swiss Franc (USD/CHF) and US Dollar vs Japanese Yen (USD/JPY).

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2014

© The Authors. Published by Elsevier B.V.
doi:10.1016/j.procs.2014.05.190

The main strategy in this market is to buy low and sell high. For example, a trader figures out that Euro will increase in price against the US dollar, so he/she will buy EUR/USD pair at lower price and when the price appreciates, sell the currency pair to gain profit.

Forex market is open 24 hours a day and 5 days of a week. Due to high volatility of the market, it is important to monitor the market constantly. It is impossible for the human to monitor the market 24 hours a day and perform the manual trading. Moreover, traders may have unrealistic expectations of return with limited risk. Fear causes traders to make poor trading decisions and lack of discipline may cause traders to violate trading rules that they promised to follow [2]. However, traders can avoid these pitfalls by using Expert Advisors. Expert Advisors are computer programs that perform automatic trading with no human emotions involved, and are based on logic and discipline. Expert advisors can monitor the market 24 hours a day and make trades based on their algorithmic discipline.

New advancements in machine learning made it possible for expert advisors to learn from pervious market data and make profitable trades [2,16]. Most of the recent works try to predict the market based on the previous prices. In [3], Yao et al proposed a forex market predictor based on a neural network, which was trained using price data and technical indicators. They tested the system using pairing of American Dollar with five other major currencies. They showed that the system provides promising results except for Japanese Yen. Neural network also used in [4] for forecasting foreign exchange (FX) rates of Australian dollar against six different currencies. In this paper they investigate using of Standard Backpropagation (SBP), Backpropagation with Bayesian Regularization (BPR) and Scaled Conjugate Gradient (SCG) for training neural network using five moving average indicators. Their result showed that SCG model outperformed the other two. In [5], Pacelli et. al designed optimal multilayer perceptron topologies using genetic algorithm multi-objective Pareto-Based. Three best designed topologies could predict three days ahead of last available prices with an accuracy of 60%, 70% and 80% on validation data. A comparison between feed forward neural network and Takagi-Sugeno type neuro-fuzzy system at [6] for forecasting the average monthly forex rates showed that neuro-fuzzy system performed better in term of root-mean-square error (RMSE) and training time. Kamruzzaman et. al [7] introduced Support Vector Machine (SVM) to the forex market. They investigate the effect of different kernel function and regularization parameters on currency trading. Their study revealed that polynomial and radial basis kernel are better choices for forex trading. Unlike the traditional models which decided based on a single model, [8] used bootstrap methods to train multiple learners and then combined the results from each model to make the final decision. They also used neural network as their base learning method. Single indicator does not always produce right signals for trading. In order to alleviate this issue, Liu et. al [9] proposed fusion of multiple indicators based on Dempster-Shafer theory. Their experiment revealed that fused indicator can produce more accurate results than a single one in forex trading. In [10], a machine learning approach with sparse grid combination technique is used for predicting the FX rates. FX rates signal is transformed to a D-dimensional regression problem and sparse grid is used to cope with the curse of dimensionality in the D-dimensional regression problem. Khashei et. al [11] introduced fuzzy autoregressive integrated moving average (FARIMA) models in combination of probabilistic neural classifier for forecasting Forex market which is robust to missing data.

Due to volatility of the market, predicting the exact FX rate is error-prone. Moreover, predicted FX rates do not provide strategies for trading in Forex market. In this paper, we present a new paradigm based on a completely different approach to tackle the problem of automated trading in the Forex market. Instead of predicting the actual FX rates, we devised a new classification approach to *identify trends* in the market. The approach is rooted in over ten years of observing Forex market by the founder of CTS Forex international currency trading company [2]. We, thus, classify Forex market trends into three classes: 1) *up trends*, when the FX rate increase by a certain amount, 2) *down trends*, when the FX rate decreases by a certain amount, and 3) *sideway trends*, when the FX rates fluctuates in a specified interval. Proposed approach uses the *zigzag indicator* to identify these trends in historical data on FX rates. We extract features from these trends using our multi-scale feature extraction approach. Multiple classifiers are trained using these features. Bayesian voting is used to create the ensemble of these classifiers which can recognize trends in the market. By predicting the trend of the market, we can buy the currency pair during up trends and sell it during down trends. The method thus is computationally efficient, and free of prediction errors.

