IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received November 20, 2019, accepted December 11, 2019, date of publication December 17, 2019,
date of current version December 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2960379

Prediction and Trading in Crude Oil Markets Using Multi-Class Classification and Multi-Objective Optimization

SHANGKUN DENG[1,2], XIAORU HUANG[1], JIASHUANG SHEN[1], HAORAN YU[1], CHENGUANG WANG[1], HONGYU TIAN[1,2], FANGJIE MA[3], AND TIANXIANG YANG[4]

[1]College of Economics and Management, China Three Gorges University, Yichang 443002, China
[2]Financial Research Institute, College of Economics and Management, China Three Gorges University, Yichang 443002, China
[3]Department of Humanity and Law, College of Science and Technology, China Three Gorges University, Yichang 443002, China
[4]School of Creative Science and Engineering, Waseda Univesity, Tokyo 169-8555, Japan

Corresponding author: Shangkun Deng (dsk8672@163.com)

This work was supported in part by the Hubei Provincial Department of Education under Grant Q20171208, in part by the Starting Grant of China Three Gorges University under Grant 20170907, and in part by the Hubei Provincial Department of Education under Grant Talent Excellence Program 2019.

ABSTRACT Crude oil price direction forecasting presents an extremely challenging task that attracts considerable attention from academic scholars, individual investors and institutional investors. In this research, we proposed an integration method by adopting the Multi-Class Support Vector Machine (MCSVM) and the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) for forecasting and trading simulation in two well-known crude oil markets. Firstly, the proposed approach applied the MCSVM to train a multi-class classification model, and it adopted the NSGA-II to optimize the threshold values of trading rules. Then, the trained MCSVM model was used to forecast the movement direction and magnitude levels. Next, the proposed method forecasted the direction of crude oil price movements one week later and executed trading simulation according to the direction and magnitude level predictions. Finally, after a testing period lasted for four years, the performances of the proposed approach were gauged in terms of direction prediction correctness and investment yields. Experimental results demonstrated that the proposed approach produced outstanding results not only on hit ratio and accumulated return but also return-risk ratio. It indicates that the proposed approach can provide beneficial suggestions for individual investors, institutional investors, as well as for government officers engaged in energy investment policies making.

INDEX TERMS Crude oil, multi-class classification, direction prediction, multi-objective optimization, simulation trading.

I. INTRODUCTION

Crude oil is recognized as one of the essential energy sources, and it is one of the largest and most actively traded commodities in the world [1]. Among crude oil markets, West Texas Intermediate (WTI) and Brent are the two most influential ones. WTI is priced by using the crude oil sourced from North America, whereas Brent crude is from North Sea [2]. A large group of individual and institutional investors, such as governments, energy-related companies, market investors, as well as hedge funds are all involved in these two crude oil markets, and their main concern about the market is the volatility and movement direction of the crude oil prices. Therefore, if the direction of crude oil price fluctuation can be predicted correctly to some extent, it will be of great benefits for them, for instance, it would be extremely beneficial for the government to formulate policies, helpful for energy-related enterprises to make decisions, and it could provide investment advice for individual investors and institutional investors.

As a traditional time series analysis method, ARIMA (Autoregressive Integrated Moving Average) model [3] has been applied to crude oil prices prediction by a lot of researchers around the world. For instance, Moshiri and Foroutan predicted crude oil future prices using an ARIMA

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Liu.

based method [4]. Mohammadi and Su applied an ARIMA based approach to predict crude oil spot prices five business days later on eleven international crude oil markets [5]. Nonetheless, due to the nonlinear and non-stationary characteristics of crude oil prices [6], there was an increasing number of researchers started to apply machine learning methods for predicting the movement of crude oil prices. Among machine learning methods, the ANN and SVM are ones that have been most widely applied by researchers. ANN is a computing system, and it was proposed by inspiration of the biological neural networks that constitute animal brains. It is renowned for its ability to perform well for solving nonlinear problems, and it has been widely used by researchers on the direction prediction of crude oil prices. For instance, Li *et al.* predicted the trends of crude oil prices by using an ANN-based method and online news articles [7]. Yu *et al.* developed an ANN-based approach for crude oil price direction prediction [8]; Wu *et al.* applied an improved neural network-based method for prediction of crude oil prices [9]. Jammazi and Aloui proposed a crude oil prices forecasting method based on ANN-based method [10]. ANN-based methods have shown the powerful capability for crude oil prices predictions, yet they are susceptible to the over-fitting problem [11]. From the perspective of machine learning methodology, SVM [12] has strong robustness to over-learning caused by numerous irrelevant features. Lots of researchers have found that SVM significantly outperformed ANN in their applications [13]–[15], and its effectiveness has also been examined in the application of crude oil price prediction or other financial market price forecasting. For example, Qi and Zhang proposed an SVM based method to forecast the price fluctuations in international crude oil markets [16]; Yasir *et al.* developed an SVM based system to predict the foreign exchange rates [17]. Khashman and Nwulu proposed an intelligent prediction model for crude oil price prediction by using an SVM based method [18]. Excellent performances in those studies indicate that SVM is a suitable method for price prediction in crude oil markets.

Although scholars in academic had achieved outstanding performances for predicting crude oil movements with SVM, in the previous studies, it was found by us that when scholars adopted the SVM to predict the direction of crude oil prices, they generally considered the forecasting task as a two-category classification problem, that is, to predict a rise or a fall of crude oil price movement. However, in the price movement of the crude oil market, most of the movements could be considered as very slight fluctuations. In our preliminary analysis of the daily returns of Brent crude oil price movements in the year 2008, among all the positive returns, the minimum value was about 0.04%, and the maximum value among all negative returns was −0.01%. The median (mean) of all positive returns was about 1.1% (1.3%), and the median (mean) of all negative returns was about −1.08% (−1.49%). Additionally, we found that about 47% (30%) of the absolute daily returns of Brent crude oil prices were smaller than 1% (0.5%), which were considered by us as small fluctuations. Thus, if the problem of crude oil direction prediction is simply regarded as a two-class problem, the prediction of small fluctuations may bring serious errors since intuitively it is not easy to predict correctly for relatively very small movements that like a noise, for instance, a return of 0.04%. Indeed, for governments, enterprises or investors, most of them concerned more about whether there will be a substantial change in crude oil prices in the near future. Therefore, if we only forecast and trade for predictions with relatively larger return changes, we may filter out the noise of small movements, and subsequently the predictor is expected to enhance the accuracy of oil market direction prediction and transaction returns.

In this study, instead of using an ordinary binary SVM, we decide to pursue the use of multi-class SVM (MCSVM) to classify the fluctuations of crude oil prices. Many researchers have applied multi-class SVM and achieved outstanding performances [19]–[21]. In our multi-class SVM model, we defined the price changes in crude oil prices into six magnitude levels: Small Positive (SP), Medium Positive (MP), and Large Positive (LP), which are three levels of positive returns, as well as Small Negative (SN), Medium Negative (MN) and Large Negative (LN), which belong to three levels of negative returns. In the direction forecasting and simulation trading, we would not forecast or trade if the movement prediction by the MCSVM belongs to SP or SN. Instead, we would conduct the direction forecasting and open the trading position when the predicted market price changes belong to large or medium movements. However, how to set accurate threshold values for each movement magnitude level becomes a problem. Indeed, whether the movement is large or medium generally are defined by professional traders according to their experiences. Also, for market participants, in order to prevent sudden huge transaction risk or accumulated profit and loss risk in the transaction, it is necessary to design a stop order in the trading rule, which is often achieved by a loss-cutting or profit-taking stop order. Whereas, there are several problems such as how to determine the accurate value of thresholds for them.

