This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3003819, IEEE Access

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Prediction-based portfolio optimization models using deep neural networks

YILIN MA, RUIZHU HAN, AND WEIZHONG WANG
School of Economics and Management, Southeast University, Nanjing 211189, China

Corresponding author: Ruizhu Han (e-mail: daodao-777@163.com)

This work was supported by the Chinese Natural Science Foundation under Grant 71390335.

ABSTRACT Portfolio optimization is a hot research topic, which has attracted many researchers in recent decades. Better portfolio optimization model can help investors earn more stable profits. This paper uses three deep neural networks (DNNs), i.e., deep multilayer perceptron (DMLP), long short memory (LSTM) neural network and convolutional neural network (CNN) to build prediction-based portfolio optimization models which own the advantages of both deep learning technology and modern portfolio theory. These models first use DNNs to predict each stock's future return. Then, predictive errors of DNNs are applied to measure the risk of each stock. Next, the portfolio optimization models are built by integrating the predictive returns and semi-absolute deviation of predictive errors. These models are compared with three equal weighted portfolios, where their stocks are selected by DMLP, LSTM neural network and CNN respectively. Also, two prediction-based portfolio models built with support vector regression are used as benchmarks. This paper applies component stocks of China securities 100 index in Chinese stock market as experimental data. Experimental results present that the prediction-based portfolio model based on DMLP performs the best among these models under different desired portfolio returns, and high desired portfolio return can further improve the performance of this model. This paper presents the promising performance of DNNs in building prediction-based portfolio models.

INDEX TERMS Deep neural network, Prediction-based portfolio, Semi-absolute deviation.

I. INTRODUCTION

PORTFOLIO optimization as a challenge and multi-objective optimization problem has received increasing attention from researchers, fund managers and individual investors. The main idea of portfolio optimization is to determine the optimal weight of each asset by maximizing its expected return and minimizing the risk simultaneously. Better portfolio optimization model owns superior efficient frontier, which can help investors obtain higher expected return at the same risk level. Thus, proposing more efficient portfolio optimization model becomes a hot topic in investment management fields.

Markowitz mean variance (MV) model as the beginning of modern portfolio theory first presents an efficient formula solution to trade-off between expected return maximization and risk minimization [1]. Since the MV model is based on many restrict hypotheses such as normal distribution of stocks' returns, which is hardly established in real stock market, many researchers try to improve the suitability of this model from different perspectives. Mean absolute deviation model is introduced by Konno et. al to replace MV model [2]. This model applies absolute deviation as risk metric, which is easy to calculate for large scale portfolio optimization problems. They also proves that absolute deviation metric is equivalent to variance metric when stocks' returns are normally distributed. Then, Speranza uses semi-absolute deviation indicator to measure the downside risk of each stock [3]. Because the downside risk of each stock is more important than the upside risk and can help investors better handle the risk of each stock and made a better choice. Since then, semi-absolute deviation metric has been widely used to measure the risk for portfolio optimization.

These classical portfolio models usually adopt the mean of historical stock returns as expected return, which is suitable for long term investment in stock market practice. But for short term investment, since stock market has many short term speculations and stock price is greatly affected by market sentiment, these models may not be suitable for short term investment. In addition, using mean historical returns as future expected returns conducts a low pass filtering influence

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3003819, IEEE Access

on the stock markets, giving imprecise predictions of future stock returns, which is adverse to the models' performance on short term investments [4].

Recently, different kinds of machine learning (ML) models have been used for short term financial market prediction [5]–[12] and have obtained satisfying results. This shows a promising direction to apply predictive future return as expected return in building portfolio models. Thus, ML techniques can be combined with modern portfolio optimization models for short term investment, which possesses the advantages of both ML techniques and modern portfolio theory [13]. In this case, some researches try to combine artificial neural networks (ANNs) with portfolio optimization models to build novel portfolio models and generate satisfying results [14]–[16]. These models utilize historical returns of individual stocks to calculate portfolio's risk. However, many empirical studies have shown that historical returns hardly obey the normal distribution hypothesis. Thus, it is not reasonable to use historical returns to calculate portfolio's risk. Fortunately, Freitas et al. discover that the normality of ANN's predictive errors is higher than the series of historical returns [17], which means predictive errors of ANN are more suitable for building MV model. Then, some researches propose different prediction-based portfolio models by using predictive errors of ML models rather than historical returns of individual stocks for portfolio formulation and obtain promising results [4], [18], [19]. However, this kind of research has received very little attention from researchers since then [20]. Thus, it is necessary and worthy to further pay attention to this research direction. As deep neural networks (DNNs) have shown better performance than traditional ML technologies in financial market prediction [21]–[25], it is of great importance to explore the application of DNNs in formulating prediction-based portfolio optimization models. Therefore, this study focuses on building prediction-based portfolio optimization models by using different DNNs' predictive results.

Among all the DNNs, deep multilayer perceptron (DMLP), long short memory (LSTM) neural network, and convolutional neural network (CNN) are frequently applied in stock market prediction, a detailed review refers to [26]. Thus, the objective of this paper is to further improve the out-of-sample performance of prediction-based portfolio optimization models by building these models with DMLP, LSTM neural network and CNN. Actually, this paper first applies DMLP, LSTM neural network and CNN for future stock return prediction and calculates the portfolio's expected return by linear combination of each stock's predictive return. Then, semi-absolute deviation metric is used to measure the risk of each stock based on their predictive errors. Finally, prediction-based portfolio optimization models are built by generalizing the frame of mean semi-absolute deviation (MSAD) portfolio model. In order to present the benefit of these models, this paper chooses three equal weighted (EW) portfolios as comparisons, where their stocks are selected by DMLP, LSTM neural network and CNN respectively, Also, two prediction-based portfolio models based on support vector regression (SVR) are applied as benchmarks, which use SVR instead of DNNs for stock prediction.

In this regard, this paper has three contributions to fill the research gap in existing literature. First, this paper investigates the performance of three frequently used DNNs, i.e., DMLP, LSTM neural network and CNN, in formulating prediction-based portfolio optimization models. As far as we know, this is the first paper to apply DNNs to build prediction-based portfolio optimization models, which extends the existing researches. Second, this paper applies semi-absolute deviation as the risk indicator to replace variance in building prediction-based portfolio models which do not need the normal distribution hypothesis and only focus on the downside risk of predictive errors. Thus, these models become more efficient for large scale portfolio optimization problems and more practical in real stock market investment. Third, this paper compares these prediction-based portfolio models with three equal weighted portfolio models in order to show their advantages. Also, two prediction-based portfolio models based on SVR are used as benchmarks. In addition, this paper uses China Securities 100 Index component stocks as experimental data. Also, this paper focuses on the historical data from 2007 to 2015 and uses the last four years' data to test the performance of these prediction-based portfolio models.

The remainder of this paper is showed as follows. Section 2 reviews some relative researches. Section 3 gives different prediction-based portfolio optimization models. Section 4 shows the whole experimental process. Experimental results of prediction-based portfolio models are presented in Section 5. Section 6 gives comparison of different models with transaction fee. Finally, Section 7 draws a conclusion.

II. LITERATURE REVIEW

Many portfolio selection models have been proposed during the past few decades. In the following, some recent researches related to this study are presented in three perspectives.