The rest of this paper is organized as follows. In section II we present our feature extraction. Zigzag indicator is introduced to find profitable times from price data. Data clearing and feature multi-scale feature extraction will be discussed. Section III describes the training phase. In this section we describe Gaussian classifiers for classifying up, down, and sideway trend and Bayesian voting is introduced to make the final decision. Section IV presents experimental results and in section V we have our concluding remarks.

2 Feature Extraction

In this section we introduce zigzag technical indicator for indicating up trends and down trends. We present data clearing, features and multi-scale feature extraction method.

2.1 Zigzag Technical Indicator

The zigzag technical indicator is used to illustrate the market trends. It ignores small fluctuation in the rate movement and represents the rate signal using monotone linear approximation. Figure 1 demonstrates the zigzag indicator on foreign exchange rates of EUR/USD. Zigzag indicator operates based on a threshold parameter that indicates the percentage of fluctuation which should be ignored in the estimation. This threshold identifies the reversal point of the trend. For example, if the threshold is set to 10% and the foreign exchange rate changes at least 10% in one direction and then changes for at least 10% in the opposite direction, then the transition point will become a new vertex for zigzag. Connecting consecutive vertices will produce the zigzag trend indicator.

Transition points of the zigzag indicator are potential times for making profitable trades. Trading between two consecutive transition points can be profitable since the market trend is completely known. This property of zigzag indicator cannot be used to trade in real time since in relies on future prices to find the current transition point. However, the pattern of local changes in the data can be useful for recognizing these transition points in real time.

Figure 1: Zigzag trend indicator (black lines) calculated based on foreign exchange rates (red and blue candles)

2.2 Data Clearing and Features

Meta Trader is a trading platform for forex market that can be used to make trades manually and also using expert advisors [2]. It also provides historical data of the market for different currency pairs. The problem with historical data provided by Meta Trader is that there are lots of missing data points which can dramatically influence on training and performance of the system.

To overcome with this problem, we combined historical rate from the Meta Trader 5 platform (which is the current version) with historical rates of Meta Trader 4. Combination of historical data from Meta Trader 5 and 4 cannot completely compensate for missing data. Table 1 shows the frequency of gaps duration in combination of historical rate data. We used bi-linear interpolation to estimate missing data in prices. After combining historical prices, zigzag indicator applied to the data.

The minimums and maximums in the zigzag curves (transition points) indicate locations that the price trend changes. These points can be potential times for selling (maximums of zigzag) or buying (minimums of zigzag) the currency pair because the price will decrease or increase by a specific amount. To extract features from these locations, we consider one hour before and one over after of these minimums or maximum locations as our raw features. These raw features may contain these minimums or maximum locations as our raw features. These raw features may contain interpolated data which are not the actual values of the market. In order to prevent training using too much artificial data, we ignored raw features that have more than 20 interpolated data points.

Raw features which contain the exact prices of the market are not robust to a bias value added to prices. To alleviate this problem we used discrete differentiate of our raw features as the adapted feature that is going to be used for training. The difference between consecutive prices contains required information of local changes regardless of the actual price of the pair.

