Prediction of Agricultural Commodities Futures Prices: A DQN-LSTM Method

Hanxiao Xu
Ocean University of China
Jie Liang
Ocean University of China
Wenchaun Zang (zang12211@126.com)
Ocean University of China

Research Article

Keywords: Machine learning, Long short-term memory, Deep Q network, Agricultural commodity futures prices, DQN-LSTM framework

Posted Date: November 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1097759/v1

Prediction of agricultural commodities futures prices: A DQN-LSTM method

HanXiao Xu[a], Jie Liang[a] and Wenchuan Zang[b,*]

[a]College of Management, Ocean University of China, No.238 Songling Rd, Qingdao, 266100, Shandong, China
[b]College of Information Science and Engineering, Ocean University of China, No.238 Songling Rd, Qingdao, 266100, Shandong, China

ARTICLE INFO

Keywords:
Machine learning
Long short-term memory
Deep Q network
Agricultural commodity futures prices
DQN-LSTM framework

ABSTRACT

This paper combines deep Q network (DQN) with long and short-term memory (LSTM) and proposes a novel hybrid deep learning method called DQN-LSTM framework. The proposed method aims to address the prediction of five Chinese agricultural commodities futures prices over different time duration. The DQN-LSTM applies the strategy enhancement of deep reinforcement learning to the structural parameter optimization of deep recurrent networks, and achieves the organic integration of two types of deep learning algorithms. The new framework has the capacity of self-optimization and learning of parameters, thus improving the performance of prediction by its own iteration, which shows great prospects for future application in financial prediction and other directions. The performance of the proposed method is evaluated by comparing the effectiveness of the DQN-LSTM method with that of traditional predicting methods such as auto-regressive integrated moving average (ARIMA), support vector machine (SVR) and LSTM. The results show that the DQN-LSTM method can effectively optimize the traditional LSTM structural parameters through policy iteration of the deep reinforcement learning algorithm, which contributes to a better long and short-term prediction accuracy. In particular, the longer the prediction period, the more obvious the advantage of prediction accuracy of a DQN-LSTM method.

1. Introduction

As an important performance indicator of agricultural market price, agricultural commodities futures prices not only provide people engaged in the agricultural production and operation with more accurate information on long-term price fluctuations, but also serves as an important basis for hedging decisions for those involved in the agricultural industry chain. However, while performing their market functions, agricultural futures are subject to the inherent high risk characteristics of derivatives. On the one hand, agricultural commodities futures prices are affected by many factors such as economic globalization, financial crisis, climate change and oil price fluctuations, being highly complex and non-linear, which makes accurate prediction of agricultural commodities futures prices extremely challenging. On the other hand, margin trading and forced liquidation systems make agricultural futures price fluctuation exponentially more profitable for hedgers and speculators. If mishandled, it can easily lead to extreme risks in financial markets and undermine agricultural production and management activities. Therefore, an in-depth analysis of agricultural futures price fluctuations and accurate analysis and prediction of trends of agricultural commodities futures prices are important for hedgers in the agricultural industry chain to hedge price risks, and for speculators to invest rationally and guide agricultural production and resource allocation.

Recently, deep learning, especially deep recurrent neural networks have been widely used in many fields [1, 2, 3, 4]. But for the the late start of the commodity futures market, the existing literature exploring the prediction of agricultural commodity futures prices is relatively limited. However, there have been many attempts by scholars to apply a variety of modeling approaches to financial time series prediction. Early scholars mainly applied the statistical analysis models to predict agricultural products and agricultural commodities futures prices based on statistical theory. Commonly used models include Autoregressive model[5], Moving Average model [6], Exponential Smoothing model [7, 8], Autoregressive Moving Average model [9], and Autoregressive Conditional Heteroskedasticity model[10, 11, 12]. However, the factors affecting agricultural commodities futures prices are complex and variable, while the data are non-linear and time-varying. Although the traditional statistical methods have good theoretical support and valid

*Corresponding author: Wenchuan Zang. College of Information Science and Engineering, Ocean University of China, China.
emmail: zang12211@126.com
ORCID(s):

explanations, it is difficult to accurately predict the complex and changeable agricultural commodities futures prices because the nature of the data patterns have linear features.

With the flourishing of machine learning, the application of artificial intelligence is gradually extended. Compared with econometric methods, machine learning methods are widely used in financial time series prediction as they can mine valuable information directly from data without pre-formulated assumptions and machine learning can better handle non-linear data. [13] and [14] used Convolutional Neural Network (CNN) and Back Propagation Neural Network (BPNN) respectively to predict ETF prices. [15] used the BAT algorithm to predict copper price fluctuations. The experimental results showed that the BAT algorithm outperformed the classical prediction method. [16] and [17] used Decision Tree Algorithm and Support Vector Machine(SVM) to predict price of copper futures in London Metal Exchange, respectively. Although traditional neural networks have good prediction capability, their accuracy is still not satisfactory when faced with dynamic and non-linear time series data.

In recent years, deep recurrent neural network such as LSTM has presented prominent performance in time series prediction. Unlike traditional RNNs, LSTM introduce memory units where the memory of previous inputs can be persistently stored in the internal states of the network, thus allowing LSTM network to explore serial data such as time series. LSTM network with a better prediction a bility can also explore the abstract features inherent in the data, grasp the hidden structure constants in the data, and then process the time series. [18] used LSTM network to predict the monthly closing prices of stocks and used the prediction results to determine portfolio weights. [19] used EMD (Empirical Mode Decomposition) and LSTM to predict stock prices and introduced a factor of Investor Sentiment to further improve the prediction accuracy. [20] combined LSTM with various GARCH and proposed a new hybrid LSTM method for predicting price volatility of stocks. These studies show that LSTM has good performance in learning the dependencies between complex non-linear time series. However, as research on deep learning methods applied to financial markets deepens, the limitations of methods such as LSTM have also veen revealed, such as overfitting of LSTM network in the prediction of time-series data, which affects the prediction accuracy[21]. The Dropout operation is a common means to address the problem of overfitting, as overfitting can be effectively solved by selecting reasonable Dropout parameters[22]. However, there is no systematic research on the selection of such parameters and the patterns are not yet summarized. In addition, it is a sequential decision problem for deep LSTM network as to which part of the structure should the Dropout parameters be applied and what is its proper value, which cannot be solved effectively by conventional means. Deep reinforcement learning is an artificial intelligence method widely used in industrial manufacturing [23], path planning[24], and gaming[25], which can solve the challenge of sequential decision making in the selection of Dropout parameters to some extent[26]. DQN is the earliest one using deep neural networks among the many deep reinforcement learning algorithms, which has the advantages of simple structure, relatively low hyperparameter proficiency and easy implementation, and has been widely used in multiple fields.

