Report on articles on stock market forecasting with stock market historical data as input

Table 1: number of articles reviewed

conferences	6
journals	12

Table 2: purpose and type of innovation of papers

Innovations	purpose
algorithm	To overcome 1-step-ahead delay of prediction
algorithm	Use PSO to optimize SVM
Algorithm dataset	Capture multi-frequencies using for novel state frequency memory recurrent network
Algorithm	Combine multi-scale property and temporal dependency of time series by proposing a multi-scale temporal dependent recurrent convolutional neural network
algorithm	To Improve accuracy of prediction propose two hybrid model
algorithm	Propose a hybrid algorithm
Data set Well-defined concept	assign weights to data according to their temporal nearness towards the data to be predicted formally trend definitions given by referencing financial theories and best practices.
algorithm	Propose a hybrid model

algorithm	Propose time effective neural network with principal component analysis model
Algorithm Dataset	propose a neural network layer architecture that incorporates the idea of bilinear projection as well as an attention mechanism that enables the layer to detect and focus on crucial temporal information.
algorithm	Propose a close-loop ML approach that allow user to interact directly with propose algorithm
dataset	two approaches for input. One approach for input data involves computation of ten technical parameters using stock trading data (open, high, low & close prices) while the second approach focuses on representing these technical parameters as trend deterministic data
Setting parameters	Select train dataset and test dataset from different company data
Feature extraction	The task of feature learning can be done more effectively
Algorithm output	Multiple associated output

Table 3: more details of papers

	Data set	Algorithm	Metrics	attributes	Target (want to predict)
1 [1]	Apple inc. from Google finance cited by: 17	Random tree, multilayer perceptron	Correlation coefficient, Mean absolute error, Root mean squared error, Relative absolute error	Date, open, high, low, volume	Closing price
2 [2]	13 company From finance. Yahoo Cited by: 69	PSO+LS_SVM Compared with: NN, LS_SVM	Predicted curve(value), error value	six inputs vectors represent the historical data and derived technical indicators: *Relative Strength Index*, Money Flow Index, *Exponential Moving Average*, *Stochastic Oscillator*, *Moving Average Convergence/Divergence*,	One output vector represents next price (value)
3 [3]	50 stocks among 10 sectors From finance. Yahoo Cited by: 103	State Frequency Memory (SFM) recurrent network Compared with: Autoregressive model and the conventional LSTM	Square error	Stock prices	Future price trend

4 [4]	Three high-frequency stock index datasets are collected from the Chinese stock market Cited by: 2	Multi-Scale Temporal Dependent Recurrent Convolutional Neural Network (MSTD-RCNN) for FTC	Precision, recall	In the proposed method, the MS features are simultaneously extracted by convolutional units to precisely describe the state of the financial market.	trend classification and simulated trading
5 [5]	KOSPI 200 Cited by: 16	MLP, 1D-CNN, LSTM, Attention networks, Weighted attention network	Hit ratio, prediction$_i$	Change ratio and various indexes as input	lookback days as p trading days as target
6 [6]	daily, weekly, and monthly time series of some stock data from Finance. Yahoo cited by: 9	Comparison between ARIMA, LSTM and BILSTM	RMSE, %reduction	It is a comparison based on RMSA, change% (percentage of reduction in RMSA	It is a comparison based on RMSA, change% (percentage of reduction in RMSA
7 [7]	daily closing price of the S&P500, HSI, DAX and SSE are selected as the original data from Finance.Yahoo cited by: 87	EMD-LSTM and CEEMDAN-LSTM	MAE, RMSE, MAPE	input of the *t*th sample is (xt, xt+1, . . . , xt+W),	the corresponding label is xt+W+1.
8 [8]	Shanghai Stock Exchange Composite (SSEC) index --- National Association of Securities Dealers Automated Quotations (NASDAQ) index---- the Standard & Poor's 500 Composite Stock Price Index(S&P500) From yahoo.Finance Cited by: 40	EMD2FNN Compared with: neural network (NN) model, the factorization machine based neural network (FNN)	Statistical metrics: Average annual return(AAR), maximum drawdown(MD), sharp ration(SR), AAR/MMD// accuracy measure: MAE, RMSE, MAPE	SSEC, NASDAQ, S&P500 Actually Date, open, high, low, volume values	daily closing prices, NASDAQ, S&P500 trend