2.3 Multi Scale Feature Extraction

It is important to have features that can represent large fluctuation of data as well as small ones. Trading platforms provide different sampling rate from prices rate signal. Most of them restrict themselves to 1, 5, 15, 30, 60-minute sampling rates. These sampling rates show the signal in a multi-scale manner. However, there is redundant information inside of the scales. For example, we can construct the whole 15-minute sampling rate data using the 5-minute data. This shows that 5 minute data have all the information of 15-minute data. In order to alleviate this problem, Instead of using traditional sampling rates, we tried to reduce this redundancy by using a subset of prime numbers as the sampling rates. We use 3, 5, 7, 11, 13, 17, 19, 23, and 29. Using 3-minute sampling rate data, we will have fine fluctuations of rate signal and by going up to the 29-minute sampling rate, small fluctuation will vanish and the coarse one will become more vivid. Using this new approach we have a multi scale features with less redundancy.

Gaps Duration (minutes)	Years													
	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
1	0.30%	0.14%	0.09%	0.06%	0.04%	0.03%	0.19%	0.60%	0.26%	0.04%	0.04%	0.01%	0.00%	0.00%
2	0.14%	0.04%	0.03%	0.02%	0.00%	0.00%	0.03%	0.28%	0.13%	0.02%	0.01%	0.01%	0.00%	0.00%
3	0.08%	0.02%	0.02%	0.01%	0.01%	0.00%	0.01%	0.15%	0.06%	0.01%	0.00%	0.00%	0.00%	0.00%
4	0.04%	0.01%	0.01%	0.01%	0.00%	0.00%	0.00%	0.09%	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%
5	0.03%	0.01%	0.01%	0.00%	0.00%	0.00%	0.00%	0.05%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%
6-10	0.06%	0.02%	0.02%	0.01%	0.01%	0.00%	0.01%	0.09%	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%
11-30	0.04%	0.02%	0.02%	0.01%	0.01%	0.00%	0.00%	0.03%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
31-60	0.01%	0.01%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Sum	**0.72%**	**0.27%**	**0.20%**	**0.13%**	**0.08%**	**0.05%**	**0.24%**	**1.28%**	**0.53%**	**0.08%**	**0.07%**	**0.02%**	**0.00%**	**0.00%**

Table 1: Frequency of different gap duration in rates data from year 2000 to 2013 after combination of MT4 and MT5 data.

3 Training

In the training phase, we proposed using multivariate Gaussian classifier [12]. Figure 1 Shows 3-minute features based on their two largest Eigen vectors. As Figure 1 (a) and (b) demonstrate, both up trend and down trend data have dens distributions. This fact leads us to use multivariate Gaussian classifiers to classify these two classes. However, Figure 1 (c) demonstrates that there is a substantial overlap between up trends and down trend features of a same scale. The proposed multi scale features can alleviate this issue by training a multivariate Gaussian classifier on each scale and then use the ensemble of these classifiers to come up with a concrete outcome. In following subsection we will introduce multivariate Gaussian Classifier and Bayesian voting for our ensemble method to make the final decision.