To find an appropriate trading rule, numerous scholars have adopted GA (Genetic Algorithm) [22], PSO (Particle Swarm Optimization) [23], or other evolutionary algorithms to address the problem of parameter selection. For instance, Hirabayashi *et al.* applied GA to generating the optimized parameters of trading rule for exchange rate trading [24]. Deng *et al.* developed a GA based method for short-term trading on three influential currency pairs [25]. Although GA can be used to address the problem of parameter optimization, it was developed and employed for single-objective optimization, which was used by most scholars in trading applications that their objective function was generally set to be investment return [26]–[28]. Nonetheless, market participants often take into account not only the investment return but also about the accuracy of direction prediction [29]. They prefer to obtain a larger return, and they also hope to predict the direction correctly. Indeed, correct direction predictions in financial markets could be beneficial for market participants that engaged in options transaction [30]. However, the maximization of direction prediction accuracy and investment return sometimes might be conflicted with each other, since market participants may suffer more losses in the incorrect direction predictions than the positive yields generated in the correct ones. Therefore, we expect that the model can perform well at both investment yield and direction prediction accuracy. In general, researchers or market participants adopt the hit ratio to gauge the correctness of direction prediction of the model. In our proposed approach, a multi-objective optimization method was subsequently adopted. Among the multi-objective optimization methods, NSGA-II is a commonly used method in the application of many research fields. For instance, Alkayem *et al*. applied the NSGA-II to selecting the optimal parameters for welding processes [31]. Huang *et al*. applied a NSGA-II based method for customer churn prediction in telecommunications [32]. Successful applications of the NSGA-II in above studies suggest that it is an alternative method for parameter optimization in solving multi-objective problems.

In the present study, we proposed an MCSVM-NSGA-II model for crude oil prices' direction prediction and trading simulation. Firstly, it collected the crude oil data and perform the pre-procession on the data. Then, it divided the data into in-sample and out-of-sample periods; Next, the levels of return movement were classified into six categories. The values of movement levels, loss-cutting and profit-taking were designed as parameters which were optimized in the NSGA-II; Next, we employed the NSGA-II to optimize the parameters, in which the objective functions were the maximization of the accumulated returns and hit ratio; Lastly, we applied the trained MCSVM model for direction prediction and trading simulation. Note that we only forecasted and traded when the predicted level was medium or large. For the selection of suitable evaluation measures, we not only chose the hit ratio and return but also considered that it is necessary to employ the Sharpe ratio to gauge the return/risk ratio of the proposed approach. Since we decided not to trade when the predictions were "Small Positive" or "Small Negative", we only performed direction prediction and trading if the predicted movements were not too slight.

The main contributions of this study could be summarized as follows: 1) Multi-class SVM was adopted to classify the magnitude level of crude oil price movements. Instead of forecasting and trading for each time, we applied it to forecasting and trading when there were relatively larger fluctuation predictions; 2) NSGA-II was applied to optimize an extremely sophisticated trading rule, which takes multiple movement levels, accumulated return and hit ratio into account comprehensively; and 3) For trading rule design, we considered the possible risks in the market. Thus, we have implemented a mechanism of profit-taking and loss-cutting to control the risk. All the parameters related to the trading rule were not from the experience of experts but learned by using the historical data, which was considered to be more trustworthy. Furthermore, experiments of many benchmark methods were conducted to examine the effectiveness of the multi-class classification and multi-objective optimization.

The rest of this article is organized as follows. Section 2 explains the background of related algorithms employed in this study, including multi-class SVM and NSGA-II; Section 3 describes the structure of the proposed method in details; Experimental design and data are explained in section 4; Experimental results and discussions are reported in section 5. Finally, section 6 summarizes this study and provides some future research directions.

II. METHODS

A. BINARY AND MULTI-CLASS SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is a machine learning algorithm which is initially developed for binary classification problems, and it has exhibited good performances on a wide range of application fields [33]–[36]. The most common type of SVM is a binary classifier, which constructs a hyperplane that maximizes the margin of separation between samples of two classes. The binary classifier SVMs can be extended into a multi-class classifier. The main objective of the SVM is to obtain a function $f(x)$ that determines the decision boundary or hyperplane. Two classes of the input data points are optimally separated by this hyperplane. MCSVM classifier positions the decision boundary by using a maximal margin among all possible hyperplanes. $\|w\|$ is minimized subject to conditions for maximizing the margin M:

$$\min \frac{\|w\|^2}{2}$$
$$\text{s.t. } y_i\left(\langle w, x_i\rangle + b\right) \geq 1, \quad i = 1, \ldots, n \tag{1}$$

where w is a vector that defines the boundary, n denotes the number of input data of MCSVM. x_i is the input data point, and b is an offset. The optimal hyperplane of SVM ($f(x)$) is as follows:

$$f(x) = \sum_{n=1}^{t} y_n a_n \langle x_n, x\rangle + b \tag{2}$$

where x_n are the support vectors. To adapt the binary SVM into a classifier for multi-class classification problems, we decide to adopt the OVO (One-Versus-One) strategy, which has been widely employed due to its efficiency, simplicity and outstanding classification performances [37]. The OVO method constructs $n(n-1)/2$ binary SVMs for an n-class classification problem by taking into consideration all binary combinations of classes. In the OVO strategy, each of the $n(n-1)/2$ SVMs are trained using the samples from two classes. Data samples are partitioned by a series of optimal hyperplanes, which indicates that the optimal hyperplane training data is maximally distant from the hyperplane itself. It can achieve the smallest classification error when performing classification of the current trading set by using this hyperplane.

After all the $n(n-1)/2$ classifiers are constructed, we employ a Max-Win voting strategy [38] in this research to count how often each class is outputted by the binary SVMs, and the testing samples are assigned to the most voted class. Each of the $n(n-1)/2$ SVM classifiers casts one vote. If the m-th SVM classified x_i as the p-th class, then the vote of x_i for the p-th SVM is added by one. Otherwise, the q-th is increased by one. The p and q denote the two arbitrary classes separated by an optimal hyperplane in n classes. Subsequently, samples x_i can be classified in the class with the highest vote.

B. NSGA-II

NSGA-II (Non-Dominated Sorting Genetic Algorithm II) has been applied in lots of applications for multi-objective optimization problems [39]–[41]. It is an algorithm that combines the genetic algorithm and the concept of non-dominance introduced by Goldberg [22]. The elitist non-dominated sorting method is the main feature of the NSGA-II algorithm, and crowding distance is incorporated to achieve the spread on the Pareto front. The main procedures of the NSGA-II are listed as follows:

Step 1) A population $\mathrm{P_t}$ of size N is randomly initialized.

Step 2) Using the genetic operators of tournament selection, crossover and mutation to generate an offspring population $\mathrm{Q_t}$.

Step 3) Combining $\mathrm{P_t}$ and $\mathrm{Q_t}$ to create a population $\mathrm{R_t}$ of size $2N$.

Step 4) Sorting the combined population $\mathrm{R_t}$ according to non-domination to obtain different non-dominated fronts $\mathrm{F_i}$.

Step 5) Generating the new population $\mathrm{P_{t+1}}$ of size N, and the non-dominated fronts $\mathrm{F_i}$ are included until the new population $\mathrm{P_{t+1}}$ is filled.

