Paper Type: Original Article

The analysis of Alibaba Group's future stock price trends and overall status based on the Random Forest mode

Qiushuang Wang

School of Finance, JiLin University of Finance and Economics, Changchun City, Jilin Province

Abstract

With the rapid development of technology and the ever-changing global landscape, technologies such as cloud computing, artificial intelligence, big data, and blockchain are being widely adopted. E-commerce is rising rapidly, and in today's era, Alibaba Group's overall development trend has gradually heated up after the "1+6+N" business model reform. For investors, stock trend prediction is an important task, but predicting stock price movements is challenging due to the influence of various factors. The Random Forest model is an important technology in the field of artificial intelligence, with significant effectiveness in simulating the specific characteristics of research objects, handling nonlinear problems, and studying non-stationary data. This paper aims to predict Alibaba Group's future development prospects and stock price fluctuations after a series of reforms by constructing a Random Forest model. A multivariate linear regression model is used for numerical analysis, while the model's predictions are systematically scored under the MAE, MSE, MAPE, and RMSE indicators. The accuracy of the Random Forest model in predicting Alibaba Group's future stock price trends is as high as 82.97%.

Keywords: Alibaba Group; Stock Price; Random Forest Model; Multivariate Linear Regression Model; "1+6+N" .

1. Company profile

Alibaba Group Holdings Ltd., now known as Alibaba (China) Co., LTD., Alibaba Group for short, was founded in 1999 by Jack Ma and 18 people in Hangzhou, Zhejiang province. Alibaba was listed on the New York Stock Exchange in September 2014 with the ticker symbol "BABA". In November 2019, it was listed on the Main board of the Board of the Hong Kong Stock Exchange with the ticker symbol "9988". Since its development in 1999, Alibaba has accumulated and formed a complete business ecosystem. Its data and traffic sharing are the core, cloud services and marketing are the foundation, and Alipay is the starting point of data integration. Alibaba's business covers e-commerce, Ant Finance, logistics, cloud services and other Internet-related services, including 11 products such as Taobao, Tmall and Alipay.

Table 1: Alibaba's business development and the trend of its employee size

fiscal year	Staff size	Business development
2021	251462	
2020	117600	Holding Gao Xin retail
2019	101958	Acquisition of koala
2018	66421	Is the acquisition hungry
2017	50092	Tmall Taobao overseas
2016	36450	Hippopotamus amphibius
2015	34985	Nail, Ali health
2014	22072	Ant Group, Ali Travel (Flying Pig)
2013	20674	Cainiao network
2012	21930	2009 Ali Cloud, 2010 Juhuasuan, AliExpress
2008	12000	2004 Alipay, 2007 Ali mother, Ali Software, 2008 Tmall
1999	18	1999 launched "1688", 2003 Taobao

As of June 30,2024, the total number of employees of Alibaba Holdings Limited is included. With the increasing efforts of the group, the group is gradually expanded, and the company's income is also increasing.

Table 2 Details of Alibaba's revenue and net profit from 2024 to 2017

year	revenue	net earnings
2024	479,739,000,000	67,569,000,000
2023	719,294,000,000	70,413,000,000
2022	660,487,000,000	43,577,000,000

2021	649,010,000,000	65,436,000,000
2020	529,894,000,000	150,938,000,000
2019	395,397,000,000	140,002,000,000
2018	283,346,000,000	56,855,000,000
2017	158,273,000,000	41,226,000,000

2. Development history

Alibaba Group Holdings Limited (hereinafter referred to as "Alibaba Group") was founded by 18 people in 1999 including Jack Ma in Hangzhou, China. Since its birth, Alibaba Group's mission has never changed, and all the founders have been convinced from the very beginning that the Internet can create a fair environment. In response to the changes in the market environment and continuous adjustment and improvement. In 2003, Taobao was officially established. In 2005, Alipay was separately spun off. In 2007, the eight-year-old B2B business was successfully landed in the Stock Exchange, raising 1.7 billion US dollars, which also marked the beginning of internationalization. At this time, Alibaba's business model has been expanded from the original B2B fast online trading platform to B2C, C2C and other diversified business forms, and Alibaba's own ecological chain has been gradually improved. Around 2010, Alibaba launched the business group system to provide more targeted management and support for each business segment. In 2011, Taobao experienced the "one demolition and three divisions". In 2012, Alibaba established seven business divisions, and then expanded to 25 business divisions in the following year. In 2014, the popularization and rapid development of mobile communication equipment provided a new opportunity for enterprises. Alibaba responded quickly and shifted its business focus from PC terminal to mobile terminal.