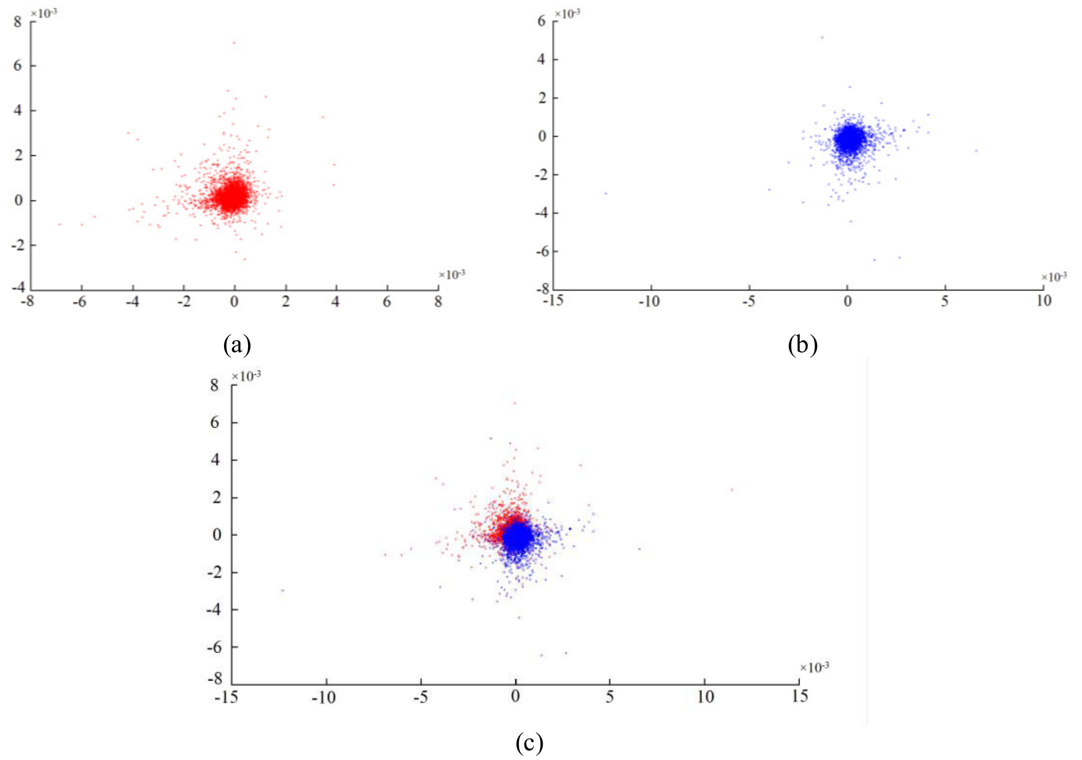

Figure 2: Demonstration of density 3-minute features based on its two largest Eigen vectors. (a) Density of down trend features (b) Density of up trend features. (c) Overlap of uptrend and down trend features.

3.1 Multivariate Gaussian Classifier

Multivariate Gaussian Classifier (MGC) [12] considers that each class of data has a multivariate Gaussian distribution. Let $x^i \in R^n$ be features of class i, if the probability distribution of x^i is Gaussian with mean vector μ and covariance matrix Σ, then [12]:

$$x^i \sim \mathcal{N}(\mu, \Sigma) \quad (1)$$

Where,

$$\mathcal{N}(x;\mu,\Sigma) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)) \quad (2)$$

The goal of this classifier is to find μ and Σ which are parameters of the Gaussian distribution. Given a training set $\{x^1, \ldots, x^m\}_c$ where x^i is a n dimensional vector representing the i^{th} training sample in the class c and X is a $m \times n$ matrix created by training set data, the parameter of Gaussian distribution for this class can be calculated as follow [12]:

- Calculate $\mu = \frac{1}{m}\sum_{i=1}^{m} x^i$, $\mu \in \mathbb{R}^n$.
- Calculate $\Sigma = \frac{1}{m}(X - 1\mu^T)^T(X - 1\mu^T)$

Having the probability distribution of each class, we can use Bayes rule to calculate the membership probability of a new observation x' to a class C_i [13]:

$$p(C_i|x') = \frac{p(x'|C_i)p(C_i)}{p(x')} \quad (3)$$

where,

$$p(x'|C_i) = \mathcal{N}(x';\mu_{C_i},\Sigma_{C_i}) \quad (4)$$

and,

$$p(C_i) \approx \frac{m_{C_i}}{\sum_{j=1}^{K} m_{C_j}} \quad (5)$$

where K is the number of classes, and,

$$p(x') = \sum_{j=1}^{K} p(x'|C_j) * p(C_j) \quad (6)$$

Having the membership probability of each new observation to each class, we can classify new observations to one of C_j classes. If the membership probability is more than an acceptable probability threshold δ it can be considered as a member of a class with an unknown probability density function.

To classify FX rates, we trained Gaussian classifiers for different scales of up trend and down trend features. For each of the trends, nine Gaussian distribution was found that are the hypotheses for that trend. A hypothesis set $H_{c_i} = \{h_1^{c_i}, \ldots, h_9^{c_i}\}$, is the set of all the different scale hypothesis for trend c_i. In this case, we have two hypotheses set $H_{c_{up}}$ and $H_{c_{down}}$. Every feature vector that is ignored by both of these hypothesis sets is classified as sideway trend.