NSGA-II adopts the use of fast non-dominated sorting to obtain the different non-dominated fronts $\mathrm{F_i}$. It sorts the population according to the level of non-domination, and every solution in the population is compared with other solutions to verify whether it is dominated. Additionally, in order to keep the population diversity in each non-dominated front, the crowding distance, which is the average distance of two points on either side if a particular solution, is calculated. The extreme points of each front are assigned with an infinite distance. The algorithm of calculating crowding distance is shown in Fig. 1.

```
n: size of the non-dominated solutions in S
m: number of the objectives
F[i]: the crowding distance of the i-th individual
f[i,j]: the j-th objective value of the i-th individual
Initialize population
for i= 1 to n do
    CD[i] = 0
end for
for j= 1 to m do
    Sort the solution S with objective j in ascending order
    CD[1] = CD[n]=inf
    for i= 2 to n– 1 do
        CD[i] = CD[i] + (f[i+1,j] –f[i–1,j])/(max(f[*,j]) –min(f[*,j]))
    end for
end for
```

FIGURE 1. Procedures of crowding distance calculation.

III. PROPOSED METHOD

A. STRUCTURE OF THE PROPOSED METHOD

The structure of the proposed method is shown in Fig. 2, in which we can find it mainly consists of the following four parts:

1) Data Pre-processing (DP)
2) Multi-class Classification and Multi-objective Optimization (MCMO)
3) Direction Forecasting and Trading (DFT)
4) Performance Evaluation (PE)

1) DP PART

This part is designed for obtaining the data and data preprocessing. We selected WTI and Brent's crude oil market price data, mainly because they are the most influential benchmarks in the international crude oil markets. The original daily spot price data of WTI and Brent were sourced from the website of U.S. EIA (Energy Information Administration). The vacancy data on the non-trading days of Brent or WTI were replaced by the spot price one day before. Next, the whole period of the data for Brent and WTI market prices were separated into four training datasets and testing datasets.

2) MCMO PART

This part first applies the multi-class SVM algorithm to identifying the price movement direction and level. In the meanwhile, it adopts the NSGA-II algorithm to optimize the movement magnitude level threshold values as well as the parameters of the trading rule. For designing the fitness function of the NSGA-II, both the hit ratio and accumulated return produced in the training period were taken into account.

3) DFT PART

In this part, firstly it generates the trained MCSVM model and the optimized trading rule from the previous MCMO part. Then, it applies the trained MCSVM model over the testing periods. Subsequently, the forecasted direction and magnitude level are generated. Then, based upon the direction and level prediction, the proposed method executes a simulation trading over the testing periods.

4) PE PART

Before the PE part, the proposed method has suggested a trading signal of a long or short-selling transaction, or there would be no trading action if the magnitude level forecasting belongs to small positive or small negative (see III.B).

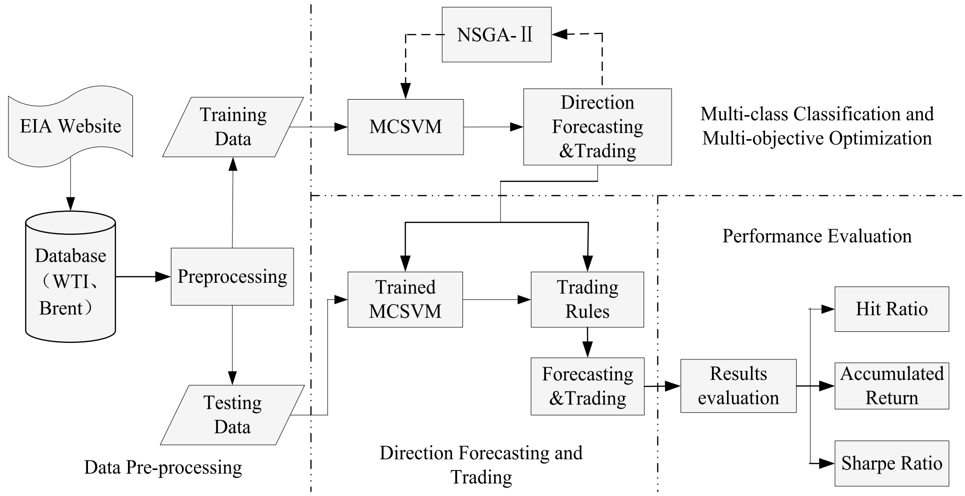


FIGURE 2. Structure of the proposed approach MCSVM-NSGA-II approach from the aspects of direction forecasting and simulation trading.

TABLE 1. Price movement class and parameters of trading rule (totally twelve parameters). N/A means no profit taking or loss cutting setting because of no trading for the movement level prediction of "Small Positive" and "Small Negative".

Level abbreviation	Level meaning	Level threshold	Direction forecasting	Transaction	Profit-Taking threshold	Loss-Cutting threshold
LP	Large Positive	$>T_1$	Upward	Long	PT_1	LC_1
MP	Medium Positive	$(T_2, T_1]$	Upward	Long	PT_2	LC_2
SP	Small Positive	$(0, T_2]$	Uncertain	Wait	N/A	N/A
SN	Small Negative	$[T_3, 0)$	Uncertain	Wait	N/A	N/A
MN	Medium Negative	$[T_4, T_3)$	Downward	Short	PT_3	LC_3
LN	Large Negative	$<T_4$	Downward	Short	PT_4	LC_4

At last, three well-known evaluation criteria, which are the hit ratio, accumulated profit, and Sharpe ratio, were employed to assess the effectiveness of the proposed approach from the aspects of direction forecasting and simulation trading.

B. OPERATION PROCEDURES OF THE PROPOSED APPROACH

In the MCSVM model for multi-class classification, there are in total six classes for classification: Large Positive (LP), Medium Positive (MP), Small Positive (SP), Small Negative (SN), Medium Negative (MN), and Large Negative (LN). The proposed approach applies the NSGA-II to optimizing the parameter of level thresholds ($T_1 \sim T_4$) and profit-taking/loss cutting thresholds ($PT_1 \sim PT_4$, and $LC_1 \sim LC_4$). After the parameter optimization, the MCSVM classify the movement direction and level by using the parameters from NSGA-II, and then the proposed approach forecasts the price movement direction and executes the trading. In the crude oil market, the position a market participant is either long or short. A long position means a market participant would buy while a short position means a short-selling transaction. We supposed that a trading rule comprises the upward/downward direction forecasting, the way to open a position based on the trading signal, and how the position is closed. Loss-cutting and profit-taking thresholds were implemented for minimizing the loss or ensuring the positive return if the market moves rapidly against the expected direction. Finally, since different movement levels might have different characteristics, our proposed method employed different profit-taking/loss cutting thresholds for different movement levels. Therefore, there were in total of twelve parameters for optimization, and their explanations are shown in Table 1.

The operation procedures of direction forecasting and trading simulation for the proposed approach is as follows:

Step 1. Obtaining the daily spot price time series: P_1, P_2, P_3, ..., P_n.

Step 2. Setting the level thresholds for multi-class classification, and designing the trading rule for simulation trading. Since there is no trade if the predicted level is SP or SN, we just implemented profit-taking threshold and loss-cutting threshold values for the class of LP, MP, MN, and LN.

Step 3. Using the NSGA-II to generate values for those parameters. The fitness functions of NSGA-II are set to be hit ratio and accumulated return in the training period.

Step 4. Calculating the labels of data for MSSVM, train the MCSVM model in the training period, and then test the performances over the testing periods.

Step 5. The NSGA-II uses hit ratio and accumulated return to evaluate the performances.

Step 6. Repeating steps 3 to 5 until the final generation of the NSGA-II algorithm. Then we can obtain the optimized parameters for level thresholds of MCSVM (T_1, T_2, T_3 and T_4), and optimized the trading rule ($PT_1 \sim PT_4$ and $LC_1 \sim LC_4$).