A. PORTFOLIOS BASED ON THE PREDICTIVE RESULTS OF MACHINE LEARNING MODELS

Lee et al. [27] first compared the performance of recurrent neural network, gated recurrent unit and LSTM neural network for stock return prediction. Experimental results showed that LSTM neural network outperformed the other models. They also proposed predictive threshold-based portfolios with the predictive results of LSTM neural network and generated satisfying performance. Krauss et al. [7] implemented and compared the performance of multilayer perceptron (MLP), gradient-boosted tree, random forest and some ensembles of these models for statistical arbitrage. Based on the predictive results of different models, portfolios were built by going long the top k stocks and going short the bottom k stocks. Experiments presented that portfolio based on a equal weighted ensemble model containing MLP,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3003819, IEEE Access

gradient-boosted tree and random forest generated return exceeding 0.45 percent per day prior to transaction fee. Fischer et al. [6] first utilized LSTM neural network, random forest, MLP and logistic regression for future stock return prediction. They found that LSTM neural network performed better than the other memory-free models. They also built a portfolio based on the predictive results of LSTM neural network by the same method in [7]. Experimental results showed that this portfolio outperformed the general market from 1992 to 2009, but deteriorated in 2010. Yang et al. [28] applied extreme learning machine to predict future stock return, then used the predictive return as a indicator to construct portfolio optimization model combining with other technical indices. Differential evolution algorithms were used to solve the portfolio optimization problem. By using the A-share market of China as experimental data, the results showed that the proposed model outperformed traditional methods, which suggested the promising effect of stock prediction for stock selection.

These models only apply ML models for stock return prediction and build portfolios simply based on the predictive results of ML models. These portfolio methods can not effectively balance return and risk because different assets own different risk. Since modern portfolio theory is proposed to solve this kind of problem, thus ML prediction model should be combined with modern portfolio optimization models for investment.

B. PORTFOLIOS BASED ON MACHINE LEARNING MODELS AND MV

Lin et al. [14] considered a dynamic portfolio selection problem, where the Elman neural network was applied to learn the dynamic stock market behavior and predict future return, and the cross-covariance matric was used to calculate the covariance matrix of stocks, then a optimal dynamic portfolio selection models was obtained. Experimental results showed that this model performed better than vector autoregression model and gave better results for dynamic portfolio optimization problem. Alizadeh et al. [29] applied an adaptive neuro-fuzzy inference system (ANFIS) for stock portfolio prediction. They showed that the performance of portfolio return prediction could be improved by using ANFIS and different input features containing technical factors and fundamental factors. Experiments presented that the proposed method outperformed the classical mean variance model, neural networks and the Sugeno-Yasukawa method. Deng et al. [30] used linear regression model containing ten variables for stock selection in US and global equities, then, they built portfolio by using MV model. Experimental results showed that the proposed model in global equity universe outperformed that of the US equity universe. Paiva et al. [15] proposed a decision making model named SVM+MV for financial trading in stock market by using support vector machine (SVM) for stock price prediction and MV model for portfolio optimization. This model first applied SVM to select better assets, then used MV model for portfolio optimization. Experimental results showed that SVM could improve the total performance of the portfolio and their decision making model owned satisfying performance in Brazilian market. Wang et al. [16] combined LSTM neural network with MV model to build portfolio. This model first used LSTM neural network to predict future moving direction of each stock, then selected the top k stock to build portfolio by using MV model. They compared their proposed model with four MV models based on three ML models and autoregressive integrated moving average model in order to show its superiority.

These studies utilize MV model to conduct the portfolio optimization based on historical returns. MV model is based on the hypotheses that the mean value of historical returns is its average value and the historical returns follow normal distribution, but the distributions of historical returns often depart from normality, exhibiting kurtosis and skewness [31], [32]. Thus, it is not rigorous to use MV model for portfolio optimization in practice.

C. PREDICTION-BASED PORTFOLIO OPTIMIZATION MODELS

Freitas et al. [18] proposed a novel portfolio optimization model, which used autoregressive neural network to predict expected returns and applied predictive errors for portfolio optimization. Experimental results showed that the proposed model outperformed the MV model and generated better return for the same risk. Freitas et al. [4] proposed a prediction-based portfolio optimization model by using autoregressive moving reference neural network (AR-MRNN) model as predictor. This paper first applied AR-MRNN model to predict future stock return and then used the variance of predictive error as risk to set up portfolio optimization model. Experimental results showed that the proposed model outperformed classical mean variance model based on the analysis of efficient frontier and real stock market performance. Hao et al. [19] presented a prediction-based portfolio selection model by using SVR for future stock return prediction and the variance of predictive errors as risk for portfolio optimization, they compared their proposed model with the model in [4]. Experimental results showed that their model performed better. Also, they mentioned that better prediction of future stock return gave better performance of their model.

These studies not only apply predictive return as expected return, but also use the predictive errors to build portfolio optimization model. Their conclusions show that this type of portfolio optimization model, i.e., prediction-based portfolio model, is promising in future stock investment. However, according to [20], the author showed that prediction-based portfolio model was an interesting research area for future research which had received very little attention from researchers. Thus, this paper tries to fill this gap by using DNNs to build prediction-based portfolio models.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3003819, IEEE Access

III. MODELS

This section shows different models used in this paper in four main parts, i.e., prediction-based portfolio optimization model based on DMLP, LSTM neural network, and CNN and SVR.

A. PORTFOLIO OPTIMIZATION MODEL BASED ON DMLP

DMLP is a classical artificial neural network which has been often used for classifications and regressions. DMLP model usually contains one input layer, multiple hidden layers and one output layer [33]. It is different from shadow networks since it consists of more hidden layers [26]. Usually more hidden layers are used to improve its learning ability. DMLP contains many hyperparameters, this paper uses grid research method to discover its optimal hyperparameters. Since relu function performs better than tanh function [34], this paper applies relu function as activation function. All the considered hyperparameters are presented in Table 1.

TABLE 1. Parameters of DMLP

Parameter	value
Hidden nodes	5,10,15,20,25,30
Hidden layers	1,2,3,...,10
Learning rate	0.0001, 0.001, 0.01, 0.1
Patient	0,5,10
Batch size	50,100,200
Activation function	relu
Loss function	Mean absolute error
Optimizer	Adam, RMSProp, AdaGrad, SGD

Stochastic gradient descent method is used to train the DMLP and earlystopping is adopted to solve overfitting problem. After many experiments, the specified topology of DMLP model is obtained. This paper sets DMLP with 10 nodes per hidden layer, 2 hidden layers, 0.01 for learning rate, 0 for patient, 100 for batch size and Adam for optimizer.

1) Prediction-based portfolio optimization model with DMLP and MSAD(DMLP+MSAD)

DMLP+MSAD first adopts DMLP to predict future stock return of each stock and then applies the semi-absolute deviation as the risk indicator to build prediction-based portfolio model. This kind of risk metric only considers the downside risk of each portfolio, which is more practical in real investment. Also absolute deviation is easier to calculate than the variance indicator which is more suitable for large scale portfolio optimization problems. In the following, this model is displayed in detail.