3.2 Bayesian Voting

Each hypothesis defines a probability distribution that can predict a class of data points based on conditional probability. In our setting for predicting Forex market trends, there are nine hypotheses for classifying each trend. More formally $H_{c_i} = \{h_1^{c_i}, \ldots, h_9^{c_i}\}$ is a hypothesis set for discriminating class c_i from the rest of classes in our problem. Prediction of each hypothesis may be different from the other. To combine the result of all the hypothesis in H_{c_i}, we used Bayesian voting technique.

Bayesian voting [14] finds the ensemble of all the hypothesis by making a weighted summation over all the hypothesis in H_{c_i}. Given a new data point x and our training set S, we are interested to find

$p(x\epsilon c_i|\, S, H_{c_i})$. In order to find this probability, summation of hypothesis in $\mathrm{H}_{\mathrm{c_i}}$ weighted by their posterior probability is used. We can write this weighted summation as follows [14]:

$$p(x\epsilon c_i|\, S, H_{c_i}) = \sum_{h\in H_{c_i}} h(x)p(h|S) \qquad (7)$$

where $h(x)$ is the response of the hypothesis h to the data point x and $p(h|S)$ is the posterior probability of hypothesis h.

4 Experimental Results

In this section we explain our training data used for creating classifiers and the test data for validating our approach. Since we do not predict FX rates, as majority of other methods do, there are no comparators in that category. Instead, performance metrics are used to evaluate the developed system.

4.1 Data Collection

FX rate of EUR/USD from the beginning of January 2000 to the end of December 2013 were extracted from both Metatrader 5 and Metatrader 4 database. After combining the results from these databases, the zigzag indicator is applied to the FX rates. FX rates in the interval of one hour before and one hour after maximum and minimum points in the zigzag were extracted as sample points for up trend and down trends. For the sideway, sample data was extracted from intervals between maximum and minimum points of zigzag. Sample points that have more than 20 missing FX rates were ignored for training of the classifier. Samples that are extracted from of January 2000 to the end of December 2010 were used for training the classifier. The rest of samples are used for testing the classifier.

4.2 Performance Metrics

To evaluate the performance of the system we used three performance metrics namely, recall, precision, and accuracy [15]. In order to define metrics, we should introduce true positive (tp), true negative (tn), false positive (fp) and false negative (fn). True positive (tp) is the number of data points that are member of class $\mathrm{C_i}$ and are classified correctly as class $\mathrm{C_i}$. True negative (tn) is the number of data points that are not members of class $\mathrm{C_i}$ and are classified correctly as not members of class $\mathrm{C_i}$. False positive (fp) is the number of data points that are not members of class $\mathrm{C_i}$ and are classified wrongly as member of class $\mathrm{C_i}$. False negative (fn) is the number of data points that are member of class $\mathrm{C_i}$ and are classified wrongly as not members of class $\mathrm{C_i}$. We can define recall, precision, and accuracy using the following formula [15]:

$$Recall = \frac{tp}{tp + fn} \qquad (8)$$

$$Percision = \frac{tp}{tp + fp} \qquad (9)$$

$$Accuracy = \frac{tp + tn}{tp + fn + fp + tn} \qquad (10)$$

Recall shows the ratio correctly classified data points of class $\mathrm{C_i}$ to the all number of all the data points classified as class $\mathrm{C_i}$. Precision is the ratio correctly classified data points of class $\mathrm{C_i}$ to the number of all data points that should be classified as class C_i. Accuracy is the proportion of correctly classified data points.