Step 7. Applying the MCSVM model and trading rule constructed by GA to the testing periods, we obtain the direction prediction and execute the simulation trading.

Step 8. Conducting out-of-sample direction predictions and simulation trading for four years.

Step 9. Evaluating the performances based on the results of hit ratio, annual accumulated return and Sharpe ratio.

IV. EXPERIMENT DESIGN

A. EXPERIMENTAL DATA

The daily based spot price data of the WTI and Brent were employed for the experiments. The data period ranges from January 1986 to December 2018, whereas for Brent crude oil, its spot price data starts from May 1987 and ends at December 2018. Both of the two datasets were sourced from the U.S. EIA website [42]. In this study, we designed four training and testing datasets, and each dataset was separated into four periods: 1) MCSVM and NSGA-II training period. It is used to perform the MCSVM model training and used for NSGA-II parameters training; 2) NSGA-II testing period; 3) MCSVM training period; 4) Direction prediction and simulation trading period. Note that we adopted a rolling window method to separate the training and testing dataset and lengths of the in-sample periods and out-of-sample test periods were fixed throughout our experiment. For each dataset of training and testing periods, the data was divided with a length ratio of 8:2:10:1. The details of each dataset and the relationships between the training and testing periods of four datasets are shown in Fig. 3 and Table 2.

B. BENCHMARK METHODS

Table 3 displays a summary of the proposed approach MCSVM-NSGA-II and the benchmark methods for experiments. Method 1 and method 2 adopt an ANN-based and an SVM based method for crude oil price direction forecasting and trading simulations, respectively. To keep consistency with the proposed approach, the input features for the ANN and SVM based methods are also the five days' returns. It should be noted that the input returns for all methods except the BAH and SAH have been pre-processed by normalization. In the experiments, as recommended by Thomason [43], a horizon of five days for direction forecasting was chosen. Moreover, we assumed that the movement's direction of five business days later is dependent on the latest five days' price movements, thus the five days' spot price sequence was selected. Method 3 and method 4 perform direction forecasting and trading simulation based on the movement level prediction by an MCSVM-based and an MCSVM-GA based method, respectively. Method 3 is used to determine whether it is necessary to use multi-objective optimization by NSGA-II, while method 4 employs the GA instead of the NSGA-II for comparing their direction prediction and trading performances. Methods 7 and 8 are two well-known benchmarks of passive trading strategies. We utilize their experimental results to compare with the proposed approach for examining whether the proposed approach can outperform

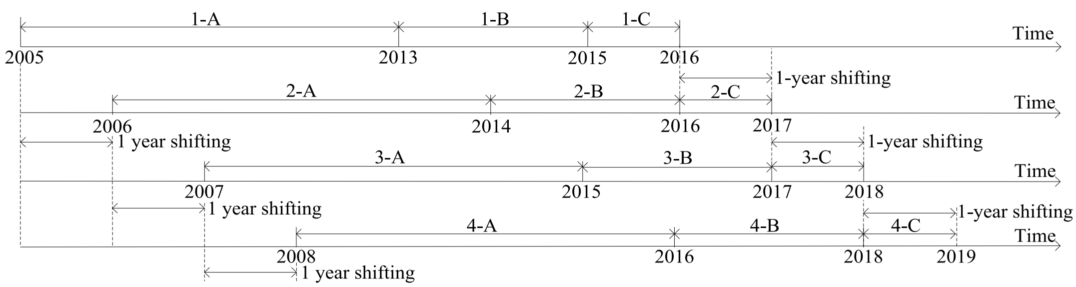


FIGURE 3. A rolling window method for data separation of training and testing periods of MCSVM and NSGA-II.

TABLE 2. Training and testing periods of four datasets for experiments.

Dataset	Period A (eight years) MCSVM + NSGA-II Training period	Period B (two years) MCSVM+NSGA-II Testing period	Periods A+ B (ten years) MCSVM Training period	Period C (one year) Prediction and Trading period
Dataset 1	2005/01-2012/12	2013/01-2014/12	2005/01-2014/12	2015/01-2015/12
Dataset 1	2006/01-2013/12	2014/01-2015/12	2006/01-2015/12	2016/01-2016/12
Dataset 2	2007/01-2014/12	2015/01-2016/12	2007/01-2016/12	2017/01-2017/12
Dataset 3	2008/01-2015/12	2016/01-2017/12	2008/01-2017/12	2018/01-2018/12

TABLE 3. A list of the proposed method and benchmark methods.

No	Method Name	Description
1	ANN	An ANN-based classification method
2	SVM	An SVM based classification method
3	MCSVM	A Multi-class SVM based classification method. No thresholds optimization
4	MCSVM-GA	A Multi-class SVM based classification method. The thresholds are optimized by the GA
5	MCSVM-NSGA-II (proposed method)	A Multi-class SVM based classification method. The thresholds are optimized by the NSGA-II
6	Buy and Hold (BAH)	Buy at the initial time and close the position at the ending time of each testing period
7	Short and Hold (SAH)	Execute a short selling transaction on the first day and close the position on the final day of each testing period

TABLE 4. The parameters optimized by the NSGA-II algorithm for the proposed method.

No	Parameter	Search Range	No	Parameter	Search Range
1	T_1	[0.04,0.08]	2	T_2	[0.02,0.04]
3	T_3	[−0.04, −0.02]	4	T_4	[−0.04, −0.08]
5	PT_1	[0.05,0.10]	6	LC_1	[−0.05, −0.03]
7	PT_2	[0.05, 0.10]	8	LC_2	[−0.05, −0.03]
9	PT_3	[0.05,0.10]	10	LC_3	[−0.05, −0.03]
11	PT_4	[0.05,0.10]	12	LC_4	[−0.05, −0.03]

the two well-known passive trading strategies. Table 4 shows the parameters optimized by the NSGA-II algorithm for the proposed approach. In order to keep consistency, the search ranges of the parameters optimized by the MCSVM-GA are identical with the proposed approach MCSVM-NSGA-II. Since there is no parameter optimization function for the MCSVM, we used the values 0.04, 0.02, −0.02, and −0.04 for the parameters T_1, T_2, T_3, and T_4, respectively. The profit-taking and loss-cutting thresholds of each movement level were set to be the fixed values 0.07 and −0.04, respectively. We also employed the SVM and ANN-based methods, which are traditional approaches that they predict the direction and execute transactions on every trading days. Since there are no movement levels of predictions for them, they will execute a long transaction if the prediction output belongs to a positive movement. In contrast, if the prediction output is negative movement, the ANN and SVM based methods will execute a short-selling transaction.