In response to low prediction accuracy caused by overfitting in the process of LSTM algorithm time series prediction and the randomness faced by the selection of Dropout parameters to solve the overfitting problem, this paper proposes a hybrid prediction method (called DQN-LSTM) combining DQN and LSTM by intelligently deciding the Dropout parameters in the structure of LSTM network through DQN algorithm, which can effectively improve the generalization and robustness of the method. Specifically, with the running status of LSTM network as the observation value of the DQN algorithm and the Dropout parameters value option as the action value of the DQN algorithm, the two is fitted using a deep reinforcement learning network. The correlation between the training data is destroyed and the independent homogeneous distribution of the method training data is improved by establishing a priority experience replay mechanism and fixed-target mechanism. The method's performance on predicting the prices of agricultural commodities is effectively improved through iterative training.

The possible contributions of this paper are summarized as follows: (1) A novel framework for predictions of agricultural commodities futures prices is proposed by combining DQN and LSTM network. The DQN-LSTM method can tap into the hidden high-level interdependencies in data of agricultural commodities futures prices and has better prediction performance by addressing the problem of differences in the performance of traditional LSTM algorithm in the training set and test set. (2) The DQN algorithm is embedded in the operation of the LSTM algorithm to make decisions on the algorithm parameters, while the overall LSTM prediction process is used as the single-act learning process of the DQN algorithm, achieving the deep coupling of the two algorithms. (3) Based on DQN's intelligent decision-making capability, the LSTM algorithm with dynamic parameters is realized, which outperforms traditional LSTM algorithm with fixed parameters. Deeper levels of deep learning methods such as LSTM help to improve the feature extraction capability of the method, while the application of fixed parameter methods limits the performance

of the method. The manual tuning of parameters or building an optimised method involves a huge workload, while artificial intelligent tuning of parameters achieves fast and dynamic calculation of optimal operating parameters.

2. Methodology

This section describes the basic architecture of DQN and LSTM network, on the basis of which the general framework of the DQN-LSTM method is introduced to explore the theoretical basis and feasibility of establishing a DQN-LSTM prediction method for agricultural commodities futures prices time series data.

2.1. The deep Q network algorithm

The deep Q network algorithm (DQN) is a technique that combines reinforcement learning and deep neural networks. A schematic representation of the DQN algorithm is shown in Fig.1. This network acts as an approximation function fitting the non-linear relationship between system observations and system actions. Given a state variable s, the action value (or Q-value) of all actions in that state is calculated as an estimate of the future cumulative reward, and the action with the highest action value is then selected using the epsilon-greedy mechanism. The optimal action policy for the problem is achieved through continuous optimization of the neural network parameters. The approximation function is denoted as $Q(s, -; w)$, where w is the set of weights of the network. When used as a controller, the action with the largest Q value (or target value) in the current state S is executed for the system as a control signal. The DQN algorithm has two important techniques, namely fixed-target and experience-replay, which are to increase the speed of convergence and improve the effectiveness of the results [25]. The target Q-network is an isolated network used to generate temporal differential signals during training. Instead of being updated at each training step, the target Q-network is updated at a fixed period to avoid the frequent follow-through of the temporal differential signal due to updates of the neural network parameters. Then the target value function of the DQN is demonstrated in Equ.1.

$$U_i = R_i + \gamma \max_a Q(S_i', a; w_{target}) \tag{1}$$

In the Eq.1, the time differential target value of the DQN is given, the reward R_i is obtained from the current step, and γ is the discount factor. S_i' is the state at the next time step and a the action taken. w_{target} is the target Q-network at time t. In this way, the training process for the DQN evaluation parameters from each extracted experience i can be expressed as Equ.2.

$$w' = w + \frac{\alpha}{|\mathcal{B}|} \cdot \sum_{i \in \mathcal{B}} [U_i - q(S_i, A_i; w)] \nabla q(S_i, A_i; w) \tag{2}$$

where w is the parameter of the evaluation network, α is the learning rate, and $\mathcal{B}$ is the batch size data sampled from an experience base consisting of all experiences.

As for experience replay, the data for DQN training is randomly drawn from an experience memory pool[27]. This pool stores the results of the operation system, as the empirical data is highly continuous and correlated, which affects the training convergence of DQN because it does not satisfy the requirement of neural network parameter training regarding the independent identical distribution of the training data, and the experience replay method can reduce the correlation of the data samples. In order to make use of high-quality experience more effectively, [28] proposed a priority experience replay mechanism, which firstly assigns a certain priority to all experiences according to their performance, then ranks the experiences according to their priority, and prioritizes the experiences with higher priority for neural network training when performing experience replay.

2.2. The long short-term memory network

LSTM, an extended version of the recurrent neural network architecture, improves on the memory module in traditional RNN model, i.e. a single tanh or sigmoid layer, and addresses the problem of not being able to preserve valid historical information in the long term due to the influence of continuous data input. It is able to deal with the long-term dependence of time-series data by introducing a gating mechanism to replace the nodes in RNNs.

The basic structure of a LSTM network is made up of a series of recurrently connected sub-networks (i.e. memory modules), each with an internal configuration (as shown in Fig.2). A memory module is basically a memory store,

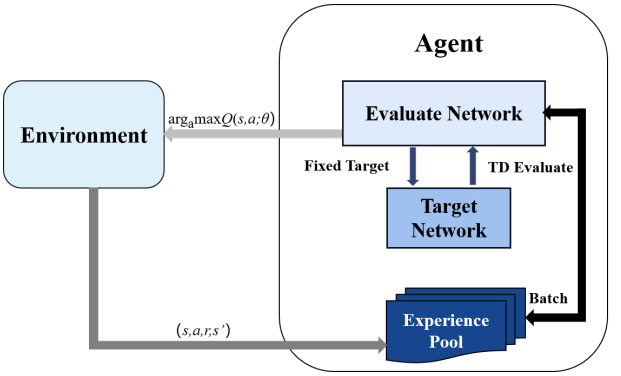

Figure 1: Schematic diagram of the DQN algorithm.

and each memory module contains of three types of gates: an input gate, a forget gate and an output gate. The LSTM network computes a mapping from the historical inputs, in which equence x to a predicted output sequence $y = (y_1, y_2, ..., y_T)$, from $t = 1$ to T, where T is the prediction period, iteratively performing the following steps:

First, the input gate determines what new information should be stored in the unit state and creates new candidates that may be added to the state $\hat{C}_t$.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\hat{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{4}$$

Then the forget gate determines which information should be cleared from the unit state.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{5}$$

The old unit state C_{t-1} is updated to the new unit state C_t by discarding some of the information from the old unit and adding filter candidates.

$$C_t = f_t \odot C_{t-1} + i_t \odot \hat{C}_t \tag{6}$$

Finally, the output gate filters the unit state and calculates the required information, with the final output shown as following:

$$o_t = \sigma(W_o \cdot [h_{t-1,x_t}] + b_o) \tag{7}$$

$$h_t = o_t \odot \tanh(C_t) \tag{8}$$

$$y_t = \phi(W_y h_t + b_y) \tag{9}$$

where C_t is the unit state, $\hat{C}_t$ is the new state after the update, h_t is the hidden layer state, i.e. the activation of the memory block. W_i, W_f, W_c, W_o and W_y denote the corresponding weight matrices, while b_i, b_f, b_c, b_o and b_y denote the corresponding deviation vectors, which can be determined by the back-propagation algorithm. Additionally, σ and tanh are the Sigmoid function and Hyperbolic tangent function respectively, where $\odot$ is the element product of the vectors and *phi* is the network output activation function.