In 2015, Alibaba began to consider the needs of business model reform, and promulgated the strategy of "large middle desk, small front desk", to relieve the pressure for the front-line business departments. Around 2020, Alibaba made a series of executive appointments and adjustments, reformed the corporate governance structure from top to bottom, and then successively promoted the management responsibility system and sector governance, in order to realize the "one center" Alibaba. Until March 2023, Alibaba made the most formal organizational structure adjustment, which is also the "1 + 6 + N" mode: 1 core investment holding group, 6 major subsidiary business groups, and N smaller and diversified business companies.

stage	a particular year	pattern
Founded in conjunction with the initial development of the No	In 1999-2000	Founded in 1999, Alibaba started with a B2B model, providing an online platform for Chinese manufacturers to enable them to connect with international buyers.
Financing and expansion	In 2000-2003	Alibaba managed to raise outside funds, including investments from Yahoo, that helped the company scale and improve the functionality and security of the platform, while the company launched its international site to further promote its international business.
Electronic payments and financial services	In 2003-2010	Alibaba has launched Alipay, a secure online payment platform, which has quickly become one of the major online payment methods in China, and has supported Alibaba's e-commerce business. In addition, Alibaba began offering small loans and other financial services to meet the financing needs of small and medium-sized enterprises.
Global expansion and listing	End of 2010-present	Alibaba has further consolidated its position in the global e-commerce market by expanding its international business through acquisitions and strategic partnerships, and the company listed on the New York Stock Exchange in 2014, the largest initial public

		offering (IPO) ever

Table 3. The Development history of Alibaba

3. State-of-the-art

(I) Alibaba's business status and challenges

At present, Alibaba Group's main business includes: e-commerce business, cloud computing business and other businesses.

- E-commerce business: The group's domestic e-commerce business has achieved double-digit growth by implementing subsidies of 10 billion yuan and returning to users. The revenue of international e-commerce business has increased by 45% year on year. With the growth of revenue, the loss of international e-commerce is also increasing.
- Cloud computing: Its business is the only segment of all Alibaba businesses that has seen revenue growth and the adjusted EBITA is positive, and is experiencing growth.
- Other businesses: Other business segments include local life, entertainment, offline retail, Ali Health, Feizhu, Dingding, etc. Most of the above businesses are in a loss state, but they are vigorously expanding and investing. At the same time, the entertainment sector is also facing fierce market competition.

At present, the market competition for Alibaba Group is becoming increasingly fierce, such as Pinduoduo, Jingdong, Vipshop and other e-commerce platforms are developing rapidly, which has brought great pressure to the future development of Alibaba. At the same time, emerging sectors such as local life business are losing large losses. With the increasing development of the Internet, network security has also become a hot issue, so in order to ensure the development of the company, it is necessary to strengthen the company's technical security measures to ensure the data security and privacy of customers and users. At the same time, with the acceleration of digital transformation, the technology in artificial intelligence and big data application is facing great challenges. Under the "1 + 6 + N" reform of business model, it will face the challenges of management mode of restructuring organizational structure, the challenges of establishing business risk management with sound financial operation, and the challenges of cultivating coordination ability.

(2) Interpretation of Alibaba's financial results

Revenue for the financial year ended March 31,2024 was RMB 941.168 billion, Up 8% year on year, Operating profit was RMB 113.350 billion yuan, Up 13% year on year, Net profit

attributable to common shareholders was RMB 79.741 billion, Net profit was 71.332 billion yuan, Up 9% year on year, Earnings per diluted American Depositary share of RMB 31.24, And diluted earnings per diluted share of RMB 3.91, Net cash flow from operating activities was RMB 182.593 billion, Down 9% from RMB 199.752 billion in the fiscal year 2023, Free cash flow was RMB 156.210 billion yuan, Down 9% from RMB 171.663 billion in the fiscal year 2023, The year-on-year decline mainly reflects the special dividend of RMB 14.464 billion collected from Ant Group in fiscal year 2023 and changes in working capital, Was offset by some unadjusted EBITA year-over-year growth. According to the analysis of the above information, the overall revenue of Alibaba Group is on the upward trend, the operating profit increases greatly, and the company's net profit is on the upward trend compared with the previous year.

The figure below shows a three-dimensional matrix drawn by Alibaba Group's 1financial strategy in recent years [].