Let $r_{it}, \hat{r}_{it}$ denote the return and predictive return of stock i at time t respectively. Then the predictive error ε_{it} is represented by

$$\varepsilon_{it} = r_{it} - \hat{r}_{it} \tag{1}$$

Thus, the time series of predictive errors of stock i is given by

$$\Delta = (\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iT}) \tag{2}$$

where T is the considered time horizon. As a non-biased estimator, the predictive errors must be statistically independent and identically distributed. Let n denotes the number of assets in portfolio, and x_i denotes the weights of each asset. Then, the predicted return, or expected return of portfolio X is represented by

$$r_X = \sum_{i=1}^{n} r_i x_i \tag{3}$$

where r_i represents the expected return of asset i. Thus, the risk of portfolio X at time t, i.e., the downside semi-absolute deviation can be represented as

$$w_t(X) = \frac{|\sum_{i=1}^{n}(r_i - r_{it})x_i| + \sum_{i=1}^{n}(r_i - r_{it})x_i}{2} \tag{4}$$

where $t = 1, 2, ..., T$. Therefore, the total risk $SAD(X)$ of portfolio X is obtained, which is presented as follows

$$SAD(X) = \frac{1}{T}\sum_{t=1}^{T} w_t(X) \tag{5}$$

With above analysis, the prediction-based portfolio optimization model is obtained, which is defined as follows

$$Min\ \frac{1}{T}\sum_{t=1}^{T} \frac{|\sum_{i=1}^{n}(r_i - r_{it})x_i| + \sum_{i=1}^{n}(r_i - r_{it})x_i}{2} \tag{6}$$

Subject to

$$\sum_{i=1}^{n} r_i x_i \geq R_p \tag{7}$$

$$\sum_{i=1}^{n} x_i = 1 \tag{8}$$

$$0 \leq x_i \leq 1 \quad i = 1, 2, ..., n \tag{9}$$

where R_p is the desired portfolio return. Eq. (6) is to minimize the portfolio's risk; Eq. (7) represents the desired return target which should more than the given threshold R_p; Eq. (8) represents the allocation of each asset; Eq. (9) represents the constraint of each asset allocation.

In the following, this prediction-base portfolio optimization model is simplified into a linear programming model. Let

$$d_t = w_t(X), t = 1, 2, ..., T \tag{10}$$

the following equivalent portfolio optimization model is obtained.

$$Min\ \frac{1}{T}\sum_{t=1}^{T} d_t \tag{11}$$

Subject to

$$d_t \geq 0 \quad t = 1, 2, ..., T \tag{12}$$

$$d_t \geq \sum_{i=1}^{n}(r_i - r_{it})x_i \quad t = 1, 2, ..., T \tag{13}$$

$$\sum_{i=1}^{n} r_i x_i \geq R_p \tag{14}$$

$$\sum_{i=1}^{n} x_i = 1 \tag{15}$$

$$0 \leq x_i \leq 1 \quad i = 1, 2, ..., n \tag{16}$$

Inspired by [35], this paper sets T as 5. The optimal values of $x_1, x_2, ..., x_n$ are obtained by solving this portfolio optimization model.

2) Equal weighted portfolio based on DMLP (DMLP+EW)

This paper applies DMLP+EW as comparison model, which builds portfolio by equally weighting stocks selected by DMLP. In this model, expected return of individual stock is first predicted by DMLP, then stocks with positive expected return are selected to build portfolios by equal weighted method.

B. PORTFOLIO OPTIMIZATION MODEL BASED ON LSTM NEURAL NETWORK

Hochreiter and Schmidhuber [36] first introduced LSTM neural network, which was designed to solve the long range dependency problem. Specifically, recurrent neural networks obtain sequential information by internal loops. This learning process is conducted through backpropagation algorithm which is adopted to adjust the weights between two layers. The slope acquires from the chain rule is sent to the activation function, then this slope becomes very small or large, which is the phenomenon of gradient vanishing or exploding. In other words, backpropagation algorithm in recurrent neural networks is fragile to the long range dependency. And, LSTM neural network is devised to solve this difficulty.

TABLE 2. Parameters of LSTM neural network

Parameter	value
Hidden nodes	5,10,15,20,25,30
Hidden layers	1,2,3,...,10
Learning rate	0.0001,0.001,0.01,0.1
Patient	0,5,10
Batch size	50,100,200
Dropout rate	0.1, 0.2,...,0.5
Recurrent dropout rate	0.1, 0.2,...,0.5
Activation function	relu
Loss function	Mean absolute error
Optimizer	RMSprop, Adam, AdaGrad, SGD

LSTM neural network contains similar hyperparameters with DMLP, all the discussed hyperparameters are presented in Table 2. Grid search is also used for searching the optimal hyperparameters and stochastic gradient descent method is adopted to train the LSTM neural network. And, earlystopping technology is used to reduce overfitting problem. Since relu function outperforms tanh function [34], relu function is adopted as activation function. After many trial and error, hidden node is set to 5, hidden layer is set to 1, learning rate is set to 0.001, patient is set to 0, batch size is set to 100, dropout rate is set to 0.1, recurrent dropout rate is set to 0.2, optimizer is set to RMSProp.

1) Prediction-based portfolio optimization model with LSTM neural network and MSAD (LSTM+MSAD)

LSTM+MSAD is similar with DMLP+MSAD, the only difference between them is the prediction model. LSTM+MSAD model uses LSTM as prediction model for future stock return prediction.

2) Equal weighted portfolio based on LSTM neural network (LSTM+EW)

This paper applies LSTM+EW as comparison model. Similar to DMLP+EW, the only difference of LSTM+EW model is that it applies LSTM for future stock return prediction.

C. PORTFOLIO OPTIMIZATION MODEL BASED ON CNN

CNN is a kind of DNN that contains convolutional layers with convolutional operation. Large numbers of CNN have been proposed in image classification and computer vision [37], [38]. CNN usually consists of convolutional layer, pooling layer and fully connected layer. Convolutional layer contains many filters and is often followed by pooling layer. Since stock price is a kind of time series, this paper uses one dimensional (1D) CNN for return prediction. Also, as CNN with 2×1 and 3×1 filter sizes perform similar, this paper applies 2×1 filter size in convolutional layer and maxpooling layer for simplicity. The consider hypeparameters of CNN are presented in Table 3. Stochastic gradient descent method is used to train the proposed CNN and earlystopping is used to reduce the overfitting problem.

TABLE 3. Parameters of CNN

Parameter	value
Filer numbers	2,4,8,16,32,64
Convolutional layers	1,2,3,...,10
Maxpooling layers	1,2,3,...,10
Fully connected layers	1,2,3
Fully connected layer nodes	2,4,8,16,32,64
Learning rate	0.0001,0.001,0.01
Patient	0,5,10
Batch size	50,100,200
Activation function	relu ,tanh
Loss function	Mean absolute error
Optimizer	SGD, Adam, AdaGrad, RMSprop

After multiple trial and error, the topology of CNN is obtained, the input layer is followed by a 1D convolutional layer (2 filters with 2×1 size), 1D maxpooling layer (2×1 size), 1D convolutional layer (2 filters with 2×1 size), 1D maxpooling layer (2×1 size), a fully connected layer (2 nodes) and the output layer. In addition, learning rate is set to 0.001, patient is set to 0, batch size is set to 100, optimizer is set to SGD, activation function is set to relu function.

1) Prediction-based portfolio optimization model with CNN and MSAD (CNN+MSAD)

Compared with DMLP+MSAD, the only difference between them is that CNN+MSAD model uses CNN instead of DMLP as prediction model for future stock return prediction.

2) Equal weighted portfolio based on CNN (CNN+EW)

Similar to DMLP+EW, the only difference of CNN+EW model is that CNN+EW model applies CNN for future stock return prediction.

D. PORTFOLIO OPTIMIZATION MODEL BASED ON SVR

SVR is a classic machine learning technique, which has been widely used in stock price prediction [15], [34]. SVR is a nonlinear kernel-based regression method which tries to locate a regression hyperplane with small risk in high-dimensional feature space. It possesses good function approximation and generalization capabilities [39].