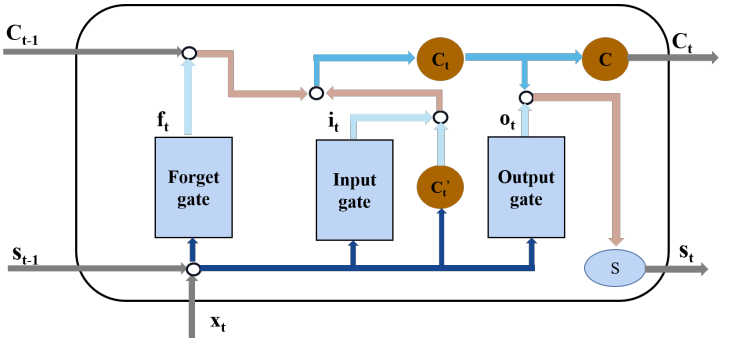

Figure 2: Schematic diagram of the basic structure of LSTM network.

2.3. The DQN-LSTM method

In view of the advantages of DQN and LSTM, this paper combines DQN and LSTM to build a new hybrid prediction algorithm: the DQN-LSTM method. LSTM is used as the prediction core of the proposed prediction mechanism, where the input data contains two sets of features, time and closing price, and the original data is reconstructed into 2-D data by sliding time step to decompose the input signal and output signal, after which their components are modeled and predicted.

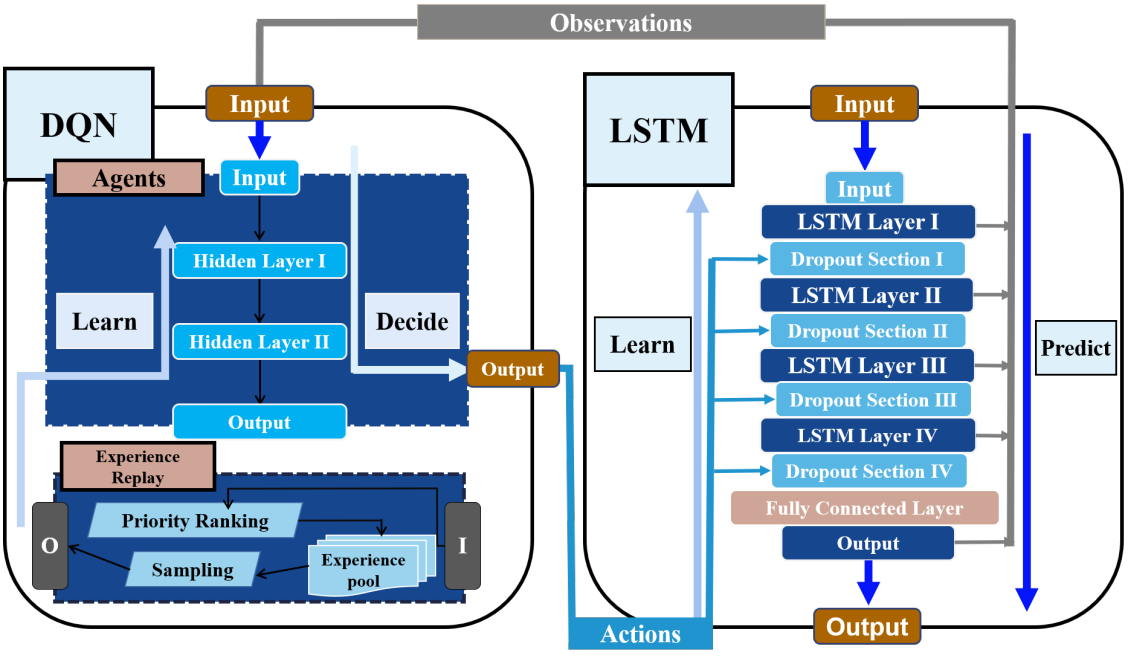

Figure 3: Schematic diagram of the hybrid DQN-LSTM method.

The DQN-LSTM method consists of two main components, the DQN and the LSTM network, which are connected through observation signals and action signals, as the Fig.3 shows. DQN consists of two structurally identical neural networks for training. The parameters of one are for strategy evaluation while the parameters of the other are for the computation of temporal differential signals and maintaining the parameters fixed for a given simulation step. Both neural network structures contain two hidden layers, each with 128 and 256 hidden neural units, respectively, and both neural networks receive observations from LSTM and output the decision Dropout parameters for transmission to the LSTM network as action parameters for DQN. The DQN method also includes an experience playback pool, where DQN does not perform training during the experience accumulation phase, but includes all collected observation-

action pairs into the experience pool, and then prioritizes them and randomly draws experience values according to the probability size determined by the priority into the DQN network for parameter training.

The LSTM network consists of four LSTM layers, four Dropout layers and one fully-connected layer. LSTM receives pre-processed 2-D data on agricultural commodities futures prices and processes and output them through the LSTM network to generate predictions of futures prices of agricultural products for the next time window. Each Dropout layer triggers the observed value generation program to send observed value to the DQN for making decisions on the dynamic parameters of the LSTM, and the Dropout parameters decided by DQN are passed back to the pending LSTM program to continue the price prediction.The inverse of the mean square error of LSTM price prediction is used as a reward signal for DQN to improve its performance in decision making. The LSTM algorithm acts as a dynamical environment for the DQN algorithm to provide direction for the iterative evolution of DQN, while the Dropout parameters of the DQN decision in turn enhances the prediction capability of LSTM, resulting in competitive iteration and co-evolution between the pair. Specifically, Algorithm.1 shows the complete pogram of the given DQN-LSTM framework.