		model, the empirical mode decomposition based neural network (EMD2NN) model and the wavelet de-noising-back propagation neural network			
9 [9]	Fifteen years' daily transaction data of five main stock indexes in China: SSE Composite Index (000001.SS), SSE 50 (000016.SS), CSI 500(000905.SS), CSI 300(000300.SS) and SZSE Composite Index(399001.SZ) From Yahoo. Finance Cited by:22	time-weighted LSTM compared with: SVM, random forest, adaboost	Accuracy	Indexes introduces in data set section Actually Date, open, high, low, volume values	Stock market trends (this paper has defined a definition for trend)
10 [10]	Standard & Poor's 500 (S&P 500 Index), Shanghai Composite Index in China, International Business Machine (IBM) from New York Stock Exchange (NYSE), Microsoft Corporation (MSFT) from National Association of Securities Dealers Automated Quotation (NASDAQ), Ping An Insurance Company of China (PAICC) From Yahoo.Finance Cited by:42	GAN (Generative Adversarial Network) Compared with: SVM, ANN, LSTM		Date, open, high, low, volume values of S&P 500 Index, NYSE, MSFT, NASDAQ, PAICC	Closing price

11 [11]	the daily data from Shanghai Stock Exchange Composite Index (SSE), Hong Kong Hang Seng 300 Index (HS300), Dow Jones Industrial Average Index (DJIA), Standard & Poor's 500 Index (S&P500) cited by: 89	*PCA-STNN Compared with:* BPNN, STNN, PCA-BPNN, SVM	MAE, RMSE, MAPE, MAPE(100)	Date, open, high, low, volume values of Shanghai Stock Exchange Composite Index (SSE), Hong Kong Hang Seng 300 Index (HS300), Dow Jones Industrial Average Index (DJIA), Standard & Poor's 500 Index (S&P500)	Returns and relative errors(proposed formula)
12 [12]	LOB Cited by: 65	*Temporal Attention-Augmented Bilinear Layer*	Accuracy, precision, recall, F1-score	Feature extraction using neural network	Average attention, accuracy, recall, computational time
13 [13]	The dataset refers to a private dataset from ATS company and it represents a recent year's operations in several markets. Cited by:- (2020)	Decision Support System for Outcome Analysis, Decision trees Compared with: Mlp, XGBoost	Classification metrics: Accuracy, recall, F1, confusion matrix Regression metrics: R^2 score, mean absolute error, pearson correlation coefficient, spearman correlation coefficient	feature set of 278 variables	TGT_LASTORDSTATUS (binary target) TGT_EXECQTYPERC (continuous target) TGT_LIMPRICE_DIFF_PERC(continues target(continuous target) TGT_COMMISSION_WEIGHT_PERC (continuous target)
14 [14]	data of total two stock price indices (CNX Nifty, S&P 239 BSE Sensex) and two stocks (Reliance Industries, Infosys Ltd.) cited by: 553	ANN, SVM, RANDOM FOREST, naïve bayes	Precision, recal, accuracy, f-meature	the actual time series(continues) i.e. ten technical indicators as input to models, (discrete) trend prediction data i.e. ten technical indicators' opinion on stock movement	Up/down movement

15 [15]	close price of daily data of iShares MSCI United Kingdom from January 2015 to June 2018 cited by: 9	Artificial neural network, SVM, random forest, deep learning, long- short term memory	MAE, MSE, RMSE, price	Close price	Closing price index
16 [16]	day-wise closing price of two different stock markets, National Stock Exchange (NSE) of India and New cited by: 93	MLP, RNN, LSTM, CNN	Mean absolute percentage error,	day-wise closing price of each stock because day wise stock price	Share price
17 [17]	market data of CSI 300 cited by: 87	Multi-filter neural networks (specifically for feature selection on financial time series samples and price movement prediction)	Accuracy, Wilcoxon signed rank sum, total return(R), annual return(AR), daily winning rate(DWR),sharp ratio, annual volatility(V),	six indicators *xt* = (*ot*, *ct*, *ht*, *lt*, *vt*, *at*) are obtained at each time step in 1-minute frequency	extreme market prediction and signal-based trading simulation
18 [18]	Shanghai composite index 000001, two stocks of PetroChina (stock code 601857) on Shanghai stock exchange and ZTE (stock code 000063) on Shenzhen stock exchange Cited by: 12	Associated deep recurrent neural network, Compared with: LSTM network, LSTM-based deep recurrent neural network	MAE, MSE	Open price, close price, lowest price, highest price, volumes, money, change	total loss and three sub-losses (opening price loss, lowest price loss, highest price loss).