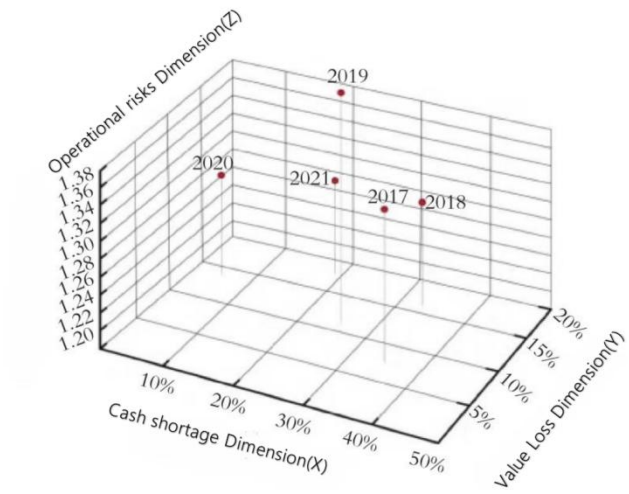

Figure 13 d matrix of Alibaba's financial strategy 2017-2021

(3) Alibaba's current business model

Alibaba's organizational structure has changed from a horizontal business company to a vertical parent-child holding group model. The corporate nature of Alibaba Group has also changed from business operation to capital operation. For the familiar "1 + 6 + N" mode, Alibaba Group has also achieved a worthy leap forward after the reform. The following is an in-depth analysis from three aspects of "1,6 and N".

1. The "1" -core

Referring to the paper "Comparative Analysis of Corporate Financial Strategy Matrix under Digital Platform Business Model- -Take Jingdong and Alibaba as an example"

"1" represents the core holding company of Alibaba Group. After years of deep cultivation and accumulation, Alibaba platform has built a huge business ecosystem, bringing together countless brands, merchants and consumers, which is the foundation of Alibaba.

2. The "6" large plate

"6" refers to the six major subsidiary business groups, which are the main business segments of Alibaba Group. Specifically, these business groups include Cloud Intelligence Group, Taotian Group, Local Life Group, Ali International Digital Business Group, Cainiao Group, and Grand Entertainment Group. In this business ecosystem, the respective development and layout of the six business segments not only highlight Alibaba's excellent market insight, but also show its accurate prediction of the future trend. The six business segments cooperate with each other and deepen deeply in their respective fields, which jointly constitute the comprehensive business ecology of Alibaba Group and promote the strategic transformation and development of the Group.

3. The "N" species have infinite possibilities

"N" refers to multiple smaller business companies that may be directly controlled by Alibaba Group or indirectly through various business groups. These companies represent Alibaba Group's investment and layout in more segmented areas, including finance, healthcare, education, new retail and so on. Thus, it can be seen that Alibaba can carry out targeted development and cooperation in large, medium and small enterprises in the future, and fully open the diversified pattern, which is conducive to exploring more channels for value creation.

(4) Alibaba's future development opportunities

With the acceleration of the global digital wave, Alibaba will also continue to push forward the digital transformation. In the future, Alibaba will further enhance the application capability of digital technology and expand the application scope of digital products and services. Alibaba will also continue to promote its internationalization strategy, strengthen its cooperation with overseas enterprises, and increase its investment in the global market. Alibaba is also expanding its business scale in cloud computing, and has now become one of the largest cloud computing service providers in China. In terms of new retail, Alibaba is exploring a new retail model, integrating offline and online, to provide consumers with a more convenient and convenient service experience.

4. Model display

(1) Index description and model assumptions

Table 4 illustrates the indicators of the model

Symbol	Description
open	at the opening
close	settlement
high	absolute valency
low	bottom price
volume	volume
money	turn volume
MACD	Exponentially smoothed similarities and similarity ages
MAE	mean absolute error
MSE	mean squared error
RMSE	root-mean-square error
MAPE	Mean absolute percentage error

Model Hypothesis:

Hypothesis 1: Suppose that a major event is not considered

Hypothesis 2: Observation error, random error and continuous problem discretization have no impact on the calculation and prediction of this problem.

Hypothesis 3: Suppose that the data collection and processing is accurate and reliable

Hypothesis 4: Suppose that there are certain correlation and mutual influence between different indicators.

Hypothesis 5: Suppose the government's policy is relatively stable, it will not affect Alibaba's stock price to produce big fluctuations.

(2) The current stock price situation of Alibaba

Through the application of matplotlib module in python software, the data of Alibaba stock price in recent years, and the application of drawing tools, it can be seen that the stock price of Alibaba Group is in a stable rising state. Specific trends are shown in the figure below.