Algorithm 1 DQN-LSTM algorithm

Input: Historical futures prices of agricultural commodities $x(t)$
Output: Predicted futures prices of agricultural commodities $y(t)$

Initializing;
  Step1: Determine the hyperparameters of DQN and LSTM networks;
  Step2: Normalize the price data;
  Step3: Construct the feature-label pair of training data according to the size of the prediction window N;
  Step4: Shuffle the training data
Forward propagation;
  # DQN-based Dropout parameter decision;
  While l in layers:
    Step1: Prediction via LSTM:$y_l = \phi(W_y h_t + b_y)$;
    Step2: Determine Dropout via DQN:$Dropout = \text{argmax} q(l, \cdot; w)$;
    Step3: Update state:$l = l + 1$;
    Step4: Calculate the loss and obtain the reward r;
    Step5: Pack the experience $(l, Dropout, r, l')$ and send the experience into the pool;
  End While
# Backward propagation;
  Step1: Calculate the gradients via $\frac{\partial J}{\partial y_t}, \frac{\partial J}{\partial c_{t+1}}, \frac{\partial J}{\partial \hat{o}_{t+1}}, \frac{\partial J}{\partial \hat{f}_{t+1}}, \frac{\partial J}{\partial \hat{i}_{t+1}}, \frac{\partial J}{\partial \hat{o}_{t+1}}$;
  Step2: Update the parameters of LSTM;
  Step3: Sample the batch data $\mathcal{B}$ from the experience pool;
  Step4: Update w via: $w = w + \alpha \frac{1}{\xi} \sum_{iin\xi} [U_i - q(S_i, A_i; w)] \nabla q(S_i, A_i; w)$;

3. Raw data discussion and evaluation criteria

3.1. Description of the data

This paper selects daily data on the closing prices of five agricultural commodity futures, including soybean No. 1, cotton, soybean meal, soybean oil and corn, which are representative of the Chinese market and can reflect the overall fluctuation of the market they are in. As can be seen from Fig.4, due to the sharp fluctuations in the original agricultural commodities futures prices series, the data distribution is highly variable and irregular, with very obvious complex features such as non-linearity and non-smoothness. Therefore, the data series of five agricultural commodity futures prices selected by this paper can evaluate the effectiveness and practicality of the proposed method more comprehensively and systematically than a single data series. The sample has 3519 sets of data in total, which is selected from 9 January 2006 to 24 June 2020, excluding the effects of holidays and other factors, and were obtained from The Wind Information Financial Terminal. In addition, in order to observe the effects of different prediction methods on short-term, medium-term and long-term predictions, the last 20, last 60 and last 365 trading days of the overall futures prices data set are taken as the test sets for short-term, medium-term and long-term predictions respectively, while the eliminated from the test set forms the training set. The training set is used to train the method parameters (e.g. weight matrix W) and the test set is used to evaluate the generalization ability of the trained method (i.e. to evaluate the prediction ability of out-of-sample time series data).

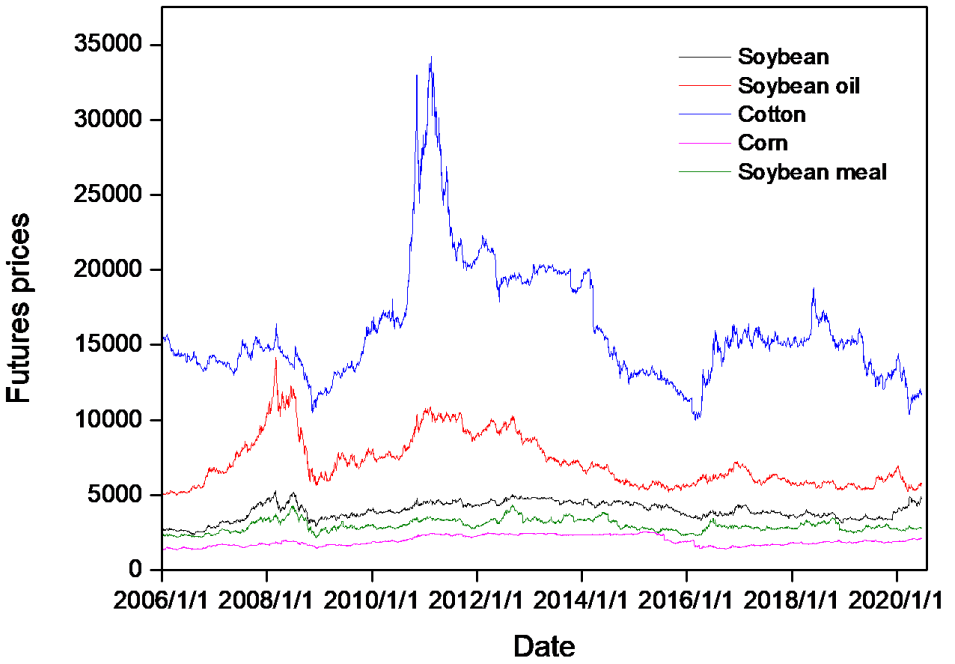

Figure 4: Agricultural commodities futures prices, 9 January 2006 to 24 June 2020.

3.2. Evaluation criteria

It is essential to apply a variety of performance metrics when evaluating the predictive capability of the method developed. In this paper, two evaluation criteria are used: horizontal and directional prediction. In order to improve the horizontal prediction accuracy, MAE, $MAPE$ and $RMSE$ evaluation metrics are used, among which, MAE is the mean absolute error, $MAPE$ the mean absolute percentage error and $RMSE$ the root mean square error. The specific calculation formulas are demonstrated through Equ.10 to Equ.12.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{10}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{|y_i|} \times 100\% \tag{11}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{12}$$

where N is the number of predicted data, y_i and $\hat{y}_i$ are the actual and predicted values. Generally, the smaller the values of the evaluation metrics of $RMSE$, MAE and $MAPE$, the closer the predicted value of the method is to the true value, which means that the method has a higher prediction accuracy. From an economic point of view, the ability to predict the correct direction is even more important than the accuracy of the directional predictions. It can

Table 1
Optimization results of DQN-LSTM with different combinations of the Dropout parameters