Concepts:

Stock vs share:

Of the two, "stocks" is the more general, generic term. It is often used to describe a slice of ownership of one or more companies. In contrast, in common parlance, "shares" has a more specific meaning: It often refers to the ownership of a particular company.

Stock market Time series:

Stock market is having a highly fluctuating and non-linear time series data. A time series is a set of data measured over time to acquire the status of some activity

Stock market index

A stock market index shows how investors feel an economy is faring. An index collects data from a variety of companies across industries. Together, that data forms a picture that helps investors compare current price levels with past prices to calculate market performance.

Some indexes focus on a smaller subset of the market. For example, the Nasdaq index closely tracks the technology sector. So if you want to know how technology companies are performing, you'd want to look at the Nasdaq stock index.

Indexes also vary in size, with some tracking just a handful of stocks and others looking at thousands. Each index serves a unique purpose because different investors are interested in different sectors.

S&P 500:

A barometer of the overall stock market's performance that contains 500 companies, weighted by market cap, from across different sectors.

The S&P 500 uses a market capitalization weighting method, giving a higher percentage allocation to companies with the largest market capitalizations.

Nasdaq:

This index includes the roughly 3,000 companies that are part of the Nasdaq stock exchange and is predominately focused on technology.

SSEC:

The SSE Composite Index also known as SSE Index is a stock market index of all stocks (A shares and B shares) that are traded at the Shanghai Stock Exchange.

Some economic metrics

Average annual return:

In its simplest terms, the average annual return (AAR) measures the money made or lost by a mutual fund over a given period.

When you are selecting a mutual fund, the average annual return is a helpful guide for measuring a fund's long-term performance. However, investors should also look at a fund's yearly performance to fully appreciate the consistency of its annual total returns.

Sharpe ratio:

is used to help investors understand the return of an investment compared to its risk. he Sharpe ratio can be used to evaluate the total performance of an aggregate investment portfolio or the performance of an individual stock.

The Sharpe ratio indicates how well an equity investment performs in comparison to the rate of return on a risk-free investment.

Generally, the greater the value of the Sharpe ratio, the more attractive the risk-adjusted return.

Maximum drawdown:

A maximum drawdown (MDD) is the maximum observed loss from a peak to a trough of a portfolio, before a new peak is attained. Maximum drawdown is an indicator of downside risk over a specified time period.

Maximum drawdown is a specific measure of drawdown that looks for the greatest movement from a high point to a low point, before a new peak is achieved. However, it's important to note that it only measures the size of the largest loss, without taking into consideration the frequency of large losses.

1.

Stock Market Data Prediction Using Machine Learning Techniques:

We have experimented with stock market data of the Apple Inc. using random trees and multilayer perceptron algorithms to perform the predictions of closing prices.

For the purpose of this paper, we have used the historical data of opening price, closing price, highest price, lowest price, and volume traded of Apple Inc. stocks which has been gathered using Google Finance.

The present work has used the WEKA software to execute the machine learning algorithms in our stock market data sets.

As previously mentioned, we considered six attributes for the algorithms: (1) Date, (2) Open, (3) High, (4) Low, (5) Close, and (6) Volume.

- Date: Market session date
- Open: The opening price for the market session
- High: The highest price reached during the market session
- Low: The lowest price reached during the market session
- Close: The closing price for the market session
- Volume: The total number of trades performed on the stock during the market session

The attribute Close is the one to be forecasted and compared to its real data, so that the accuracy of the algorithms can be tested and measured for errors and over fitting.

A deep increasing–decreasing-linear neural network for financial time series prediction:

Several neural network models have been proposed in the literature to predict the future behavior of financial time series. However, an intrinsic limitation arises from this particular prediction task with modeling via neural networks, since the prediction, when sampled in daily frequency, have 1-step-ahead delay with respect to real time series observations. In order to overcome such drawback, we present a deep increasing–decreasing-linear neural network (wherein each layer is composed of a set of increasing– decreasing-linear processing units) to predict the behavior of financial time series. In addition, we present a learning process to train the proposed model using a descending gradient-based approach

Cont….

2.