This study uses radial basis function as kernel function of SVR, which is defined as follows:

$$K(v_i, v_j) = \exp(-\gamma \parallel v_i - v_j \parallel^2) \tag{17}$$

where γ is the parameter of radial basis function and v_i means the features of its training sample. The hyperparameters of SVR mainly contain C and γ, which are presented in Table 4. C represents the regularization parameter of SVR. Grid search is also used for searching the optimal hyperparameters. This paper applies SVR to build two prediction-based portfolios, i.e., SVR+MSAD and SVR+MV, as benchmarks.

TABLE 4. Parameters of SVR

Parameter	value
C	$2^0, 2^1, ..., 2^5$
γ	$2^{-5}, 2^{-4}, ..., 2^0$

1) Prediction-based portfolio optimization model with SVR and MSAD (SVR+MSAD)

In order to show the advantage of semi-absolute deviation in building prediction-based portfolio model, this paper formulates SVR+MSAD model by replacing the DMLP of DMLP+MSAD with SVR.

2) Prediction-based portfolio optimization model with SVR and MV(SVR+MV)

SVR+MV first uses SVR for stock return prediction, then builds prediction-based portfolio model with predictive errors of SVR measured by variance metric. The only difference between SVR+MV with SVR+MSAD is that SVR+MV uses variance metric instead of semi-absolute deviation metric to build prediction-based portfolio model.

IV. EXPERIMENTS

This section first presents the experimental data and the details of data preprocessing, then shows applied evaluation metrics.

A. DATA AND PREPROCESSING

This paper applies the China Securities 100 Index component stocks' historical data as experimental data. China Securities 100 Index is selected from the Shanghai and Shenzhen 300 Index component stocks, which represents the whole situation of the largest capitalization companies in Chinese stock market. This paper selects experimental data between January 4, 2007 and December 31, 2015, and the remainder of the component stocks consists of 49 stocks after neglecting some stocks that are halted or unlisted during this period. The final selected stocks' tickers are presented in Table 5.

TABLE 5. Selected stocks' tickers

000001	000002	000063	000069	000538	000625
000651	000725	000858	000895	002024	300059
600000	600010	600011	600015	600016	600018
600019	600028	600030	600031	600036	600048
600050	600104	600111	600115	600150	600276
600340	600372	600398	600485	600518	600519
600585	600637	600690	600795	600837	600886
600887	600893	600900	601006	601111	601398
601988					

The daily growth rate of close price, open price, high price, low price and volumes are used as input features. Actually, the growth rate $r(t)$ of close price $p(t)$ at time t is defined as follows

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}} \tag{18}$$

Similar to [6], [7], this paper applies past 20 days' daily growth rate of closing, opening, high, low prices and volumes as input features to predict the next day's return, i.e., the total number of input features is 20×5 for each prediction. Next, the process of data preprocessing is presented. For each input feature series $\{d_i\}$, d_i is modified as follows

$$d_i = \begin{cases} d_m + 5d_{mm} & if \;\; d_i \geq d_m + 5d_{mm}, \\ d_m - 5d_{mm} & if \;\; d_i \leq d_m - 5d_{mm}. \end{cases} \tag{19}$$

where d_m is the median of series $\{d_i\}$ and d_{mm} is the median of series $\{|d_i - d_m|\}$. Then, in order to unify the fluctuation range for model training, each modified input feature is standardized as follows.

$$\hat{x}_i = \frac{x_i - \mu}{\sigma} \tag{20}$$

where μ and σ denote mean and standard deviation of series $\{x_i\}$. After checking the daily return, this paper discovers that its value is relative small which is almost between -0.1 and 0.1. Thus, the daily return r_t is enlarged as sample target, which is presented as follows

$$r_t = \begin{cases} min(10r_t, 1) & if \;\; 10r_t \geq 1, \\ max(10r_t, -1) & if \;\; 10r_t \leq -1. \end{cases} \tag{21}$$

B. EVALUATION METRICS

The total experimental data consists of 9 years's data. Sliding window is used in the experiments, i.e., the first 4 years's data is training data, and the following year's data is validation data, then the next year's data is test data. Thus, the last four years' data ($2012-2015$) is used to measure the proposed portfolio model's performance. In addition, DMLP, LSTM neural network and CNN are implemented based on Keras deep learning package, and SVR is conducted by using Scikit-learn machine learning package.

This experiment applies mean squared error (MSE) and mean absolute error (MAE) to measure the predictive errors of different models. Two metrics are defined as follows

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(r_t - \hat{r}_t)^2 \tag{22}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|r_t - \hat{r}_t| \tag{23}$$

Also, the Hit Rates H_R, H_{R+}, H_{R-} are used to measure the prediction performance of different models, which are defined as follows

$$H_R = \frac{Count_{t=1}^{n}(r_t\hat{r}_t > 0)}{Count_{t=1}^{n}(r_t\hat{r}_t \neq 0)} \tag{24}$$

$$H_{R+} = \frac{Count_{t=1}^{n}(r_t > 0 \ \ AND \ \ \hat{r}_t > 0)}{Count_{t=1}^{n}(\hat{r}_t > 0)} \tag{25}$$

$$H_{R-} = \frac{Count_{t=1}^{n}(r_t < 0 \ \ AND \ \ \hat{r}_t < 0)}{Count_{t=1}^{n}(\hat{r}_t < 0)} \tag{26}$$

where H_R represents the total hit rate of model prediction, H_{R+} represents the accuracy of positive prediction and H_{R-} denotes the accuracy of negative prediction.

V. EXPERIMENTAL RESULTS OF PREDICTION-BASED PORTFOLIOS

In this section, this paper first applies five evaluation metrics to measure the predictive abilities of DMLP, LSTM neural network and CNN. Then, trading simulation without transaction fee is conducted to research the investing performance of different models under three desired portfolio returns.

A. PREDICTION OF DNNS

This section first measures the predictive performance of DMLP, LSTM neural network and SVR during the whole test period which consists of four years (2012-2015). This paper applies five evaluation metrics, i.e., mean absolute error (MAE), mean squared error (MSE), H_R, H_{R+} and H_{R-} to comprehensively measure their predictive abilities.

Since predictive errors are directly correlated with the prediction-based portfolio models, the MAE and MSE metrics are regarded as the key indicators among all the evaluation metrics. As we can see from Table 6-9, DMLP's mean predictive return errors measured by MAE and MSE metric are lower than the other models each year, and although their standard deviations are not the lowest, their difference is relatively small. Also, the mean H_R of DMLP is pretty high among these predictive models, and its standard deviation is low. However, for the H_{R-} and H_{R+} metric, DMLP performs no better than the other two models. Based on the above analysis, the predictive performance of DMLP outperforms the other models. Therefore, DMLP is a better model than the others in stock return prediction.

This result is consistent with the conclusion in [21]. This phenomenon is probably because of input features that influence the performance of different models. Since this paper only uses the past 20 days' historical data as input features, these input features have few time series information for LSTM neural network to learn long term correlations. Also, the limited information of input features is difficult for CNN to give full play to its advantages since CNN is known to be outstanding in image recognition, where pictures usually contain a lot of information. If the input features contain more historical data, the performance of LSTM neural network and CNN will be improved.