4.3 Test Results

Two sets of hypothesis for up trend and down trend data were trained. For each set we trained nine hypotheses which are correspondence with different sampling rate of FX rates. We used Bayesian voting to combine each set of hypothesis and make decisions for classifying test data. If the membership probability of a new data point is more than an acceptance probability threshold (δ), then it is considered as a member of the corresponding class. If the new data point does not classify as up or down trend, we consider it to be a sideway trend. The performance of the system is measured in term of recall, precision, and accuracy. In order to show the superiority of ensemble method performance measures is reported for each of the individual hypothesis and ensemble hypothesis. Table 2, 3, and 4 show the precision, recall and accuracy of each hypothesis individually for different acceptance probability threshold (δ). In the last row of each table the performance measure is reported for the ensemble of classifiers (H). Performance measures in each table show that ensemble classifier H can outperform every individual classifier. Recall of ensemble hypothesis shows that a promising number of up and down trends are recognized correctly. Due to similarity of patterns in sideway trend and up and down trend, some of sideway trends were wrongly recognized as up and down trends. The low precision value is due to this similarity. However, these misclassifications do not result in losing of money in trades since it is not misclassification of up and down trends.

		Acceptance Probability Threshold (δ)								
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
Hypotheses	h_1	0.992	0.987	0.977	0.961	0.928	0.882	0.819	0.721	0.548
	h_2	0.997	0.995	0.987	0.978	0.955	0.911	0.835	0.721	0.540
	h_3	0.997	0.993	0.988	0.979	0.958	0.917	0.838	0.724	0.537
	h_4	0.998	0.996	0.990	0.981	0.962	0.908	0.830	0.705	0.520
	h_5	0.998	0.995	0.991	0.983	0.959	0.910	0.825	0.699	0.506
	h_6	0.998	0.996	0.994	0.982	0.962	0.917	0.832	0.702	0.515
	h_7	0.997	0.995	0.988	0.969	0.944	0.885	0.779	0.644	0.456
	h_8	0.996	0.993	0.986	0.977	0.947	0.885	0.782	0.630	0.444
	h_9	0.997	0.996	0.989	0.970	0.941	0.875	0.768	0.629	0.440
	H	**1.000**	**0.999**	**0.996**	**0.990**	**0.970**	**0.927**	**0.854**	**0.749**	**0.554**

Table 2: Recall of each hypothesis based on the test data

		Acceptance Probability Threshold (δ)								
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
Hypotheses	h_1	**0.041**	0.045	0.051	0.059	0.069	0.082	0.099	0.124	0.156
	h_2	0.040	0.044	0.050	0.058	0.069	0.085	0.107	0.135	0.172
	h_3	0.040	0.044	0.050	0.057	0.069	0.086	0.108	0.135	0.174
	h_4	0.040	0.044	0.049	0.057	0.069	0.087	0.111	0.141	0.181
	h_5	0.040	0.044	0.049	0.057	0.069	0.087	0.109	0.140	0.180
	h_6	0.040	0.044	0.049	0.057	0.070	0.088	0.110	0.139	0.182
	h_7	0.039	0.043	0.048	0.055	0.068	0.086	0.110	0.141	0.182
	h_8	0.039	0.043	0.048	0.056	0.068	0.086	0.109	0.136	0.178
	h_9	0.040	0.043	0.048	0.056	0.068	0.084	0.106	0.134	0.171
	H	0.040	**0.046**	**0.053**	**0.064**	**0.074**	**0.098**	**0.123**	**0.165**	**0.230**

Table 3: Precision of each hypothesis based on the test data

	Acceptance Probability Threshold (δ)								
Hypotheses	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
h_1	0.162	0.248	0.340	0.436	0.531	0.625	0.717	0.803	0.877
h_2	**0.149**	0.228	0.318	0.422	0.533	0.642	0.741	0.823	0.890
h_3	0.147	0.228	0.316	0.419	0.533	0.645	0.743	0.824	0.892
h_4	0.139	0.215	0.305	0.413	0.533	0.652	0.754	0.835	0.898
h_5	0.137	0.216	0.307	0.413	0.533	0.651	0.752	0.834	0.899
h_6	0.138	0.217	0.309	0.413	0.533	0.651	0.751	0.833	0.899
h_7	0.126	0.201	0.291	0.403	0.532	0.658	0.764	0.845	0.907
h_8	0.126	0.202	0.293	0.405	0.532	0.657	0.762	0.843	0.906
h_9	0.128	0.204	0.296	0.408	0.532	0.653	0.758	0.841	0.903
H	0.146	**0.254**	**0.355**	**0.469**	**0.558**	**0.688**	**0.775**	**0.855**	**0.917**