A Machine Learning Model for Stock Market Prediction

This paper proposes a machine learning model to predict stock market price. The proposed algorithm integrates Particle swarm optimization (PSO) and least square support vector machine (LS-SVM). The PSO algorithm is employed to optimize LS-SVM to predict the daily stock prices. Thirteen benchmark financials datasets and compared with artificial neural network with Levenberg-Marquardt (LM) algorithm.

The proposed model is based on the study of historical data, technical indicators and optimizing LS-SVM with PSO algorithm to be used in the prediction of daily stock prices.

Levenberg-Marquardt (LM) algorithm is used as a benchmark for comparison with LS-SVM and LS-SVM-PSO models. The proposed model architecture contains six inputs vectors represent the historical data and derived technical indicators and one output represents next price.

3

Stock Price Prediction via Discovering Multi-Frequency Trading Patterns:

We propose a novel State Frequency Memory (SFM) recurrent network to capture the multi-frequency trading patterns from past market data to make long- and short-term predictions over time.

Stock investors attempt to discover latent trading patterns in stock market to forecast the future price trends for seeking profit-maximization strategies.

These patterns with various frequencies provide useful hints on the future trend, which is confirmed by the experiments we conducted.

Autoregressive (AR) model is a classical method that is widely used for time-series forecasting. It predicts the future market price from the prices of several past steps. Specifically, its prediction is made as a linear combination of the past prices subject to a Gaussian noise term

4.

Multi-Scale RCNN Model for Financial Time-series Classification

classifying sequence of financial time-series data into different categories by their movement direction. Actually, financial time-series classification

In the proposed method, the MS features are simultaneously extracted by convolutional units to precisely describe the state of the financial market.

Three high-frequency stock index datasets are collected from the Chinese stock market.

The proposed method is evaluated on three financial time-series datasets which source from the Chinese stock market. Extensive experimental results indicate that our model achieves the state-of-the-art performance in trend classification and simulated trading

5.

FINANCIAL SERIES PREDICTION USING ATTENTION LST

In this paper, we will compare various deep learning models, such as multilayer perceptron (MLP), one-dimensional convolutional neural networks (1D CNN), stacked long short-term memory (stacked LSTM), attention networks, and weighted attention networks for financial time series prediction.

Technical analysis is a traditional method that uses historical stock prices and trading volumes to determine the trends of future stock movements. This method is based on supply and demand in financial markets. It facilitates the easy construction of models as it only considers numerical variables.

Fundamental analysis predicts stock prices by using intrinsic values. When using this method, stock values are determined by financial news, market sentiments and economic factors; investors estimate the profits of firms and evaluate whether they are suitable for investment.

The KOSPI 200 index is a weighted combination of the 200 most traded securities in the Korean stock exchange. we take the change ratios of the KOSPI 200, and various indexes such as the currency, commodities, and global indexes, which are closely related with Korean financial markets in terms of fundamental analysis, as input data.

6.

A Comparative Analysis of Forecasting Financial Time Series Using ARIMA, LSTM, and BiLSTM

This paper reports a behavioral analysis and comparison of BiLSTM and LSTM models. The objective is to explore to what extend additional layers of training of data would be beneficial to tune the involved parameters.

previously collected data, in which daily, weekly, and monthly time series of some stock data for the period of Jan 1985 to Aug 2018 were extracted from the Yahoo finance Website1.

7.

Financial time series forecasting model based on CEEMDAN and LSTM

In order to improve the accuracy of the stock market prices forecasting, two hybrid forecasting models are proposed in this paper.

The financial time series is a kind of non-linear and non-stationary random signal, which can be decomposed into several intrinsic mode functions of different time scales by the original EMD and the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN).

Because of the supervised learning method, the input time series is converted to one sample with one label. The length of the training set time series is T, the window size is W, the input of the tth sample is (xt, xt+1, . . . , xt+W), and the corresponding label is xt+W+1.

8.

EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction

The goal of this paper is to introduce a new hybrid, end to end approach containing two stages, the empirical mode decomposition and factorization machine based neural network to predict the stock market trend.

to predict the daily closing prices from the Shanghai Stock Exchange Composite (SSEC) index, the National Association of Securities Dealers Automated Quotations (NASDAQ) index and the Standard & Poor's 500 Composite Stock Price Index (S&P 500), which respectively exhibit oscillatory, upward and downward patterns.