B. TRADING SIMULATION WITH DIFFERENT DESIRED PORTFOLIO RETURNS

This section presents trading simulation experiments to compare the performance of different prediction-based portfolio models during the whole test period. To be specific, this paper simulates trading behaviors like an ordinary investor. This investor decides to buy and sell certain scale of stocks in stock market each trading day after achieving the calculated proportion of each stock. For simplicity, trading costs, dividends and correlated taxes are set aside, also leveraging and short selling are neglected.

Since the prediction-based portfolio model needs an desired portfolio return R_p and the daily return of assets is mainly between -0.1 and 0.1, this paper only considers three values, i.e., $R_p = 0.001, 0.02, 0.04$, which represent different types of desired portfolio returns, i.e., low desired return, medium desired return and high desired return. In addition, in order to show the advantages of portfolio optimization models based semi-absolute deviation metric, this paper investigates the performance of equal weighted portfolio models, i.e., DMLP+EW, LSTM+EW and CNN+EW, respectively.

This trading simulation applies excess return, standard deviation, information ratio, total return, maximum drawdown, turnover rate and net value as evaluation metrics in order to comprehensively compare their investing abilities. To be specific, excess return is the acquired return after deducting the average return of total assets, and the standard deviation represents the volatility of excess return each month, information ratio denotes the excess return under unit risk, total return measures the total profits during the whole test period, maximum drawdown means the maximum holding risk based on historical net value graph, and turnover rate measures transaction fee caused by turnover. In addition, information ratio, maximum drawdown and turnover rate are defined as follows.

TABLE 6. The predictive performance of DMLP, LSTM neural network and CNN in 2012

Model		MAE	MSE	H_R	H_{R+}	H_{R-}
DMLP	mean	0.1406	0.0393	49.04%	47.69%	48.49%
	standard deviation	0.0425	0.0227	0.0366	0.1013	0.0373
LSTM	mean	0.1444	0.0411	48.23%	48.46%	48.87%
	standard deviation	0.0435	0.0235	0.0420	0.0593	0.2441
CNN	mean	0.1471	0.0463	48.45%	48.06%	45.79%
	standard deviation	0.0493	0.0356	0.0413	0.1547	0.1625

TABLE 7. The predictive performance of DMLP, LSTM neural network and CNN in 2013

Model		MAE	MSE	H_R	H_{R+}	H_{R-}
DMLP	mean	0.1729	0.0605	49.20%	44.38%	49.90%
	standard deviation	0.0472	0.0315	0.0358	0.1134	0.0396
LSTM	mean	0.1785	0.0636	47.64%	46.87%	49.76%
	standard deviation	0.0477	0.0335	0.0438	0.1051	0.1724
CNN	mean	0.1755	0.0634	48.78%	46.22%	50.40%
	standard deviation	0.0465	0.0328	0.0377	0.0955	0.1652

TABLE 8. The predictive performance of DMLP, LSTM neural network and CNN in 2014

Model		MAE	MSE	H_R	H_{R+}	H_{R-}
DMLP	mean	0.1528	0.0509	48.38%	50.62%	46.00%
	standard deviation	0.0377	0.0244	0.0395	0.0563	0.0385
LSTM	mean	0.1645	0.0597	48.79%	50.65%	41.23%
	standard deviation	0.0407	0.0339	0.0336	0.0741	0.1347
CNN	mean	0.1561	0.0535	48.18%	50.14%	48.02%
	standard deviation	0.0396	0.0265	0.0390	0.0778	0.1164

TABLE 9. The predictive performance of DMLP, LSTM neural network and CNN in 2015

Model		MAE	MSE	H_R	H_{R+}	H_{R-}
DMLP	mean	0.2617	0.1336	49.32%	49.46%	47.54%
	standard deviation	0.0555	0.0495	0.0340	0.0514	0.0335
LSTM	mean	0.2728	0.1457	49.60%	51.58%	50.81%
	standard deviation	0.0573	0.0578	0.0318	0.1298	0.1182
CNN	mean	0.2661	0.1413	48.85%	49.70%	47.26%
	standard deviation	0.0567	0.0560	0.0376	0.1157	0.0991

$$Information\ ratio = \frac{excess\ \ return}{stanard\ \ deviation} \tag{27}$$

$$Maximum\ \ drawdown = \max_{p<q} \frac{Nev_p - Nev_q}{Nev_p} \tag{28}$$

$$Turnover\ \ rate = \sum_{i=1}^{n} |x_{i,t} - x_{i,t-1}| \tag{29}$$

where Nev_p means the net value at time p, $x_{i,t}$ represents the proportion of stock x in the portfolio at time t and n is the total number of stocks in the portfolio. Net value tracks the performance of different models during the whole test period. Note that, this paper sets excess return as the key metric since it represents the profitability of different portfolio models.

1) Trading simulation with low desired return

This section presents the performance of different prediction-based portfolio models (i.e., DMLP+MSAD, LSTM+MSAD and CNN+MSAD) and equal weighted portfolio models (i.e., DMLP+EW, LSTM+EW and CNN+EW) under low desired return, i.e., $R_p = 0.001$. The experimental results are showed in Table 10 in detail, and their net value graphs are presented in Figure 1.

First, DMLP+MSAD and DMLP+EW are compared. From Table 10, we can obtain that DMLP+MSAD owns higher excess return and total return, DMLP+EW has lower standard deviation, maximum drawdown and turnover rate. Also, DMLP+EW's information ratio is higher. In order to further compare these two models, Mann-Whitney test is conducted to compare their excess returns, the test $p-$value equals to 0.032, which means that there is significant difference statistically between these models. Thus, DMLP+MSAD is a better model for investment.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3003819, IEEE Access

TABLE 10. The performance of different models when $R_p = 0.001$

Model	ER	SD	IR	TOR	MD	TUR
DMLP+MSAD	28.28%	0.4224	0.6696	152.27%	53.73%	41.78%
DMLP+EW	13.73%	0.1651	0.8315	116.33%	45.62%	18.16%
LSTM+MSAD	−1.11%	0.0562	−0.1972	53.81%	46.12%	11.17%
LSTM+EW	2.09%	0.0556	0.3752	76.94%	42.41%	7.66%
CNN+MSAD	14.95%	0.2148	0.6961	156.48%	37.93%	58.15%
CNN+EW	5.48%	0.0946	0.5791	87.42%	40.75%	33.78%

ER means excess return, SD means standard deviation, IR means information ratio, TOR means total return, MD means maximum drawdown, TUR means turnover rate.

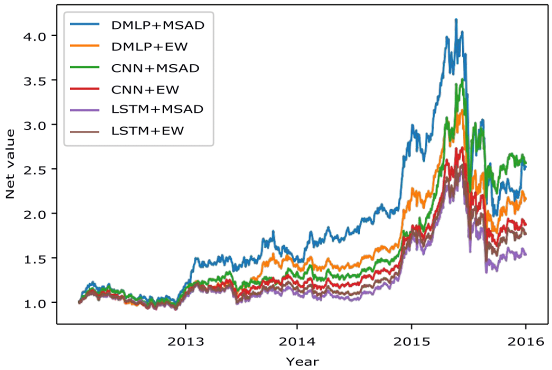


FIGURE 1. Net value of different models when $R_p = 0.001$.

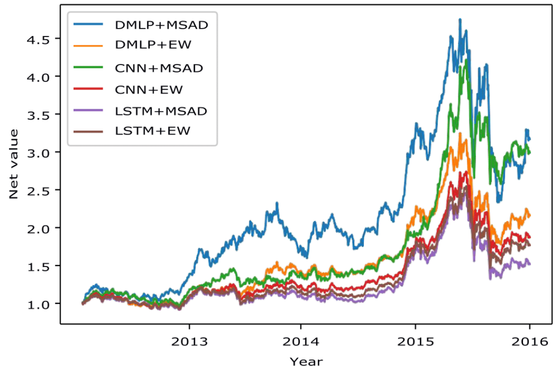


FIGURE 2. Net value of different models when $R_p = 0.02$.