Table 4: Average Accuracy of each hypothesis based on the test data

Figure 3 shows the Receiver Operating Characteristic (ROC) curve for every hypothesis (h_i) and the ensemble (H). As the figure 3 illustrates, for a specific false positive rate, ensemble classifier always has the highest true positive rate.

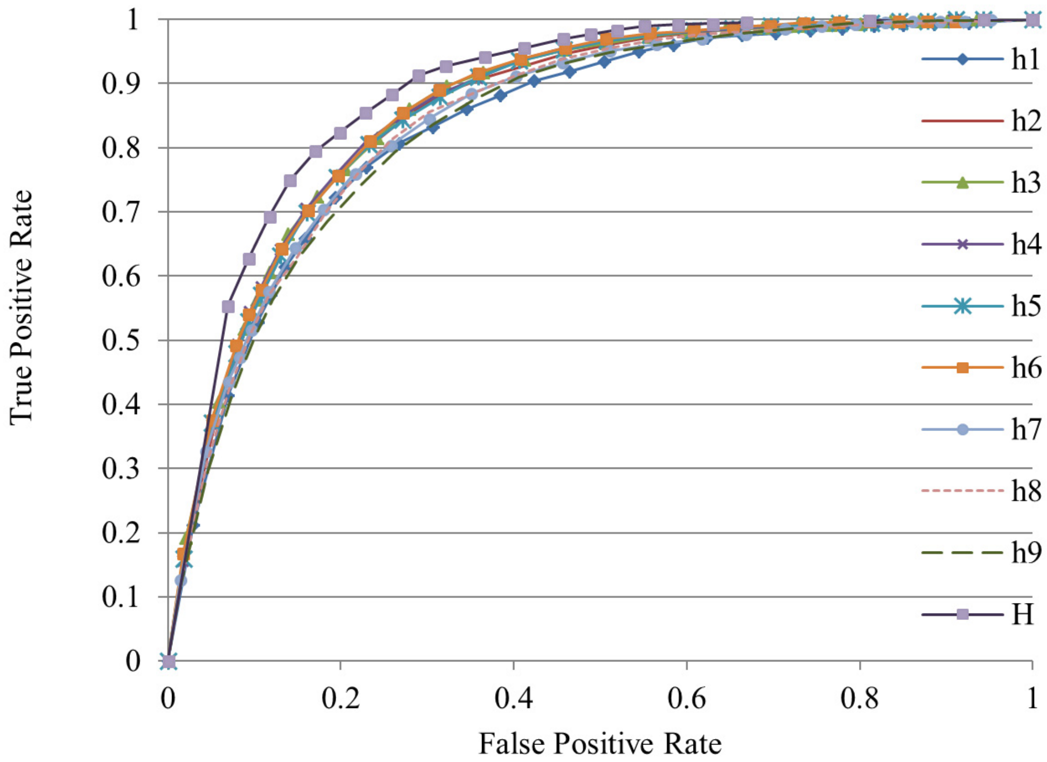

Figure 3: Receiver Operating Characteristic (ROC) curve of different hypotheses based on test data.

5 Concluding Remarks

In this paper, we presented a radically new approach for automated trading in the Forex market. The key methodological development is in introducing a classification method which uses multi-scale features extracted from FX rate. The underlying distribution of each scale feature was calculated as a classifiers and Bayesian voting method used to find the ensemble of these classifiers. Recall, precision, and average accuracy showed the superiority of the ensemble classifier. Experimental results showed that the proposed system is able to identify up and down trends in the FX rate signal accurately.