In addition, Table 3 lists all statistical metrics used for examining the trading performance in this study. The Sharpe Ratio (SR) measures the risk-adjusted return. AAR/MD is slightly modified from the Calmar ratio and is calculated as the Average Annual Return (AAR) derived by the Maximum Drawdown (MD) for the whole duration considered. For the MD, defined as the largest accumulated percentage loss due to a sequence of drops over a investment horizon, a lower output means a better performance. For the Average Annual Return, Sharpe 285 Ratio and AAR/MD, a higher output is better.

9.

Time-weighted LSTM Model with Redefined Labeling for Stock Trend Prediction

In this paper, we attempted to exploit the time attribute of stock data to improve prediction accuracy. Firstly, instead of treating data indiscriminately, we used time weight function to carefully assign *weights* to data according to their temporal nearness towards the data to be predicted. Secondly, the stock trend definitions were formally given by referencing financial theories and best practices. Lastly, Long Short-Term Memory (LSTM) network was customized to discover the underlying temporal dependencies in data.

we put forward a way to quantify trend. It provide a trend labeling algorithm.

10.

Stock Market Prediction Based on Generative Adversarial Network

In this paper, we propose a novel architecture of Generative Adversarial Network (GAN) with the Multi-Layer Perceptron (MLP) as the discriminator and the Long Short-Term Memory (LSTM) as the generator for forecasting the closing price of stocks.

We choose the daily data on S&P 500 Index and several stocks in a wide range of trading days and try to predict the daily closing price

GAN is a new framework which trains two models like a zero-sum game. The generator of our model is designed by LSTM with its strong ability in processing time series data.

We choose the daily data in the last 20 years with 7 financial factors to predict the future closing price. The 7 factors of the stock data in one day are High Price, Low Price, Open Price, Close Price, Volume, Turnover Rate and Ma5 (the average of closing price in past 5 days).

Standard & Poor's 500 (S&P 500 Index), Shanghai Composite Index in China, International Business Machine (IBM) from New York Stock Exchange (NYSE), Microsoft Corporation (MSFT) from National Association of Securities Dealers Automated Quotation (NASDAQ), Ping An Insurance Company of Chine.

11.

Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks

A stochastic time effective function neural network (STNN) with principal component analysis (PCA) developed for financial time series prediction is presented in the present work. In the training modeling, we first use the approach of PCA to extract the principal components from the input data, then integrate the STNN model to perform the financial price series prediction.

we consider the statistical behaviors of price returns and relative errors of the PCA-STNN forecasting results. Let $S(t)$ ($t = 1, 2, \cdot\cdot\cdot$) denote the price sequences of SSE, HS300, S&P500 and DJIA at time t, then the formula of stock logarithmic return and relative error are given as follows

12.

Temporal Attention-Augmented Bilinear Network for Financial Time-Series Data Analysis

we propose a neural network layer architecture that incorporates the idea of bilinear projection as well as an attention mechanism that enables the layer to detect and focus on crucial temporal information.

In stock markets, traders buy and sell stocks through an order-driven system that aggregates all out-standing limit orders in LOB. A limit order is a type of order to buy or sell a certain amount of a security at a specified price or better.

We evaluate our proposed architecture with the task of predicting the future movement of midprice given the past bid and ask prices with respective volumes. We use the publicly available data set provided in [65], known as FI-2010 data set.

The data were collected from five different Finnish stocks in NASDAQ Nordic coming from different industrial sectors.

13.

Machine Learning in Capital Markets: Decision Support System for Outcome Analysis

The aim of this work is the proposal of a closed-loop ML approach based on decision tree (DT) model to perform outcome analysis on financial trading data. The overall approach is integrated in a Decision Support System for Outcome Analysis (DSS-OA). Taking into account the model complexity, the DT algorithm enables to generate explanations that allow the user to understand (i) how this outcome is reached (decision rules) and (ii) the most discriminative outcome predictors (feature importance). The closed-loop approach allows the users to interact directly with the proposed DSS-OA by retraining the algorithm with the goal to a fine-grained outcome analysis.

The dataset refers to a private dataset from ATS company andit represents a recent year’s operations in several markets.

Trading History data are anonymous in terms of customers and instrument and their use, detention and conservation are regulated by an agreement between ATS company, Università Politecnica delle Marche and data owners. The dataset refers to a private dataset from ATS company and it represents a recent year's operations in several markets. The dataset was used because (i) it is an example of how trading operations data can be made available from a trading platform and (ii) it contains a good variety of operations on different markets/asset types over a consistent period.