Second, comparison of LSTM+MSAD and LSTM+EW is conducted. Table 10 shows that all the metrics of LSTM+EW perform better than LSTM+MSAD. Therefore, LSTM+EW outperforms LSTM+MSAD.

Third, the performance of CNN+MSAD and CNN+EW is discussed. Table 10 presents that CNN+MSAD owns higher excess return, information ratio and total return, and its maximum drawdown is lower. CNN+EW has lower standard deviation and turnover rate. Thus, Mann-Whitney test is used to further compare their excess returns. Test $p-$value equals to 0.001, which indicates that the null hypothesis is rejected and there is significant difference between these two models. Therefore, CNN+MSAD is a better choice for investment.

Fourth, DMLP+MSAD, CNN+MSAD and LSTM+EW are further compared. Table 10 shows that DMLP+MSAD owns higher excess return, CNN+MSAD has higher information ratio, total return and lower maximum drawdown, LSTM+EW possesses lower standard deviation and turnover rate. In order to further compare their differences, Mann-Whitney test is conducted to measure their excess returns. The test $p-$values of DMLP+MSAD and CNN+MSAD, CNN+MSAD and LSTM+EW, equal to 0.003 and 0.000 respectively, which means that their null hypotheses are rejected and their differences are significant statistically. Thus, DMLP+MSAD is a better choice compared with CNN+MSAD and LSTM+EW.

Based on the above analysis, this section concludes that DMLP+MSAD performs the best among all these models under low desired return.

2) Trading simulation with medium desired return

This section considers trading simulation based on desired return $R_p = 0.02$, which can be regarded as portfolio with medium desired return. The experimental results are presented in Figure 2 and Table 11. Figure 2 presents the net values of different models and Table 11 shows their detailed performance under multiple metrics.

Table 11 presents that DMLP+MSAD owns the highest excess return, information ratio and total return, LSTM+EW has the lowest standard deviation and turnover rate, CNN+MSAD possesses the lowest maximum drawdown. Thus, DMLP+MSAD, LSTM+EW and CNN+MSAD are selected for further comparisons. In order to distinguish these models, Mann-Whitney test is conducted to measure their excess returns. The test's result shows that $p-$values of DMLP+MSAD and CNN+MSAD, CNN+MSAD and LSTM+EW equal to 0.002 and 0.000, which means that their differences are significant statistically. Therefore, DMLP+MSAD outperforms the other two models.

Based on the above analysis, DMLP+MSAD performs the best among these models. Next, the performance of DMLP+MSAD under low and medium desired return is compared. Table 10-11 show that DMLP+MSAD with medium desired return owns higher excess return, information ratio

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3003819, IEEE Access

TABLE 11. The performance of different models when $R_p = 0.02$

Model	ER	SD	IR	TOR	MD	TUR
DMLP+MSAD	43.34%	0.4593	0.9436	216.86%	50.84%	52.31%
DMLP+EW	13.73%	0.1651	0.8315	116.33%	45.62%	18.16%
LSTM+MSAD	−1.38%	0.0581	−0.2370	52.04%	46.43%	11.78%
LSTM+EW	2.09%	0.0556	0.3752	76.94%	42.41%	7.66%
CNN+MSAD	22.92%	0.3350	0.6844	199.06%	38.92%	75.34%
CNN+EW	5.48%	0.0946	0.5791	87.42%	40.75%	33.78%

ER means excess return, SD means standard deviation, IR means information ratio, TOR means total return, MD means maximum drawdown, TUR means turnover rate.

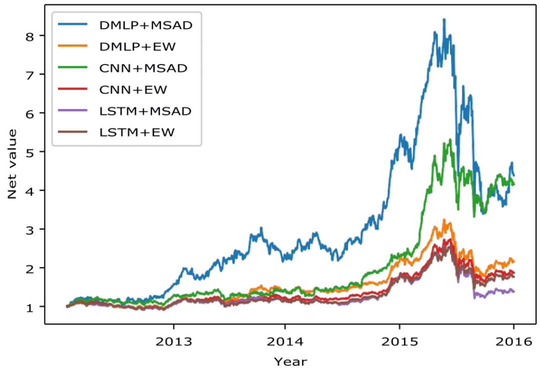


FIGURE 3. Net value of different models when $R_p = 0.04$.

and total return, lower maximum drawdown. But its standard deviation, turnover rate is higher. Also, T-test is conducted to measure their excess returns, test's result shows that $p-$value equals to 0.101, which means that there is no significant difference between these two models statistically. Thus, the profitability of DMLP+MSAD with medium desired return is not markedly improved compared with low desired return.

3) Trading simulation with high desired return

This section discusses the performance of different models under high desired return $R_p = 0.04$. Table 12 shows the experimental results of different metrics and Figure 3 presents their net value graphs.

Table 12 presents that DMLP+MSAD owns higher excess return, information ratio and total return, LSTM+EW possesses lower standard deviation and turnover rate, CNN+MSAD has lower maximum drawdown. Then, the performance of DMLP+MSAD, CNN+MSAD and LSTM+EW is further compared by using Mann-Whitney test to measure their excess returns. Test's result shows $p-$values of DMLP+MSAD and CNN+MSAD, CNN+MSAD and LSTM+EW equal to 0.002 and 0.000 respectively, which indicates there are significant differences between these models. Therefore, DMLP+MSAD outperforms the other models for investment.

Last, above analysis shows that the performance of DMLP+MSAD with low desired return and medium desired return is similar. Now, the performance of DMLP+MSAD under high desired return with medium desired return is compared. Table 11-12 show that DMLP+MSAD with high desired return owns higher excess return and total return, but its standard deviation, maximum drawdown and turnover rate are higher. Further, T-test is conducted to measure their excess returns, test's result presents $p-$value equals to 0.005, which means that the difference between them is statistically significant. Thus, DMLP+MSAD with high desired return is a better choice for investment.

Therefore, different desired returns do influence the performance of prediction-based portfolio models for investing. DMLP+MSAD model always performs the best under different desired returns, and high desired return is more suitable for DMLP+MSAD. This is mainly because DMLP owns the lower predictive errors among these models and the prediction-based portfolio models are built based on the predictive errors of different models. Thus, this paper only researches the performance of different models with transaction fee under high desired return for simplicity in the following.

VI. MODEL COMPARISON WITH TRANSACTION FEE

As is known to all, transaction fee can greatly influence the performance of trading strategy, and high turnover rate causes high transaction fee. Thus, it is meaningful to test the practical performance of prediction-based portfolio models after deducting their transaction fees. This paper only considers turnover fee of 0.05% per unit to research the performance of different models for simplicity. This section applies $R_p = 0.04$ to discuss the abilities of different prediction-based portfolio models.