C. EVALUATION MEASURES

To measure the performance of the proposed methods for price movement direction prediction and trading simulation, we employed three evaluation methods, including the hit ratio, accumulated return, and Sharpe ratio. Among them, hit ratio is generally used to measure the correctness rate for price movement direction forecasting. Suppose that in one testing period, there were in total N times of direction forecasting, in which the right times of "upward" predictions and "downward" forecasting were *CR* and *CF*, respectively. Since in our magnitude level design for MCSVM, there are no direction prediction and trading if the movement prediction level is SP or SN, the proposed approach would not forecast the direction every day. Thus, the calculation of the hit ratio for the proposed method is as follows:

$$HR = \frac{CR + CF}{N} \tag{3}$$

It should be noted that even if the price movement direction predictions have a large hit ratio, we might fail to make a positive yield if we do not produce more positive returns than the absolute value of the losses when the direction prediction is incorrect. Therefore, one appropriate measure of the correctness of the predictions should be the profit. Suppose a method forecasts that the crude oil price will go up when the predicted level is LP or MP, then a "long" transaction is executed. In contrast, if the method forecasts the crude oil price to go down when the predicted level is LN or MN, then a "short-selling" transaction would be taken up. The holding period of the transactions is five business days; that is, the position will be closed after five business days unless the return reaches the profit-taking or loss-cutting threshold of the predicted level. Then, after one testing period lasts for one, the Accumulated Return (AR) of that trading year can be calculated by:

$$AR = \sum\nolimits_{i=1}^{b} \frac{1}{c} \frac{(C_{i+a} - P_i)}{P_i} \times \operatorname{sgn}((P_{i+a} - P_i)(C_{i+a} - P_i))$$
$$\operatorname{sgn}(x) = \begin{cases} 0, & x = 0 \\ 1, & x > 0 \\ -1, & x < 0 \end{cases} \tag{4}$$

where a refers to the steps of a-day ahead prediction (5-day ahead prediction), b denotes the number of days for one testing period that will last for one year), while c is the position percentage for every transaction. In the present study, since a trading position may last for at most five business days, to ensure that we can execute the transaction when there is a long or short trading signal, c is set to be 20%, so the ratio of the base amount relative to the initial capital was set to 20%. Additionally, transaction cost was not taken into account for the calculation of accumulated returns, since unlike the trading in security markets or currency markets, we supposed that governments or institutional investors could make transactions in crude oil market without brokers.

Furthermore, high returns are in general accompanied by the potential for high risk. Thus, other than the hit ratio or trading return, trading risk also plays an essential role in gauging the performance of the investment portfolio. Among the performance evaluation criteria, the risk-adjusted return is an index that can take both investment yields and trading risks into consideration. Among the evaluation criteria to evaluate the risk-adjusted return, Sharpe ratio is one of the well-known and classic indicators that generally utilized for measuring the return-risk ratio, named after William F. Sharpe [44]. It is defined as follows:

$$SR = \frac{E[R_{asset} - R_{free}]}{\sigma_{asset}} = \frac{E[R_{asset} - R_{free}]}{\sqrt{var[R_{aseet}]}} \tag{5}$$

where R_{asset} denotes the investment return of the asset, R_{free} denotes the return of a risk-free investment, $E\left[R_{asset} - R_{free}\right]$ refers to the expected value of the asset excess return, and σ is the standard deviation of the asset returns. The following Table 5 shows a list of evaluation measures employed in this research.

V. EXPERIMENTAL RESULTS

A. HIT RATIO RESULTS

To measure the performance of crude oil price direction forecasting for the proposed method MCSVM-NSGA-II, the ANN, SVM, MCSVM and MCSVM-GA based methods were chosen as the benchmarks, and their experiments were conducted on the crude oil price of Brent and WTI. The following Table 6 and Table 7 report the hit ratio results of Brent and WTI, respectively, over the testing periods from 2015 to 2018 for crude oil price direction forecasting.

Generally, a higher hit ratio result indicates a superior direction prediction accuracy. From the hit ratio results of the proposed approach and benchmark methods for

TABLE 5. A list of the evaluation measures for the proposed method and benchmark methods.

Evaluation Measure	Calculation Formula	Description
Hit Ratio (HR)	$HR = \frac{CR+CF}{N}$	N denotes the total number of trading times; CR is the number of right forecasting for the upward movements, while CF is the number of right forecasting for the downward movements
Accumulated Return (AR)	$AR = \sum_{i=1}^{b} \frac{1}{c} \frac{(C_{i+a} - P_i)}{P_i} \times \mathrm{sgn}((P_{i+a} - P_i)(C_{i+a} - P_i))$	a is the length of time horizon for direction forecasting; b is the number of forecasting and trading times; c denotes the trading amount percentage of each transaction
Sharpe Ratio (SR)	$SR = \frac{E[R_{asset} - R_{free}]}{\sigma_{asset}} = \frac{E[R_{asset} - R_{free}]}{\sqrt{\mathrm{var}[R_{asset}]}}$	R_{asset} refers to the asset return; R_{free} is the return of a risk-free asset; E[R_{asset}-R_{free}] is the expected value of the excess return of the asset return over the risk-free benchmark return; σ_{asset} denotes the variances of the yearly asset returns

TABLE 6. Hit ratio results for crude oil price movement direction forecasting in WTI market.

Method	ANN(%)	SVM(%)	MCSVM(%)	MCSVM-GA(%)	MCSVM-NSGA-II(%)
2015	41.96	45.88	40	**52.17**	49.52
2016	**60.78**	51.76	46.98	58.57	50.78
2017	45.66	55.46	35.71	50	**87.17**
2018	46.06	51.18	32.32	53.93	**61.17**
Average	48.62	51.07	38.75	53.67	**62.16**

TABLE 7. Hit ratio results for crude oil price movement direction forecasting in Brent market.

Method	ANN(%)	SVM(%)	MCSVM(%)	MCSVM-GA(%)	MCSVM-NSGA-II(%)
2015	47.84	48.62	43.63	**51.85**	47.79
2016	55.68	46.27	**60.71**	60.08	60
2017	58.98	60.15	50	51.13	**70.27**
2018	47.24	52.75	57.14	45.27	**61.78**
Average	52.44	51.95	52.87	52.08	**59.96**

WTI market shown in Table 6, we can observe that the average hit ratio over four testing years produced by the ANN, SVM, MCSVM and MCSVM-GA were 48.62%, 51.07%, 38.75%, and 53.67%, respectively, while our proposed method MCSVM-NSGA-II produced the highest average hit ratio of 62.16%, which was significantly better than other benchmark methods; Within the four years from 2015 to 2018, the proposed method performed best in two years (2017 and 2018). Although the ANN and MCSVM-GA performed best in 2015 and 2016, respectively, they obtained smaller hit ratio results in other years. In addition, we find the SVM produced better results than MCSVM, possibly because of the pre-set parameters of MCSVM were not suitable for price movement level determination. Therefore, the parameters without optimization might cause a smaller hit ratio result. In addition, the MCSVM-GA outperformed the SVM and MCSVM, which indicates that parameter optimization using the GA was beneficial for producing a better hit ratio result. Moreover, it is observed by us that the average hit ratio result of the proposed approach MCSVM-NSGA-II was better than MCSVM-GA, demonstrating that the multi-objective optimization in the proposed method successfully enhanced the direction forecasting accuracy of the price movements. Additionally, for the crude oil price direction prediction in Brent market that shown in Table 7, we find that the average hit ratio result of ANN, SVM, MCSVM, and MCSVM-GA based methods were 52.44%, 51.95%, 52.87%, and 52.08%, respectively, while the best result was produced by the proposed method MCSVM-NSGA-II that it obtained average hit ratio results of 59.96%. In 2017 and 2018, the proposed method produced the best hit ratio results. In spite that the MCSVM-GA and MCSVM obtained the best hit ratio result in 2015 and 2016, respectively, they obtained relatively small hit ratio result in other years (for instance, 43.63% in 2015 for the MCSVM) thus their average performance is worse than the proposed method. We also observe that the average hit ratio performance of ANN and SVM were 51.44% and 51.95%, respectively, which indicates that ANN performed better than SVM in terms of average hit ratio results. This finding is consistent with the conclusion of related literature that ANN can sometimes be superior to SVM [45], [46].

In addition, unlike the results of the MCSVM and MCSVM-GA in the WTI market, they produced very close average hit ratio results in the Brent market. This may have been because of the fitness function of the MCSVM-GA was the accumulated return in the training period, while a larger return would not always lead to a larger hit ratio result, and vice versa. Hence, in our proposed method, we set both the hit ratio and accumulated return as the objective values, its better average hit ratio results than MCSVM-GA as well as other benchmark methods demonstrate that the multi-objective optimization and multi-class classification successfully improved the hit ratio results. In summary, the hit ratio results in Brent and WTI crude oil markets indicate that the proposed method was extremely better than benchmark methods in terms of the direction prediction accuracy.