According to the ATS company, four target variables have been identified on the *Trading History* dataset. Our goal is to perform individually the outcome analysis of these 4 tasks.

14.

Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques

This paper addresses problem of predicting direction of movement of stock and stock price index for 23 Indian stock markets. The study compares four prediction models, Artificial Neural Network (ANN), support vector machine (SVM), random forest and Naive-Bayes with two approaches for input to these models. The first approach for input data involves computation of ten technical parameters using stock trading data (open, high, low & close prices) while the second approach focuses on representing these technical parameters as trend deterministic data.

Two approaches for the representation of the input data are employed in this study. The first approach uses continuous value representation, i.e., the actual time series while the second one uses trend deterministic representation (which is discrete in nature) for the inputs. Both the representations are discussed here.

15.

Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms

This study seeks to evaluate the prediction power of machine- learning models in a stock market. The data used in this study include the daily close price data of iShares MSCI United Kingdom exchange-traded fund from January 2015 to June 2018. The prediction process is done through four models of machine-learning algorithms.

16.

NSE Stock Market Prediction Using Deep-Learning Models

The various algorithms used for forecasting can be categorized into linear (AR, MA, ARIMA, ARMA) and non-linear models (ARCH, GARCH, Neural Network).In this paper, we are using four types of deep learning architectures i.e. Multilayer Perceptron (MLP), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) for predicting the stock price of a company based on the historical prices available. Here we are using day-wise closing price of two different stock markets, National Stock Exchange (NSE) of India and New York Stock Exchange (NYSE). The network was trained with the stock price of a single company from NSE and predicted for five different companies from both NSE and NYSE.

The only problem with linear models like AR, ARMA, ARIMA are, that they work only for a particular time series data, i.e. the model identified for a particular company won't perform well for another.

Dataset is taken from highly traded stocks of three different sectors which are Automobile, Banking and IT sectors from NSE. Each contains information like stock symbol , stock series, stock date and previous closing, opening ,high ,low ,last , closing and average prices, total traded quantity, turnover and no : of trades.

From these datasets, we extract only the day-wise closing price of each stock because day wise stock price is preferred since investors make decision on buying which stock or forfeiting which stock based on the closing price of the market

17.

Deep Learning-Based Feature Engineering for Stock Price Movement Prediction

With the growth in deep learning, the task of feature learning can be performed more effectively by purposely designed network. In this paper, we propose a novel end-to-end model named multi-filters neural network (MFNN) specifically for feature extraction on financial time series samples and price movement prediction task. Both convolutional and recurrent neurons are integrated to build the multi-filters structure, so that the information from different feature spaces and market views can be obtained. We apply our MFNN for extreme market prediction and signal-based trading simulation tasks on Chinese stock market index CSI 300.

Experimental results show that our network outperforms traditional machine learning models, statistical models, and single-structure(convolutional, recurrent, and LSTM) networks in terms of the accuracy, profitability, and stability.

In the task of stock price prediction, we define *x*1, *x*2, *x*3...*xt*, ... as indicators sequences, in our works, six indicators *xt* = (*ot*, *ct*, *ht*, *lt*, *vt*, *at*) are obtained at each time step in 1-minute frequency, including open price(stock price at the start of each time step), close price(stock price at the end of each time step), highest price(the highest price among each time step), lowest price(the lowest price among each time step), volume(the number of shares traded in a security during each time step), amount(the amount of money traded in a security during each time step). At each time step, features of each sample is composed of the six indicator sequences over past 120 minutes, which can be de noted as *Xt* = (*xT t*-119, *xT t*-118, ..., *xT t*)*T* = (*Ot*, *Ct*, *Ht*, *Lt*, *Vt*, *At*). Features of each sample are then scaled(refer to paper).

Good define the economic metrics.

Market data at each time step within a sample partly result from historical behaviors and contain different amounts of information in different sub-periods. So to capture these information contained in different sub-periods, the right part of multi-filters module extracts features by recurrent neurons on raw inputs, which can be formulated as:

when *tforward* = 5 and θ = 0.1, samples whose prices rise for more than 0.26% in future 5 minutes are labeled +1, and whose prices fall for more than -0.25% in future 5 minutes are labeled -1, and the other samples are labeled 0.