Time-span	Dropouts of layers (action)	Reward	MAE	$MAPE(\%)$	$RMSE$	$D_{stat}(\%)$
Short-term	(2, 2, 1, 1)	0.029	8.481	1.349	12.120	96.234
	(1, 1, 1, 1)	0.027	10.059	1.412	13.851	92.590
	(2, 2, 0, 0)	0.024	13.044	5.754	15.950	93.641
	(0, 0, 0, 0)	0.021	15.894	5.801	18.457	91.381
	(0, 0, 0, 1)	0.003	17.160	6.102	18.590	89.332
Medium-term	(0, 0, 2, 2)	0.029	8.832	1.500	12.260	92.372
	(2, 2, 2, 2)	0.028	11.365	1.579	14.641	90.590
	(2, 2, 0, 0)	0.026	12.990	4.760	17.873	89.524
	(0, 0, 0, 0)	0.017	15.533	5.149	19.653	88.330
	(0, 0, 1, 1)	0.003	17.570	5.247	21.740	83.730
Long-term	(2, 2, 0, 0)	0.029	10.648	4.617	13.570	90.623
	(1, 1, 1, 1)	0.027	12.56	4.888	14.251	84.930
	(0, 0, 1, 1)	0.024	19.127	5.072	16.045	82.852
	(0, 0, 0, 0)	0.003	24.658	5.184	21.067	78.275
	(1, 1, 0, 0)	0.001	25.064	6.355	26.739	75.340

be measured by the directional statistic D_{stat}. It is usually defined as Equ.13, where the A_i can be further expressed as Equ.14

$$D_{stat} = \frac{1}{N}\sum_{i=1}^{N} A_i \tag{13}$$

$$A_i = \begin{cases} 1, & (y_{i+1} - y_i)(\hat{y}_{i+1} - y_i) \geq 0 \\ 0, & otherwise \end{cases} \tag{14}$$

4. Empirical results and analysis

4.1. Optimization results of the DQN-LSTM method

The optimization algorithm for the LSTM network training in this paper uses the DQN algorithm to make decisions on the Dropout parameters. To prevent over-fitting of method training, the training process further employs Dropout to eliminate certain hidden units in the hidden layer. Table.1 lists the DQN-LSTM results for different parameter combinations. In order to save space, this paper only lists the DQN-LSTM optimization results of prediction of short-term futures prices of corn with different parameter combinations. It can be seen that the different Dropout parameters combinations significantly affect the prediction performance, where the higher the reward value, the higher the prediction accuracy. The best Dropout parameters combination $(2, 2, 1, 1)$ improves the MAE, $MAPE$, $RMSE$ and $Dstst$ values by $8.679\%, 4.753\%, 6.470\%$ and 6.902% respectively compared with the worst parameter combination $(0, 0, 0, 1)$. Therefore, in the following analysis, this paper selects the highest REWARD value dropout parameters combinations for further comparative analysis.

4.2. Comparative analysis of short-term prediction results

This paper measures the short-term performance of four prediction methods, namely the DQN-LSTM method, the SVR method, the ARIMA method and the LSTM network, using a time duration of 20 trading days. Table.2 reports the comparison of prediction results of short-term agricultural commodities futures prices of different prediction methods.

In terms of horizontal prediction accuracy, overall, the DQN-LSTM method has the best prediction accuracy, which is followed by the LSTM method, the SVR method, while the lowest prediction accuracy is from the ARIMA method.

Table 2
Short-term agricultural commodities futures prices predictions results

Futures varieties	Method	MAE	$MAPE(\%)$	$RMSE$	$D_{stat}(\%)$
Corn	ARIMA	92.803	5.100	93.896	88.512
	SVR	12.346	1.703	16.409	73.732
	LSTM	10.199	1.377	12.530	93.412
	DQN-LSTM	**8.481**	**1.349**	**12.120**	**96.234**
Soybean meal	ARIMA	60.958	2.180	64.834	86.559
	SVR	30.002	7.693	42.200	80.486
	LSTM	**18.325**	1.739	52.460	91.111
	DQN-LSTM	18.776	**1.723**	**25.490**	**94.932**
Soybean oil	ARIMA	49.811	0.761	60.867	84.609
	SVR	60.33	11.687	64.064	80.235
	LSTM	44.765	1.440	60.610	92.272
	DQN-LSTM	**34.800**	**0.657**	**48.110**	**93.640**
Soybean	ARIMA	125.926	3.922	132.184	76.760
	SVR	67.843	8.546	78.373	53.135
	LSTM	34.194	3.866	47.890	90.184
	DQN-LSTM	**27.696**	**3.783**	**43.270**	**91.134**
Cotton	ARIMA	101.404	1.852	119.622	75.139
	SVR	186.367	13.146	101.860	57.716
	LSTM	124.853	1.594	174.983	87.174
	DQN-LSTM	**85.486**	**0.861**	**101.210**	**90.586**

The LSTM method only slightly outperforms the DQN-LSTM method in terms of MAE value for soybean meal. The MAE, $MAPE$ and $RMSE$ values of the method are 8.481, 1.349 and 12.120, an increase of 84.322, 3.751 and 81.776 respectively compared with the ARIMA method; an increase of 3.865, 0.354 and 4.289 respectively compared with the SVR method; and an increase of 1.718, 0.028 and 0.410 respectively compared with the LSTM method.

In terms of directional prediction accuracy, the D_{stat} value of the DQN-LSTM method are all higher than those of the other three prediction methods, followed by the LSTM method and the ARIMA method, and the SVR method with the lowest directional prediction accuracy. The D_{stat} value of the DQN-LSTM method is 96.234%, which is 7.722% higher than that of ARIMA method, 22.502% higher than that of the SVR method and 2.822 higher than that of the LSTM method. Based on the successful combination of these two algorithms, the proposed DQN-LSTM method has the best prediction results among all the methods considered due to the unique advantages of DQN and LSTM in terms of parameter self-iterations and time series processing, respectively. The pattern of variation are generally consistent in the prediction of futures prices of soybean meal, soybean oil, soybean and cotton, reflecting the robustness of the prediction method.

4.3. Comparative analysis of medium-term prediction results

Also, this work measures the medium-term performance of above mentioned four prediction methods using a time duration of 60 trading days. Table.3 reports a comparison of the medium-term agricultural commodities futures prices prediction results of the different prediction methods.

In terms of horizontal prediction accuracy, overall, the DQN-LSTM method has the best prediction accuracy, which is followed by the LSTM network method, the SVR method, and the ARIMA method with the lowest prediction accuracy. The LSTM only slightly outperforms the DQN-LSTM method in terms of $RMSE$ value for soybean meal. Taking the corn futures prices prediction results as an example, the MAE, $MAPE$ and $RMSE$ values of the DQN-LSTM method are 8.832, 1.496 and 12.260, an increase of 89.269, 3.725 and 86.674 respectively compared with the ARIMA method; and an increase of 3.865, 0.354 and 12.526 respectively compared with the SVR method. Compared With the LSTM method, the increase are 2.71, 0.142 and 0.240 respectively.