B. ACCUMULATED RETURN RESULTS

Other than the price direction prediction, investors generally would also consider the market return as an extremely essential evaluation criterion of the approach. Table 8 and Table 9 show the accumulated returns of the proposed approach and benchmark methods over 2015 to 2018 for simulation trading in the Brent and WTI crude oil market, respectively. Note that values of accumulated return were measured relative to the initial investment at the beginning of each year. Additionally, the accumulated return results of the passive trading strategies BAH and SAH were provided and compared with the proposed approach.

In general, a larger accumulate return indicates a superior profit-making ability. For WTI results shown in Table 8, we can observe that the average accumulated return of four testing years produced by the ANN, SVM, MCSVM,

TABLE 8. Yearly accumulated returns (in percentage) for trading simulation in WTI crude oil market.

Method	ANN (%)	SVM (%)	MCSVM (%)	MCSVM-GA (%)	BAH (%)	SAH (%)	MCSVM-GA-II (%)
2015	−24.39	−14.71	7.83	36.48	−29.57	29.57	**52.85**
2016	**58.53**	22.36	18.45	17.15	46.02	−46.02	40.54
2017	−20.32	**20.19**	−1.85	−5.70	15.46	−15.46	9.88
2018	−21.18	11.72	−6.56	8.90	−25.21	**25.21**	8.43
Average	−1.84	9.89	4.47	14.21	1.68	−1.68	**27.92**

TABLE 9. Yearly accumulated returns (in percentage) for trading simulation in Brent crude oil market.

Method	ANN (%)	SVM (%)	MCSVM (%)	MCSVM-GA (%)	BAH (%)	SAH (%)	MCSVM-GA-II (%)
2015	−22.60	−9.63	3.78	15.36	−33.89	**33.89**	29.48
2016	**53.12**	−29.43	−13.99	−7.55	51.48	−51.48	4.66
2017	**22.39**	19.57	−0.01	3.40	21.21	−21.21	19.40
2018	−16.43	12.23	2.94	23.44	−24.12	**24.12**	21.26
Average	9.12	−1.82	−1.82	8.66	3.67	−3.67	**18.7**

MCSVM-GA, BAH, and SAH were −1.84%, 9.89%, 4.47%, 14.21%, 1.68%, and −1.68%, respectively, while our proposed method MCSVM-NSGA-II produced the largest average accumulated return of 27.92%. In spite that the ANN, SVM, and SAH yielded the best accumulated return in 2016, 2017 and 2018 respectively, they incurred from huge losses in other years, which indicates that these models were not sufficiently powerful for generating stable positive returns in crude oil markets. For instance, the ANN produced negative returns in 2015, 2016 and 2018; SVM generated a loss of 14.71% in 2015; The SAH trading strategy obtained negative yields during 2016-2017. Among all the methods, only the proposed method yielded all positive returns in all four trading years, and it performed the best in terms of average return.

Moreover, we find that the average return of SVM (9.89%) was superior to that of ANN (−1.84%), which agrees with the average hit ratio results shown in Table 6 that the average hit ratio of SVM was better than that of ANN. The average return of MCSVM (4.47%) was smaller than that of SVM (9.89%), which reveals that the preset multi-class classification of the price movements and profit-taking/loss-cutting failed to improve the trading returns. The MCSVM-GA yielded a superior average return than SVM and MCSVM, indicating that the parameters optimization by using GA successfully improved the average trading return. However, MCSVM-GA was incapable of producing positive returns in all four years that it suffered a negative return of −5.7% in 2017. The proposed method MCSVM-NSGA-II obtained the best average return, and it consistently outperformed the MCSVM-GA in all years except 2018, demonstrating that the multi-objective optimization enhanced the trading returns and it could be adopted as an efficient trading approach for trading in WTI crude oil market.

Next, we focus on the trading performance in Brent crude oil market. From the experimental results reported in Table 9, we find that the average hit ratio of the ANN, SVM, MCSVM, MCSVM-GA, BAH, and SAH were 9.12%, −1.82%, −1.82%, 8.66%, 3.67%, and −3.67%, respectively. Similar as for the WTI crude oil market, the best return was also produced by the proposed method MCSVM-NSGA-II that it obtained an average accumulated return of 18.7%. Among all the methods, only the proposed approach consistently yielded positive returns in all four testing years. It demonstrates that the proposed method was extremely better than the benchmark methods at producing positive returns.

In addition, it is found by us that the ANN (9.12%) produced better average return than SVM (−1.84%), which is consistent with the hit ratio results reported in Table 7 that ANN was superior to SVM in terms of average hit ratio results. However, we observe that the MCSVM (52.87%) obtained the better average hit ratio result than ANN (52.44%) and SVM (51.95%) in the Brent market, its average return was −1.82%, which reveals that it suffered more losses in the incorrect direction predictions while it gained less in total in the correct ones. It also provides evidence that results of hit ratio and accumulates return sometimes were conflicting with each other. The passive trading strategy SAH produced the best accumulated return in 2015 (33.89%) and 2018 (24.12%), whereas it incurred vast losses in 2016 (−51.48%) and 2017 (−21.21%), which indicates that it did not yield satisfying returns stably. Furthermore, the MCSVM-GA is found to generate a higher average return than the MCSVM, indicating that the parameter optimization of the trading rules and movement levels enhanced the average trading returns. Finally, the average return of the proposed approach MCSVM-NSGA-II was superior to that of MCSVM-GA, indicating that the multi-objective optimization substantially benefited the accumulated return performance.

C. SHARPE RATIO RESULTS

Besides the direction forecasting accuracy and investment return, generally, the market participants would take into account the risk of their trading. In the present study, the Sharpe ratio is employed to gauge the goodness of return/risk ratio of the proposed method and benchmark methods. In general, a larger Sharpe ratio value indicates a better return/risk ratio performance. Risk-free interest rate is often used to calculate the Sharpe ratio value. In this study, a one-year treasury bond is considered a risk-free investment. Thus, in our experiment, its yields over 2015 to 2018 that provided by the website of the U.S. department of the Treasury [47] were utilized. Subsequently, for convenience, the average interest rate of the one-year treasury bond of these four years (about 0.89%) was employed as the risk-free yield to calculate the Sharpe ratio results of the benchmark methods and the proposed method. The Sharpe ratio results of the proposed method MCSVM-NSGA-II and the benchmark methods are reported in Table 10.

From the Sharpe ratio results reported in Table 10, we can find that the proposed approach produced the highest Sharpe

TABLE 10. The Sharpe ratio values of the proposed method and benchmark methods. The values in brackets next to the Sharpe ratio value means the (average excess return/standard deviation of the returns).

Method	ANN	SVM	MCSVM	MCSVM-GA	BAH	SAH	MCSVM-NSGA-II
WTI	−0.068 (−0.0273/ 0.4028)	0.528 (0.0900/ 0.1703)	0.341 (0.0378/ 0.1108)	0.757 (0.1332/ 0.1760)	0.022 (0.0078/ 0.3585)	−0.071 (−0.0256/ 0.3585)	**1.214** (0.2703/ 0.2226)
Brent	0.232 (0.0823/ 0.3545)	−0.122 (−0.0271/ 0.222)	−0.328 (−0.0271/ 0.0827)	0.572 (0.0777/ 0.1359)	0.070 (0.0278/ 0.3990)	−0.114 (0.0456/ 0.3990)	**1.724** (0.1781/ 0.1033)

ratio values in both Brent and WTI markets (1.214 for WTI market and 1.724 for Brent market), which were significantly superior to that of the method MCSVM-GA performed best in the benchmarks (0.757 for WTI market and 0.572 for Brent market). Additionally, the Sharpe ratio values of the proposed approach were markedly better than that of SVM and MCSVM based methods, which indicates that employment of NSGA-II for threshold value optimization successfully enhanced the Sharpe ratio results.