Future direction for improvement includes extracting more features from FX rate and analyzing the performance of other ensemble methods to combine the results of classifiers.

Acknowledgements. Authors would like to acknowledge support from NSERC, MITACS and CEO of CTSForex for the conducted research.

References

[1] "Triennial Central Bank Survey: Foreign Exchange turnover in April 2013: preliminary global results." Bank for International Settlements (2013).
[2] Hoang, Winsor. "The Bull, the Bear, and the Baboon: Fx Lessons Learned the Hard Way", Createspace, (2013)
[3] Yao, Jingtao, and Chew Lim Tan. "A case study on using neural networks to perform technical forecasting of forex." Neurocomputing 34.1, 79-98. (2000)
[4] Kamruzzaman, Joarder, and Ruhul A. Sarker. "Forecasting of currency exchange rates using ANN: A case study." Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on. Vol. 1. IEEE, 793-797. (2003).
[5] Pacelli, Vincenzo, Vitoantonio Bevilacqua, and Michele Azzollini. "An Artificial Neural Network Model to Forecast Exchange Rates." JILSA 3.2A, 57-69. (2011)
[6] Abraham, Ajith, and Morshed U. Chowdhury. "An intelligent forex monitoring system." Info-tech and Info-net, 2001. Proceedings. ICII 2001-Beijing. 2001 International Conferences on. Vol. 3. 523-528 IEEE. (2001)
[7] Kamruzzaman, Joarder, Ruhul A. Sarker, and Iftekhar Ahmad. "SVM based models for predicting foreign currency exchange rates." Data Mining, 2003. ICDM 2003. Third IEEE International Conference on. IEEE, 557-560. (2003).
[8] He, Haibo, and Xiaoping Shen. "Bootstrap methods for foreign currency exchange rates prediction." Neural Networks, 2007. IJCNN 2007. International Joint Conference on. IEEE, 1272-1277. (2007)
[9] Liu, Zhihong, and Deyun Xiao. "An automated trading system with multi-indicator fusion based on DS evidence theory in forex market." Fuzzy Systems and Knowledge Discovery, 2009. FSKD'09. Sixth International Conference on. Vol. 3. IEEE, 239-243. (2009)
[10] Garcke, Jochen, Thomas Gerstner, and Michael Griebel. "Intraday foreign exchange rate forecasting using sparse grids." Sparse grids and applications. Springer Berlin Heidelberg, 81-105. (2013)
[11] Khashei, Mehdi, Farimah Mokhatab Rafiei, and Mehdi Bijari. "Hybrid Fuzzy Auto-Regressive Integrated Moving Average (FARIMAH) Model for Forecasting the Foreign Exchange Markets." International Journal of Computational Intelligence Systems 6.5, 954-968. (2013)
[12] Redner, Richard A., and Homer F. Walker. "Mixture densities, maximum likelihood and the EM algorithm." SIAM review 26.2, 195-239. (1984)
[13] Domingos, Pedro, and Michael Pazzani. "On the optimality of the simple Bayesian classifier under zero-one loss." Machine learning 29.2-3, 103-130. (1997)

[14] Dietterich, Thomas G. "Ensemble methods in machine learning." Multiple classifier systems. Springer Berlin Heidelberg, 1-15. (2000)
[15] Sokolova, Marina, and Guy Lapalme. "A systematic analysis of performance measures for classification tasks." Information Processing & Management 45.4, 427-437. (2009)
[16] Tian, Yousheng, Yingxu Wang, Marina L. Gavrilova, and Guenther Ruhe. "A formal knowledge representation system (FKRS) for the intelligent knowledge base of a cognitive learning engine." International Journal of Software Science and Computational Intelligence (IJSSCI) 3, no. 4: 1-17. (2011)