In terms of directional prediction accuracy, overall, the DQN-LSTM method has the hightest D_{stat} value, followed

Table 3
Medium-term agricultural commodities futures prices prediction results

Futures varieties	Method	MAE	$MAPE(\%)$	$RMSE$	$D_{stat}(\%)$
Corn	ARIMA	98.101	5.221	98.934	84.398
	SVR	24.786	1.705	26.936	67.874
	LSTM	11.542	1.638	14.500	90.270
	DQN-LSTM	**8.832**	**1.496**	**12.260**	**92.372**
Soybean meal	ARIMA	61.961	2.201	64.918	78.936
	SVR	30.429	7.840	42.918	81.152
	LSTM	20.157	2.114	**26.820**	89.438
	DQN-LSTM	**19.858**	**2.101**	28.080	**92.561**
Soybean oil	ARIMA	50.114	4.970	63.664	59.195
	SVR	63.096	12.176	65.496	79.669
	LSTM	47.566	1.496	63.540	85.679
	DQN-LSTM	**39.855**	**0.927**	**56.340**	**89.448**
Soybean	ARIMA	122.299	4.001	126.754	78.997
	SVR	66.225	4.497	75.705	56.549
	LSTM	32.651	3.982	44.130	83.384
	DQN-LSTM	**24.564**	**3.927**	**38.460**	**90.662**
Cotton	ARIMA	246.793	3.886	268.620	70.857
	SVR	201.713	14.660	224.880	50.598
	LSTM	134.126	3.554	178.630	83.709
	DQN-LSTM	**128.902**	**1.131**	**168.996**	**88.582**

by the LSTM method and ARIMA method, and the SVR method with the lowest directional prediction accuracy. The DQN-LSTM method has a D_{stat} value of 92.372%, which is 7.974% higher than that of the ARIMA method, 24.498% higher than that of the SVR method, and 2.102 higher than that of the LSTM method. This pattern also applies in the predictions of the futures prices of soybean meal, soybean oil, soybean and cotton, reflecting the robustness of the prediction method. It can be seen that the DQN-LSTM method is clearly superior for medium-term prediction of agricultural commodities futures prices, and can improve the accuracy of medium-term prediction of most agricultural commodities futures prices.

4.4. Comparative analysis of long-term prediction results

In this paper, a time duration of 365 trading days is used to measure the long-term performance of four prediction methods, namely the DQN-LSTM method, the SVR method, the ARIMA method and the LSTM network. Table.4 reports a comparison of the long-term agricultural commodities futures prices prediction results of the different prediction methods.

In terms of horizontal prediction accuracy, overall, the DQN-LSTM method has the best prediction accuracy, which is followed by the LSTM method, the SVR method, and the ARIMA method with the lowest prediction accuracy. The LSTM method only slightly outperforms the DQN-LSTM method in terms of MAE and $RMSE$ values for cotton. Taking the predictions of the futures prices of corn as an example, the $MAE, MAPE$ and $RMSE$ values of the DQN-LSTM method are 10.648, 4.612 and 13.570, an increase of 97.354, 0.619 and 95.008 respectively compared with the ARIMA method, and an increase of 18.188, 0.135 and 17.088 respectively compared with the SVR method, and an increase of 6.228, 0.005 and 6.04 respectively compared with the LSTM method.

In terms of directional prediction accuracy, overall, the DQN-LSTM method has the highest D_{stat} value, followed by the LSTM and ARIMA methods, while the SVR method has the lowest directional prediction accuracy. The LSTM method only has a slightly higher D_{stat} value than the DQN-LSTM method for soybean meal. Taking the prediction results of futures prices of corn as an example, the D_{stat} value of the DQN-LSTM method is 90.618%, an increase of 19.987% compared with the ARIMA method; 21.164% compared with the SVR method; and 3.281% compared with the LSTM method. The above patterns of variation are largely consistent across futures prices predictions of

Table 4
Long-term agricultural commodities futures prices prediction results

Futures varieties	Method	MAE	$MAPE(\%)$	$RMSE$	$D_{stat}(\%)$
Corn	ARIMA	108.002	5.231	108.578	70.631
	SVR	28.836	4.747	30.658	69.454
	LSTM	16.876	4.617	19.610	87.337
	DQN-LSTM	**10.648**	**4.612**	**13.570**	**90.618**
Soybean meal	ARIMA	63.444	2.299	68.901	76.091
	SVR	36.591	8.071	50.636	78.263
	LSTM	21.133	5.552	28.850	86.178
	DQN-LSTM	**20.307**	**5.481**	**52.820**	**85.896**
Soybean oil	ARIMA	51.159	6.359	68.157	74.674
	SVR	63.721	12.299	67.512	78.769
	LSTM	101.181	2.734	94.521	**86.118**
	DQN-LSTM	**40.079**	**0.865**	**56.050**	87.361
Soybean	ARIMA	116.451	4.945	120.525	67.687
	SVR	57.218	4.667	66.764	79.895
	LSTM	36.409	3.667	55.370	83.128
	DQN-LSTM	**34.309**	**3.188**	**43.980**	**93.659**
Cotton	ARIMA	336.682	8.823	375.900	66.980
	SVR	223.080	14.757	248.199	67.769
	LSTM	**137.763**	6.310	**180.320**	81.846
	DQN-LSTM	138.687	**5.770**	188.800	**84.148**

soybean meal, soybean oil, soybean and cotton, reflecting the robustness of the prediction method. In terms of the overall level of long-term prediction of agricultural commodities futures prices, the DQN-LSTM method outperforms the other three control methods in terms of both horizontal and directional prediction accuracy. It can be seen that the DQN-LSTM method has exponential long-term prediction superiority.

4.5. Comparative analysis of method prediction effects at different durations

To show the prediction effects of different methods more intuitively, Fig.5 further presents the prediction effects of the DQN-LSTM method and other prediction methods on the dynamic trends of agricultural commodities futures prices at different durations. As can be seen from Fig.5, compared with the three single-method prediction methods of ARIMA, SVR and LSTM, all the curves of the DQN-LSTM method prediction results are roughly close to their true values, despite the different degrees of volatility of different agricultural commodities futures prices, further confirming the applicability and effectiveness of the DQN-LSTM method for agricultural commodities futures prices data prediction, and showing that the DQN-LSTM method has excellent generalization ability in prediction of agricultural commodities futures prices.