In addition, it is observed that the benchmark methods ANN, SVM and MCSVM failed to produce positive Sharpe ratio results in two crude oil market. The MCSVM-GA yielded relatively good excess return results (13.32% in WTI market and 7.77% in Brent market) obtained worse Sharpe ratio results than the proposed method, in essence, because the proposed method yielded a highly better excess return than MCSVM-GA in both WTI and Brent crude oil markets. Furthermore, although the Sharpe ratio values of the most famous passive trading strategy BAH were positive (0.022 for WTI and 0.070 for Brent), their values were quite smaller than the Sharpe ratio values of the proposed approach. In summary, the proposed approach MCSVM-NSGA-II performed better than the benchmark methods in terms of not only hit ratio and yearly accumulated return, but also the return/risk ratio.

VI. CONCLUSION

In this paper, we developed a novel method integrates MCSVM and NSGA-II for price direction prediction and trading simulation in two international crude oil markets, where MCSVM was adopted to build a multi-class prediction model, and the NSGA-II was applied to generating an optimal trading rule by maximizing multiple objective functions. The proposed approach mainly consists of four parts. First, the DP part is used to derive the original data of crude oil daily spot prices and conduct the data pre-processing on them; Next, the MCMO part is used to optimize the level parameters and profit-taking/loss cutting thresholds of the trading rule; Then, the direction forecasting and simulation trading are conducted in the DPT part. Finally, we evaluate the forecasting and trading performance in the PE part. Experimental results showed that the proposed approach produced hit ratios of 62.16% and 59.96% for WTI and Brent, respectively. The average annual returns for WTI and Brent were 27.92% and 18.70%, respectively. Additionally, the average hit ratio and average return results of the proposed approach over a four-year testing period were consistently the best among all of the investigated methods. Furthermore, the proposed approach not only performed well in terms of hit ratio and trading return but also in Sharpe ratio results which were 1.214 for WTI and 1.724 for Brent. In short, the proposed method consistently produced favorable hit ratios and annual accumulated returns with low volatility over a four-year testing period. The experimental results demonstrate that the proposed method could be employed as an alternative approach for direction prediction and actual trading.

There are several research directions to be expanded in the future. For example, hit ratio and accumulated return were employed as fitness functions in our proposed method, while other researchers might examine different objective functions for optimization. In addition, researchers might design more magnitude categories of positive and negative movement levels. Moreover, in this paper, the proposed method has been examined in two crude oil markets, an investigation of its performance in other financial markets such as security markets or futures markets will be a future direction of our research.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their comments and discussions.

REFERENCES

[1] W. Xu, J. Wang, and J. Ma, "Forecasting crude oil demand using a hybrid SVR and Markov approach," in *Business Intelligence in Economic Forecasting Technologies & Techniques*, J. Wang and S. Wang, Eds. Hershey, PA, USA: IGI Global, 2010, pp. 235–247.

[2] A. Charles and O. Darne, "The efficiency of the crude oil markets: Evidence from variance ratio tests," *Energy Policy*, vol. 37, no. 11, pp. 4267–4272, Nov. 2009.

[3] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1994.

[4] S. Moshiri and F. Foroutan, "Forecasting nonlinear crude oil futures prices," *Energy J.*, vol. 27, no. 4, pp. 81–95, 2006.

[5] H. Mohammadi and L. Su, "International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models," *Energy Econ.*, vol. 32, no. 5, pp. 1001–1008, Sep. 2010.

[6] H. Chiroma, S. Abdul-Kareem, A. S. M. Noor, A. I. Abubakar, N. S. Safa, and L. Shuib, "A review on artificial intelligence methodologies for the forecasting of crude oil price," *Intell. Autom. Soft Comput.*, vol. 22, no. 3, pp. 449–462, Jan. 2016.

[7] J. Li, Z. Xu, H. Xu, L. Tang, and L. Yu, "Forecasting oil price trends with sentiment of online news articles," *Asia Pacific J. Oper. Res.*, vol. 34, no. 2, 2017, Art. no. 1740019.

[8] L. Yu, S. Wang, and K. K. Lai, "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm," *Energy Econ.*, vol. 30, no. 5, pp. 2623–2635, Sep. 2008.

[9] Q. Wu, G. Hao, and X. Cheng, "Crude oil price forecasting with an improved model based on wavelet transform and RBF neural network," in *Proc. Int. Forum Inf. Technol. Appl. (IFITA)*, Chengdu, China, May 2009.

[10] R. Jammazi and C. Aloui, "Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling," *Energy Econ.*, vol. 34, no. 3, pp. 828–841, May 2012.

[11] L. Zhang, F. Wang, B. Xu, W. Chi, Q. Wang, and T. Sun, "Prediction of stock prices based on LM-BP neural network and the estimation of overfitting point by RDCI," *Neural Comput. Appl.*, vol. 30, no. 5, pp. 1425–1444, Sep. 2018.

[12] V.N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag, 1995.

[13] M. CiMen and O. KiSi, "Comparison of two different data-driven techniques in modeling lake level fluctuations in Turkey," *J. Hydrol.*, vol. 378, nos. 3–4, pp. 253–262, Nov. 2009.

[14] S. Deng, Y. Xiang, and B. Nan, "A hybrid method for crude oil price direction forecasting using multiple timeframes dynamic time wrapping and genetic algorithm," *Soft Comput.*, to be published, doi: 10.1007/s00500-019-04304-9.

[15] A. Qureshi, N. Thakur, and M. Kumar, "VIRsiRNApred: A Web server for predicting inhibition efficacy of siRNAs targeting human viruses," *J. Transl. Med.*, vol. 11, no. 305, p. 305, Dec. 2013.

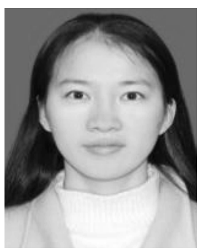
[16] Y. L. Qi and W. J. Zhang, "The improved SVM method for forecasting the fluctuation of international crude oil price," in *Proc. Int. Conf. Electron. Commerce Bus. Intell.*, Beijing, China, Jun. 2009.