18.

Study on the prediction of stock price based on the associated network model of LSTM

In order to predict multiple values in one model, it need to design a model which can handle multiple inputs and produces multiple associated output values at the same time. For this purpose, it is proposed an associated deep recurrent neural network model with multiple inputs and multiple outputs based on long short-term memory network.

The associated network model can predict the opening price, the lowest price and the highest price of a stock simultaneously.

Each data set has seven technical parameters. It is used these technical parameters as basic input attributes, and the OP, LP and HP of the next day as output values of the model.

The output of the Associated Net is the total loss and his three sub-losses (opening price loss, lowest price loss, highest price loss).

References

[1] M. Hernández-Álvarez, E. A. T. Hernández, and S. G. Yoo, “Stock Market Data Prediction Using Machine Learning Techniques,” in *International Conference on Information Technology & Systems*, 2019, pp. 539–547.

[2] O. Hegazy, O. S. Soliman, and M. A. Salam, “A machine learning model for stock market prediction,” *arXiv Prepr. arXiv1402.7351*, 2014.

[3] L. Zhang, C. Aggarwal, and G.-J. Qi, “Stock price prediction via discovering multi-frequency trading patterns,” in *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, 2017, pp. 2141–2149.

[4] L. Guang, W. Xiaojie, and L. Ruifan, “Multi-Scale RCNN Model for Financial Time-series Classification,” *arXiv Prepr. arXiv1911.09359*, 2019.

[5] S. Kim and M. Kang, “Financial series prediction using Attention LSTM,” *arXiv Prepr. arXiv1902.10877*, 2019.

[6] S. Siami-Namini, N. Tavakoli, and A. S. Namin, “A comparative analysis of forecasting financial time series using arima, lstm, and bilstm,” *arXiv Prepr. arXiv1911.09512*, 2019.

[7] J. Cao, Z. Li, and J. Li, “Financial time series forecasting model based on CEEMDAN and LSTM,” *Phys. A Stat. Mech. its Appl.*, vol. 519, pp. 127–139, 2019.

[8] F. Zhou, H. Zhou, Z. Yang, and L. Yang, “EMD2FNN: A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction,” *Expert Syst. Appl.*, vol. 115, pp. 136–151, 2019.

[9] Z. Zhao, R. Rao, S. Tu, and J. Shi, “Time-weighted LSTM model with redefined labeling for stock trend prediction,” in *2017 ieee 29th international conference on tools with artificial intelligence (ictai)*, 2017, pp. 1210–1217.

[10] K. Zhang, G. Zhong, J. Dong, S. Wang, and Y. Wang, “Stock market prediction based on generative adversarial network,” *Procedia Comput. Sci.*, vol. 147, pp. 400–406, 2019.

[11] J. Wang and J. Wang, “Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks,”

Neurocomputing, vol. 156, pp. 68–78, 2015.

[12] D. T. Tran, A. Iosifidis, J. Kanniainen, and M. Gabbouj, "Temporal attention-augmented bilinear network for financial time-series data analysis," *IEEE Trans. neural networks Learn. Syst.*, vol. 30, no. 5, pp. 1407–1418, 2018.

[13] R. Rosati, L. Romeo, C. A. Goday, T. Menga, and E. Frontoni, "Machine Learning in Capital Markets: Decision Support System for Outcome Analysis," *IEEE Access*, vol. 8, pp. 109080–109091, 2020.

[14] 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. 42, no. 1, pp. 259–268, 2015.

[15] M. Nikou, G. Mansourfar, and J. Bagherzadeh, "Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms," *Intell. Syst. Accounting, Financ. Manag.*, vol. 26, no. 4, pp. 164–174, 2019.

[16] M. Hiransha, E. A. Gopalakrishnan, V. K. Menon, and K. P. Soman, "NSE stock market prediction using deep-learning models," *Procedia Comput. Sci.*, vol. 132, pp. 1351–1362, 2018.

[17] W. Long, Z. Lu, and L. Cui, "Deep learning-based feature engineering for stock price movement prediction," *Knowledge-Based Syst.*, vol. 164, pp. 163–173, 2019.

[18] G. Ding and L. Qin, "Study on the prediction of stock price based on the associated network model of LSTM," *Int. J. Mach. Learn. Cybern.*, vol. 11, no. 6, pp. 1307–1317, 2020.