In addition, Fig.5, in conjunction with the above data specifically, shows that in terms of horizontal prediction accuracy, for example, for futures prices of corn, the DQN-LSTM method improves prediction accuracy by 84.322 (short term), 89.269 (medium term) and 97.354 (long term), respectively, compared with the ARIMA method under the evaluation criteria of MAE. Compared with the SVR method, the prediction accuracy of the DQN-LSTM method improved by 3.865 (short term), 15.954 (medium term) and 18.188 (long term), respectively. Compared with the LSTM method, the prediction accuracy of the DQN-LSTM method improved by 1.718 (short term), 2.71 (medium term) and 6.228 (long term), respectively. In terms of directional prediction accuracy, the D_{stat} value obtained by the DQN-LSTM method ARE 96.234% (short-term), 92.372% (medium-term) and 90.618% (long-term), which are all higher than the other three prediction methods, i.e. the DQN-LSTM method has better directional prediction ability than the other three methods. Compared with the ARIMA method, the DQN-LSTM method increases the directional prediction accuracy by 7.722% (short-term), 24.498% (medium-term) and 19.987% (long-term), respectively. Compared with the SVR method, the increases are 22.502% (short-term), 24.498% (mid-term) and 2.102% (long-term), respectively.

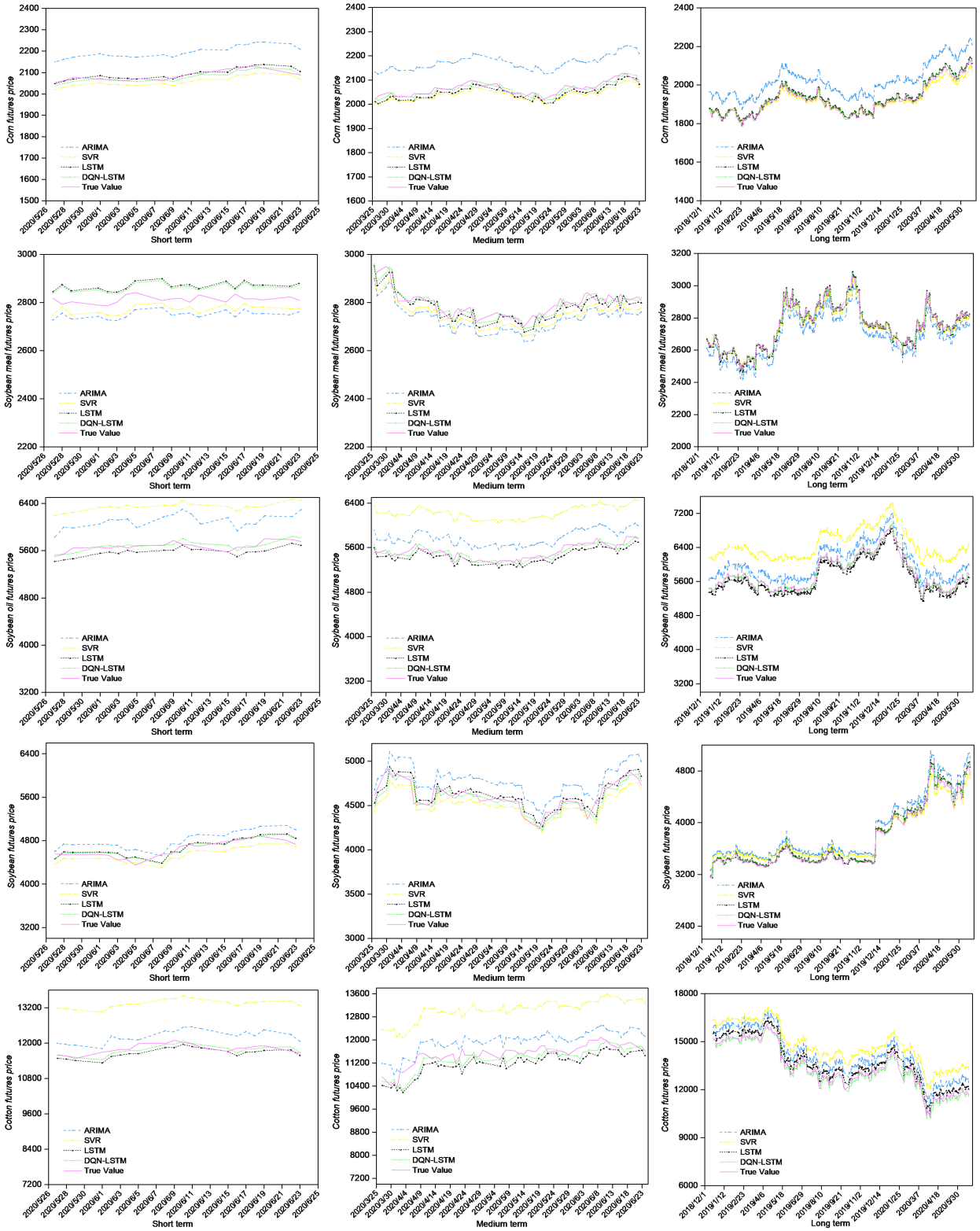

Figure 5: Agricultural commodities futures prices, 9 January 2006 to 24 June 2020.

Compared with the LSTM method, the increases are 2.822% (short term), 2.10% (medium term) and 3.281% (long term) respectively. It can be seen that the longer the prediction time period of each method, the more severe the method failure. DQN-LSTM can effectively control the fluctuation of the prediction error, and thus the longer the prediction period, the more obvious the advantage of prediction accuracy of the DQN-LSTM method is compared with other prediction methods.

5. Conclusion and Discussion

The changes in agricultural commodities futures prices have a bearing on agriculture and even the national economy. To improve the prediction accuracy of agricultural commodities futures prices, this paper proposes a new DQN-LSTM prediction method. The proposed method solves the over-fitting problem of LSTM for time-series data learning and prediction by intelligently deciding the Dropout parameters in the structure of the LSTM method through the DQN algorithm. Such mechanism effectively improves the generalization and robustness in predicting the agricultural commodities futures prices. In order to examine the prediction performance of the method, this paper selects daily data of futures prices of corn, soybean meal, soybean oil, soybean and cotton for testing experiments. The prediction results are compared with those of three methods, ARIMA, SVR and LSTM, for three different duration of short, medium and long term. The results show that the DQN-LSTM method exhibits excellent long-term and short-term prediction accuracy in both horizontal and directional predictions. In particular, the longer the prediction duration, the better the accuracy of the DQN-LSTM method compared with other methods. In addition, the prediction accuracy of the DQN-LSTM method for the futures prices of five agricultural commodities is relatively stable, indicating that the method has the generalization ability in prediction agricultural commodities futures prices.