First, the performance of DMLP+MSAD, CNN+MSAD and LSTM+MSAD is discussed. Table 13 presents that DMLP+MSAD's excess return, information ratio and total return are the highest among these models, and CNN+MSAD owns the lowest maximum drawdown. Then, Mann-Whitney test is conducted to further compare the performance of DMLP+MSAD and CNN+MSAD, test's result presents that $p-$value equals to 0.000, which means that there is statistically significant difference between them. Therefore, after deducting transaction fee, DMLP+MSAD still performs the best among prediction-based models. Not that LSTM+MSAD's excess return even becomes negative,

TABLE 12. The performance of different models when $R_p = 0.04$

Model	ER	SD	IR	TOR	MD	TUR
DMLP+MSAD	87.49%	0.9452	0.9256	338.41%	59.74%	82.24%
DMLP+EW	13.73%	0.1651	0.8315	116.33%	45.62%	18.16%
LSTM+MSAD	−0.67%	0.1040	−0.0649	38.91%	52.41%	18.74%
LSTM+EW	2.09%	0.0556	0.3752	76.94%	42.41%	7.66%
CNN+MSAD	38.41%	0.5296	0.7254	315.54%	37.58%	111.63%
CNN+EW	5.48%	0.0946	0.5791	87.42%	40.75%	33.78%

ER means excess return, SD means standard deviation, IR means information ratio, TOR means total return, MD means maximum drawdown, TUR means turnover rate.

TABLE 13. The performance of different models with transaction fee

Model	ER	SD	IR	TR	MD
DMLP+MSAD	52.71%	0.5821	0.9055	195.85%	61.17%
CNN+MSAD	9.14%	0.2024	0.4514	143.84%	39.19%
LSTM+MSAD	−4.11%	0.0975	−0.4211	26.96%	52.70%
SVR+MSAD	−13.98%	0.1608	−0.8696	−4.16%	51.74%
SVR+MV	−16.77%	0.4117	−0.4072	−12.33%	54.70%

ER means excess return, SD means standard deviation, IR means information ratio, TR means total return, MD means maximum drawdown.

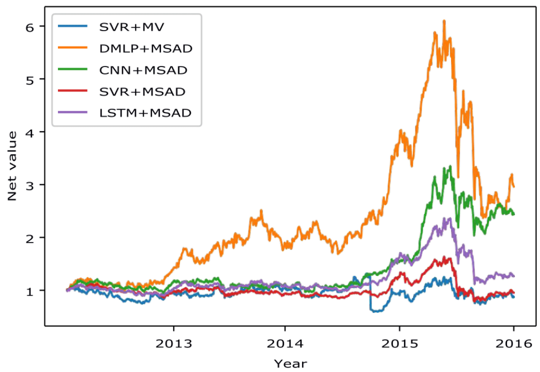


FIGURE 4. Net value of different models with transaction fee.

which means that it has no excess return after deducting its transaction fee cased by turnover.

Second, the performance of SVR+MSAD and SVR+MV is compared. Table 13 shows that all the metrics of SVR+MSAD outperform SVR+MV. Therefore, SVR+MSAD is a better choice for investment. In other words, semi-absolute deviation metric is more suitable for building prediction-based portfolio model than variance metric.

Last, comparison of DMLP+MSAD and SVR+MSAD is conducted. Table 13 shows that DMLP+MSAD owns higher excess return, information ratio and total return than SVR+MSAD, but its standard deviation and maximum drawdown are higher. Further, T-test is conducted to measure their excess returns, the test $p-$value equals to 0.000, which means there is statistically significant difference between these two models. Thus, DMLP+MSAD is better than SVR+MSAD.

In addition, all the above models' net value graphs are presented in Figure 4. It directly shows the performance of different models during the test period. Based on above analysis, this paper can deduct that DMLP+MSAD is a promising prediction-based portfolio model for practical investment.

VII. CONCLUSION

This paper applies three DNNs to build prediction-based portfolio optimization models, i.e., DMLP+MSAD, LSTM+MSAD and CNN+MSAD. These models consist of two parts that DNNs are first used for stock return prediction and their predictive errors measured by semi-absolute deviation metric are then utilized to build portfolio optimization models. DMLP, LSTM neural network and CNN are three frequently used DNNs which have been proved to own better learning abilities than traditional ML technologies. And semi-absolute deviation metric is more suitable than variance to measure risk. These models are compared with three equal weighted portfolio models (i.e., DMLP+EW, LSTM+EW and CNN+MSAD), SVR+MSAD and SVR+MV to show their superiorities.

First, five evaluation metrics, i.e., MAE, MSE, H_R, H_{R+} and H_{R-} are used to comprehensively measure to the predictive abilities of DMLP, LSTM neural network and CNN. Experiments show that DMLP outperforms the others in stock return prediction since it is more compatible with input features. Second, trading simulation is conducted to research the investing performance of DMLP+MSAD, LSTM+MSAD and CNN+MSAD without transaction fee. Three different desired returns are utilized to explore their performance. The experiments show that the DMLP+MSAD model always outperforms other models under different desired returns since DMLP owns the lowest predictive errors. And, high desired return is more suitable for DMLP+MSAD, which complies with the conclusion in [4]. Third, DMLP+MSAD, LSTM+MSAD and CNN+MSAD models are compared with two benchmark models, i.e., SVR+MSAD and SVR+MV, for investment with transaction fee. Experimental results show that the DMLP+MSAD model performs the best, and even deducting the transaction fee caused by turnover, it still earns considerable profits. In conclusion, this paper presents the promising ability of DNNs in prediction-based portfolio construction and encourages investors to apply the DMLP+MSAD model with higher desired return for practical investing.

This research further extends the literature concerning prediction-based portfolio optimization models by using DNNs for return prediction. As far as we know, this is the first attempt to use DNNs in building prediction-based portfolio optimization models, which fills the research gap in existing works. Also, the proposed prediction-based portfolio optimization models apply semi-absolute deviation as risk metric, which enables their applications in large scale portfolio optimization problems.

This paper also has limitations since it only applies simple historical data as input features in stock prediction process. Technical indicators, financial indicators, economic indicators can further improve the performance of DNNs in stock prediction. Also, there may exist other risk metrics that are more suitable than semi-absolute deviation in building portfolio optimization models. Future studies can apply more input features and better risk metric in building prediction-based portfolio models, and further improve the out-of-sample performance.