[17] M. Yasir, M. Y. Durrani, S. Afzal, M. Maqsood, F. Aadil, I. Mehmood, and S. Rho, "An intelligent event-sentiment-based daily foreign exchange rate forecasting system," *Appl. Sci.*, vol. 9, no. 15, p. 2980, Jul. 2019.
[18] A. Khashman and N. I. Nwulu, "Intelligent prediction of crude oil price using support vector machines," in *Proc. IEEE Int. Symp. Appl. Mach. Intell. Inform. (SAMI)*, Smolenice, Slovakia, Jan. 2011, pp. 165–169.
[19] C. N. Phyo, T. T. Zin, and P. Tin, "Complex human-object interactions analyzer using a DCNN and SVM hybrid approach," *Appl. Sci.*, vol. 9, no. 9, p. 1869, May 2019.
[20] S. G. A. Usha and S. Vasuki, "A novel method for segmentation and change detection of satellite images using proximal splitting algorithm and multiclass SVM," *J. Indian Soc. Remote*, vol. 47, no. 5, pp. 853–865, May 2019.
[21] T. Bao, B. N. Klatt, S. L. Whitney, K. H. Sienko, and J. Wiens, "Automatically evaluating balance: A machine learning approach," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 2, pp. 179–186, Feb. 2019.
[22] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA, USA: Addison-Wesley, 1989.
[23] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, Perth, WA, Australia, vol. 4, Nov./Dec. 1995, pp. 1942–1948, doi: 10.1109/ICNN.1995.488968.
[24] A. Hirabayashi, C. D. C. Aranha, and H. Iba, "Optimization of the trading rule in foreign exchange using genetic algorithm," in *Proc. Conf. Genetic Evol. Comput.*, Montreal, QC, Canada, Jul. 2009, pp. 1529–1536.
[25] S. Deng, K. Yoshiyama, T. Mitsubuchi, and A. Sakurai, "Hybrid method of multiple kernel learning and genetic algorithm for forecasting short-term foreign exchange rates," *Comput. Econ.*, vol. 45, no. 1, pp. 49–89, Jan. 2015.
[26] C. F. Huang, C. J. Hsu, C. C. Chen, B. Chang, and C. A. Li, "An intelligent model for pairs trading using genetic algorithms," *Comput. Intell. Neurosci.*, pp. 1–10, Mar. 2015, doi: 10.1155/2015/939606.
[27] S. Deng, Y. Xiang, Z. Fu, M. Wang, and Y. Wang, "A hybrid method for crude oil price direction forecasting using multiple timeframes dynamic time wrapping and genetic algorithm," *Appl. Soft. Comput.*, vol. 82, Sep. 2019, Art. no. 105566.
[28] T. Ware, "Adaptive statistical evaluation tools for equity ranking models," *Can. Appl. Math. Q*, vol. 13, no. 3, pp. 265–275, 2005.
[29] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
[30] S. Vine, *Options: Trading Strategy and Risk Management*. Hoboken, NJ, USA: Wiley, 2011.
[31] N. F. Alkayem, B. Parida, and S. Pal, "Optimization of friction stir welding process using NSGA-II and DEMO," *Neural Comput. Appl.*, vol. 31, pp. 947–956, Feb. 2019.
[32] B. Huang, B. Buckley, and T.-M. Kechadi, "Multi-objective feature selection by using NSGA-II for customer churn prediction in telecommunications," *Expert Syst. Appl.*, vol. 37, no. 5, pp. 3638–3646, May 2010.
[33] C. H. Park and M. Lee, "A SVM-based discretization method with application to associative classification," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 4784–4787, Apr. 2009.
[34] S. Zidi, T. Moulahi, and B. Alaya, "Fault detection in wireless sensor networks through SVM classifier," *IEEE Sensors J.*, vol. 18, no. 1, pp. 340–347, Jan. 2018.
[35] B. Mcdermott, M. O'Halloran, E. Porter, and A. Santorelli, "Brain haemorrhage detection using a SVM classifier with electrical impedance tomography measurement frames," *PLoS ONE*, vol. 13, no. 7, Jul. 2018.
[36] S. Deng, T. Mitsubuchi, and A. Sakurai, "Stock price change rate prediction by utilizing social network activities," *Sci. World J.*, Mar. 2014, doi: 10.1155/2014/861641.
[37] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 415–425, Mar. 2001.
[38] U. Krebel, "Pairwise classification and support vector machines," in *Advances in Kernel Methods-Support Vector Learning*. Cambridge, U.K.: MIT Press, Jan. 1999, pp. 255–268.
[39] Y. Makaremi, A. Haghighi, and H. R. Ghafouri, "Optimization of pump scheduling program in water supply systems using a self-adaptive NSGA-II; a review of theory to real application," *Water Resour. Manag.*, vol. 31, no. 4, pp. 1283–1304, Mar. 2017.
[40] M. Rabbani, H. Farrokhi-Asl, and B. Asgarian, "Solving a bi-objective location routing problem by a NSGA-II combined with clustering approach: Application in waste collection problem," *J. Ind. Eng. Int.*, vol. 13, no. 1, pp. 13–27, Mar. 2017.
[41] Z. Masoumi, J. Maleki, M. S. Mesgari, and A. Mansourian, "Using an evolutionary algorithm in multiobjective geographic analysis for land use allocation and decision supporting," *Geogr. Anal.*, vol. 49, no. 1, pp. 58–83, Jan. 2017.
[42] *Energy Information Administration*. Accessed: Oct. 2019. [Online]. Available: https://www.eia.gov/
[43] M. Thomason, "The practitioner methods and tool," *J. Comput. Int. Finance*, vol. 7, pp. 36–45, 1999.
[44] W. F. Sharpe, "The Sharpe ratio," *J. Portfolio Manage.*, vol. 21, no. 1, pp. 49–58, 1994.
[45] I. Belakhdar, W. Kaaniche, R. Djmel, and B. Ouni, "A comparison between ANN and SVM classifier for drowsiness detection based on single EEG channel," in *Proc. Int. Conf. Adv. Technol. Signal Image Process. (ATSIP)*, Monastir, Tunisia, Mar. 2016.
[46] D. N. Cosenza, H. G. Leite, G. E. Marcatti, D. H. B. Binoti, A. E. M. D. Alcantara, and R. Rode, "Site classification with support vector machine and artificial neural network," *Forest Sci.*, vol. 43, no. 108, pp. 955–963, 2015.
[47] *U.S. Department of the Treasury*. Accessed: Oct. 6, 2019. [Online]. Available: https://www.treasury.gov

SHANGKUN DENG received the Ph.D. degree from Keio University, Japan, in 2015.

He worked with Hitachi Corporation (Tokyo), from 2015 to 2016. He is currently a Researcher with the College of Economics and Management, China Three Gorges University, Yichang, China. His research interests include movement analysis in financial markets, machine learning applications, and financial market regulation.

XIAORU HUANG received the B.S. degree in information management and information system from the Xi'an University of Finance and Economics (XUFE), Xi'an, China, in 2018. She is currently pursuing the master's degree in applied economics with China Three Gorges University (CTGU), Yichang, China. Her research interests include price fluctuations and forecasts in financial markets.



JIASHUANG SHEN is currently pursuing the bachelor's degree with the College of Economics and Management, China Three Gorges University, Yichang, China.

Her research interests include energy price prediction and machine learning.

HAORAN YU is currently pursuing the bachelor's degree with the College of Economics and Management, China Three Gorges University, Yichang, China.

His research interests include energy price prediction and machine learning.

CHENGUANG WANG is currently pursuing the bachelor's degree with the College of Economics and Management, China Three Gorges University, Yichang, China.

His research interests include machine learning algorithms and stock price prediction.

HONGYU TIAN received the Ph.D. degree from the College of Economics and Management, Southwest University, Chongqing, China, in 2016.

He is currently a Researcher with the College of Economics and Management, China Three Georges University, Yichang, China. His research interests include time series prediction and agriculture products price analysis.

FANGJIE MA received the bachelor's degree from China Three Gorges University, Yichang, China, in 2017.

She is currently a Teaching Assistant with the College of Science and Technology, China Three Gorges University. Her research interests include big data analysis, machine learning, and financial market prediction.

TIANXIANG YANG was born in Wenzhou, China in October 2, 1988. He received the M.E. degree in global information and telecommunication studies from Waseda University, Tokyo, Japan, in 2015, where he is currently pursuing the Ph.D. degree with the Department of Industrial and Management Systems Engineering.

He is also a Research Associate with the Department of Industrial and Management Systems Engineering, Waseda University. His research interest includes purchase history data analysis based on machine learning algorithms.