This paper proactively explores the cutting edge technology of deep learning in financial prediction, and verifies the powerful self-learning capability, excellent generalization ability and high adjustability of the DQN-LSTM method. In view of the highly adjustable nature of the neural network, this paper can be improved in various directions, such as setting the network depth, the number of hidden units and the learning rate in LSTM as parameters to be decided by reinforcement learning, so that the automatic artificial intelligence design of the LSTM network can be realized. A variety of non-homogeneous information can be added as input to the neural network, with data such as wavelet decomposition or principal component analysis and other pre-processing techniques for method optimization being supplemented. Additionally, the the neural network can be further optimized structurally. The application of deep learning technology in financial prediction is only the first step in the development of financial intelligence, and can be furthered in two major themes. First, the introduction of deep neural network cutting-edge methods in the field of financial risk management, taking the advantages of big data to effectively conduct risk identification and risk measurement. Second, the application of deep learning methods in the field of investment can help financial institutions to quickly identify investment opportunities and promote the development of intelligent investment in financial market.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in Wind at https://www.wind.com.cn/. The Wind database is subscription based. Five categories of agricultural futures are selected from the database: soybean No. 1, cotton,soybean meal, soybean oil and corn. The data of futures prices are the daily closing prices and, in all cases, our sample period starts from 9 January 2006 to 24 June 2020.

Conflict of Interest

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled, “Prediction of agricultural commodities futures prices: A DQN-LSTM method”.

Funding

This work is supported by the Fundamental Research Funds for the Central Universities (grant number 202061004).

References

[1] Hang Su, Andrea Mariani, Salih Ertug Ovur, Arianna Menciassi, Giancarlo Ferrigno, and Elena De Momi. Toward teaching by demonstration for robot-assisted minimally invasive surgery. *IEEE Transactions on Automation Science and Engineering*, pages 1–11, 2021.

[2] Hang Su, Yingbai Hu, Hamid Reza Karimi, Alois Knoll, Giancarlo Ferrigno, and Elena De Momi. Improved recurrent neural network-based manipulator control with remote center of motion constraints: Experimental results. *Neural Networks*, 131:291–299, 2020.

[3] Hang Su, Wen Qi, Yingbai Hu, Hamid Reza Karimi, Giancarlo Ferrigno, and Elena De Momi. An incremental learning framework for human-like redundancy optimization of anthropomorphic manipulators. *IEEE Transactions on Industrial Informatics*, 2020.

[4] Jiehao Li, Junzheng Wang, Hui Peng, Longbin Zhang, Yingbai Hu, and Hang Su. Neural fuzzy approximation enhanced autonomous tracking control of the wheel-legged robot under uncertain physical interaction. *Neurocomputing*, 410:342–353, 2020.

[5] DU Wen and H Holly Wang. Price behavior in china's wheat futures market. *China economic review*, 15(2):215–229, 2004.

[6] Yi-Chun Weng and Jyh-Woei Shy. The comparison of the forecasting models of the raw material futures' prices. *Journal of Information and Optimization Sciences*, 31(1):129–145, 2010.

[7] Paulo Cortez, Miguel Rocha, and José Neves. Evolving time series forecasting arma models. *Journal of Heuristics*, 10(4):415–429, 2004.

[8] Alexander Hess, Stefan Spinler, and Matthias Winkenbach. Real-time demand forecasting for an urban delivery platform. *Transportation Research Part E: Logistics and Transportation Review*, 145:102147, 2021.

[9] Ayodele Ariyo Adebiyi, Aderemi Oluyinka Adewumi, and Charles Korede Ayo. Comparison of arima and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*, 2014, 2014.

[10] Chia-Lin Chang, Michael McAleer, and Roengchai Tansuchat. Modelling long memory volatility in agricultural commodity futures returns. *Annals of Financial Economics*, 7(02):1250010, 2012.

[11] Syed Abul Basher and Perry Sadorsky. Hedging emerging market stock prices with oil, gold, vix, and bonds: A comparison between dcc, adcc and go-garch. *Energy Economics*, 54:235–247, 2016.

[12] Zhe Lin. Modelling and forecasting the stock market volatility of sse composite index using garch models. *Future Generation Computer Systems*, 79:960–972, 2018.

[13] M Ugur Gudelek, S Arda Boluk, and A Murat Ozbayoglu. A deep learning based stock trading model with 2-d cnn trend detection. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–8. IEEE, 2017.

[14] Shahriar Yousefi, Ilona Weinreich, and Dominik Reinarz. Wavelet-based prediction of oil prices. *Chaos, Solitons & Fractals*, 25(2):265–275, 2005.

[15] Hesam Dehghani and Dejan Bogdanovic. Copper price estimation using bat algorithm. *Resources Policy*, 55:55–61, 2018.

[16] Chang Liu, Zhenhua Hu, Yan Li, and Shaojun Liu. Forecasting copper prices by decision tree learning. *Resources Policy*, 52:427–434, 2017.

[17] Julio López, Sebastián Maldonado, and Miguel Carrasco. Double regularization methods for robust feature selection and svm classification via dc programming. *Information Sciences*, 429:377–389, 2018.

[18] Zeynep Cipiloglu Yildiz and Selim Baha Yildiz. A portfolio construction framework using lstm-based stock markets forecasting. *International Journal of Finance & Economics*, 2020.

[19] Zhigang Jin, Yang Yang, and Yuhong Liu. Stock closing price prediction based on sentiment analysis and lstm. *Neural Computing and Applications*, pages 1–17, 2019.

[20] Ha Young Kim and Chang Hyun Won. Forecasting the volatility of stock price index: A hybrid model integrating lstm with multiple garch-type models. *Expert Systems with Applications*, 103:25–37, 2018.

[21] Yishun Liu, Chunhua Yang, Keke Huang, and Weihua Gui. Non-ferrous metals price forecasting based on variational mode decomposition and lstm network. *Knowledge-Based Systems*, 188:105006, 2020.

[22] Vu Pham, Théodore Bluche, Christopher Kermorvant, and Jérôme Louradour. Dropout improves recurrent neural networks for handwriting recognition. In *2014 14th international conference on frontiers in handwriting recognition*, pages 285–290. IEEE, 2014.

[23] Gaith Rjoub, Jamal Bentahar, Omar Abdel Wahab, and Ahmed Saleh Bataineh. Deep and reinforcement learning for automated task scheduling in large-scale cloud computing systems. *Concurrency and Computation: Practice and Experience*, page e5919, 2020.

[24] Byungjin Ko, Seunghan Ryu, Byungkyu Brian Park, and Sang Hyuk Son. Speed harmonisation and merge control using connected automated vehicles on a highway lane closure: a reinforcement learning approach. *IET Intelligent Transport Systems*, 14(8):947–957, 2020.

[25] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *nature*, 518(7540):529–533, 2015.

[26] Leemon C Baird III. Reinforcement learning through gradient descent. Technical report, CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER SCIENCE, 1999.

[27] Long-Ji Lin. Self-improving reactive agents based on reinforcement learning, planning and teaching. *Machine learning*, 8(3-4):293–321, 1992.

[28] Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. Prioritized experience replay. *arXiv preprint arXiv:1511.05952*, 2015.