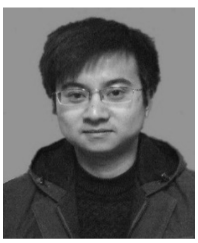
REFERENCES

[1] H. Markowitz, "Portfolio selection," in J. Financ., vol. 7, pp. 77-91, 1952.
[2] H. Konno and H. Yamazaki, "Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market," Manage. Sci., vol. 37, pp. 519–531, May. 1991.
[3] M. G. Speranza, Linear programming models for portfolio optimization. Financ., vol. 14, pp. 107-123, Jan. 1993.
[4] F. D. Freitas, A. F. De Souza and A. R. de Almeida, "Prediction-based portfolio optimization model using neural networks," in Neurocomputing, vol. 72, no. 10-112, pp. 2155-2170, Jun. 2009.
[5] E. Chong, C. Han and F. C. Park, "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies," in Expert Syst. Appl., vol. 83, pp. 187–205, Oct. 2017.
[6] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," in Eur. J. Oper. Res., vol. 270, pp. 654–669, Oct. 2018.
[7] C. Krauss, X. A. Do and N. Huck, "Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500," Eur. J. Oper. Res., vol. 259, no. 2, pp. 689–702, Jun. 2017.
[8] J. Patel, S. Shah, P. Thakkar and K. Kotecha, "Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques," Expert Syst. Appl., vol. 41, no. 1, pp. 259-268, Jan. 2015.
[9] R. Singh and S. Srivastava, "Stock prediction using deep learning," Multimed. Tools Appl., vol. 76, no. 18, pp. 18569-18584, Sep. 2017.
[10] K. Żbikowski, "Using Volume Weighted Support Vector Machines with walk forward testing and feature selection for the purpose of creating stock trading strategy," Expert Syst. Appl., vol. 42, no. 4, pp. 1797–1805, Oct. 2015.
[11] X. Yuan, J. Yuan, T. Jiang and Q U. Ain, "Integrated Long-term Stock Selection Models Based on Feature Selection and Machine Learning Algorithms for China Stock Market," in IEEE Access, vol. 8, pp. 22672–22685, Feb. 2020.
[12] J. Zhang, Y. H. Shao, L. W. Huang, et al. "Can the exchange rate be used to predict the Shanghai Composite Index?," in IEEE Access, vol. 8, pp. 2188-2199, Jan. 2019.
[13] Y. Y. Zhang, X. Li and S. N. Guo, "Portfolio selection problems with markowitz's mean-variance framework: a review of literature," in Fuzzy Optim. Decis. Ma., vol. 17, pp. 125–158, Jun. 2018.
[14] C. M. Lin, J. J. Huang, M. Gen and G. H. Tzeng, "Recurrent neural network for dynamic portfolio selection," in Appl. Math. Comput., vol. 175, no. 2, pp. 1139–1146, Apr. 2006.
[15] F. D. Paiva, R. T. N. Cardoso, G. P. Hanaoka and W. M. Duarte, "Decision-making for financial trading: A fusion approach of machine learning and portfolio selection," Expert Syst. Appl., vol. 115, pp. 635–655, Jan. 2019.
[16] W. Y. Wang, W. Z. Li, N. Zhang and K. C. Liu, "Portfolio formation with preselection using deep learning from long-term financial data," in Expert Syst. Appl., vol. 143, pp. 113042, 2020.
[17] F. D. Freitas, A. F. De Souza and A. R. Almeida, "Autoregressive neural network predictors in the Brazilian stock markets," in Proc. IEEE Lat. Am. Robot. Symp., pp. 1-8, 2005.
[18] F. D. Freitas, A. F. De Souza and A. R. Almeida, "A prediction-based portfolio optimization model," in Proc. 5st Int. Symp. Robot. Automat pp. 520-525, 2006.
[19] C. Hao, J. Wang, W. Xu and Y. Xiao, "Prediction-Based Portfolio Selection Model Using Support Vector Machines," in Proc. 6st Int. Conf. Bus. Intell. Financ. Eng, pp. 567-571, 2013.
[20] A. M. Rather, V. N. Sastry and A. Agarwal, Stock market prediction and Portfolio selection models: a survey. Opsearch, vol. 54, pp. 558–579, Jan. 2017.
[21] A. J. P. Samarawickrama and T. G. I. Fernando, "A Recurrent Neural Network Approach in Predicting Daily Stock Prices An Application to the Sri Lankan Stock Market," in Proc. IEEE Int. Conf. Ind. Inform. Syst., 2017.
[22] W. Long, Z. C. Lu and L. X. Cui, "Deep learning-based feature engineering for stock price movement prediction," in Knowl-Based Syst., vol. 164, pp. 163–173. Jan. 2019.
[23] B. Moews, J. M. Herrmann and G. Ibikunle, "Lagged correlation-based deep learning for directional trend change prediction in financial time series," in Expert Syst. Appl., vol. 120, pp. 197–206, Apr. 2019.
[24] D. M. Q. Nelson, A. C. M. Pereira and R. A. D. Oliveira, "Stock market's price movement prediction with LSTM neural networks," in Proc. Int. Conf. Joint Conf. Neur. Net., 2017.
[25] H. C. Lee and B. Ko, "Fund Price Analysis Using Convolutional Neural Networks for Multiple Variables," in IEEE Access, vol. 7, pp. 183626-183633, Nov. 2019.
[26] O. B. Sezer, M. U. Gudelek and A. M. Ozbayoglu, "Portfolio selection problems with markowitz's mean-variance framework: a review of literature," in Appl. soft comput., vol. 90, pp. 106181, 2020.
[27] S. I. Lee and S. J. Yoo, "Threshold-Based Portfolio: The Role of the Threshold and Its Applications," in J. Supercomput., https://doi:10.1007/s11227-018-2577-1
[28] F. Yang, Z. Chen, J. Li and L. Tang, "A novel hybrid stock selection method with stock prediction," in Appl. Soft Comput., vol. 80, pp. 820-831, Jul. 2019.
[29] M. Alizadeh, R. Rada, F. Jolai and E. Fotoohi, "An adaptive neuro-fuzzy system for stock portfolio analysis," in Int. J. Intell. Syst., vol. 26, no. 2, pp. 99–114, Feb. 2011.
[30] S. J. Deng and X. Y. Min, "Applied Optimization in Global Efficient Portfolio Construction Using Earning Forecasts," in J. Invest., vol. 22, no. 4, pp. 104–114, Nov. 2013.
[31] S. J. Kon, "Models of stock returns-a comparison," in J. Financ., vol. 39, no. 1, pp. 147–165, Mar. 1984.
[32] E. F. Fama, "Portfolio analysis in a stable Paretian market," in Manage. Sci., vol. 11, no. 3, pp. 404–419, Nov. 1965.
[33] A. L. D. Loureiro, V. L. Miguéis and L. F. M. da Silva, "Exploring the use of deep neural networks for sales forecasting in fashion retail," in Decis. Support Syst., vol. 114, pp. 81-93, Oct. 2018.
[34] L. O. Orimoloye, M. C. Sung, T. J. Ma and J. E. V. Johnson, "Comparing the effectiveness of deep feedforward neural networks and shadow architectures for predicting stock price indices," in Expert Syst. Appl., vol. 139, pp. 112828, 2020.
[35] J. R. Yu, W. J. P. Chiou, W. Y. Lee and S. J. Lin, "Portfolio models with return forecasting and transaction costs," in Int. Rev. Econ. Financ., vol. 66, pp. 118-130, Mar. 2020.
[36] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural comput., vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
[37] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in Proc. Int. Conf. Learn. Represent., pp. 7-9, 2015.

[38] H. Gunduz, Y. Yaslan, Z. Cataltepe, “Intraday prediction of Borsa Istanbul using convolutional neural networks and feature correlations,” in Knowledge-Based Systems, vol. 137, pp. 138-148, Dec. 2017.
[39] C. Y. Yeh, C. W. Huang and S. J. Lee, “A multiple-kernel support vector regression approach for stock market price forecasting,” in Expert Syst. Appl., vol. 38, no. 3, pp. 2177–2186, Mar. 2011.

YILIN MA received the master’s degree from the School of Mathematics, Southeast University, China, in 2017. He is currently pursuing the Ph.D degree with the School of Economics and Management, Southeast University. His research interests include deep learning and portfolio theory.

RUIZHU HAN received the Ph.D. degree in the School of Economics of Management, Southeast University, China, in 2003. She is currently a professor in the School of Economics of Management, Southeast University. Her research interests include logistics and supply chain management, complex social and economic system.



WEIZHONG WANG received the master’s degree in management science and engineering from China University of Mining and Technology, Xuzhou, in 2014. He is currently pursuing the Ph.D degree in the School of Economics of Management, Southeast University. His research results have been published in IISE Transactions, IEEE Transactions on reliability, Safety Science, Computers & industrial Engineering, International Journal of Industrial Ergonomics, among others. His current research interests include decision analysis, risk analysis, accident analysis.