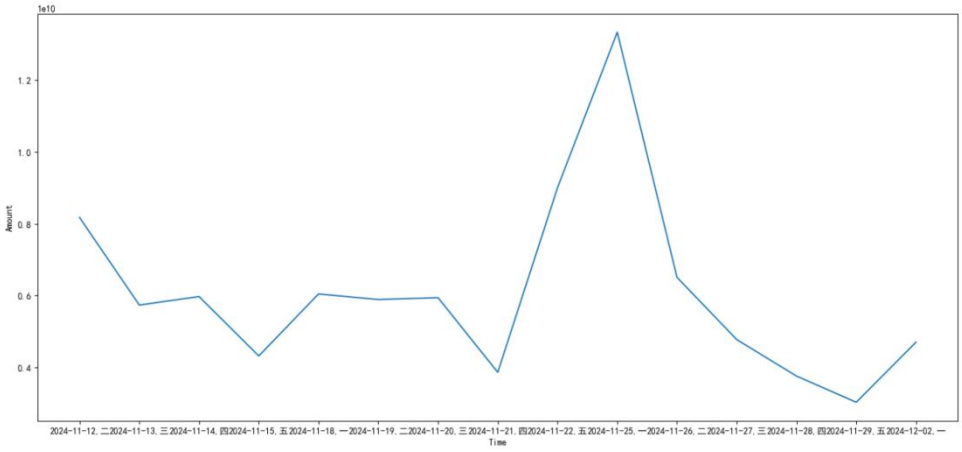

Figure 2. Alibaba's stock price status quo

According to get from flush software to alibaba group stock data, through the price, time and total hand and turnover rate analysis, in nearly 15 days, alibaba stock from price and total hand down trend can be seen, alibaba group in a recent period of time after a big shock, the overall share price has a trend of recovery.

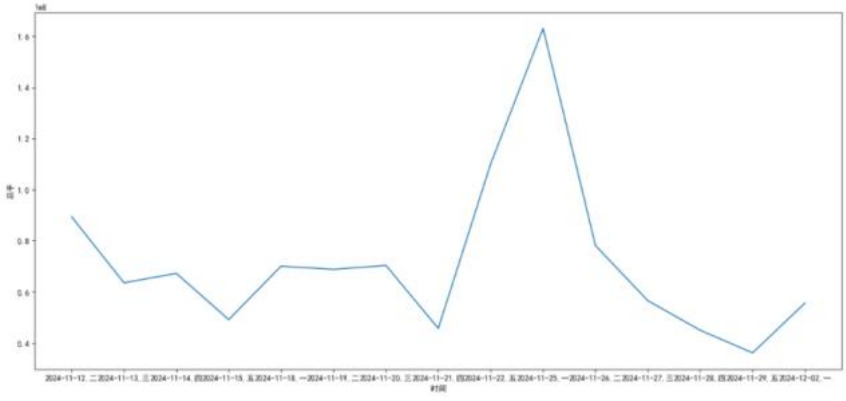

Figure 3 Up and downs of Alibaba Group Manager

(3) Construction of the random forest model

Random forest model (Random Forest) is an integrated learning algorithm based on decision tree, which can deal with regression problem. It builds a powerful prediction model by combining multiple decision trees, each of which is trained on randomly selected samples and features, and is voted to determine the final prediction results. Specifically, each decision tree in a random forest is constructed by bootstrap sampling and random feature selection. Bootstrap sampling means extracting a subset of samples of the same size from the original data set for training each decision tree. Random feature selection means that at each node, a part of all features is randomly

selected to split.

The steps of the random forest algorithm are as follows: First, for each tree Ti, randomly sample n samples from the training set D to obtain a sample subset Di, then randomly select k feature subset to train the decision tree; with Di and the selected feature subset. Second, for a new input vector X, the final prediction value is obtained by calculating the average of the output of all trees:

$$F(X)=\frac{1}{T}\sum_{i=1}^{T}T_i(X) \quad \text{Formula (1)}$$

Where T is the number of decision numbers.

(4) Construction of multiple linear regression model

The following multiple linear regression model was constructed:

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\ldots+\beta_kx_k+\varepsilon$$

The dependent variable y is the ε stock price, the random β_0 error term $\beta_1,\ \beta_2,\ \beta_3\ldots\beta_k$, the regression constant, and the regression coefficient, so that the estimated multiple linear regression model can be constructed:

$$\hat{y}=\hat{\beta}_0+\hat{\beta}_1x_1+\hat{\beta}_2x_2+\ldots\hat{\beta}_kx_k$$

The determination coefficient R2 is used to describe the statistical indicators of the data fitted by the regression model, which helps to evaluate the validity and accuracy of the regression model. RMSE is the root mean square error and is the square root of the difference between all predicted and actual values. MAE is the average absolute error, which is the average of the difference between the predicted value and the actual value, which can better reflect the actual situation of the predicted value error. MSE is the minimum squared error and is a measure of the accuracy of the data prediction model, and is often used to assess the degree of difference between the predicted value and the true value.

$$RMSE=\sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i-\hat{y}_i)^2} \quad \text{Formula (2)}$$

$$MAE=\frac{1}{m}\sum_{i=1}^{m}\left|\hat{y}_i-y_i\right| \quad \text{Formula (3)}$$

$$MSE=\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i-y_i)^2 \quad \text{Formula (4)}$$

$$MAPE=\frac{1}{m}\sum_{i=1}^{m}\left|\frac{\hat{y}_i-y_i}{y_i}\right| \quad \text{Formula (5)}$$

(5) General thinking of model construction

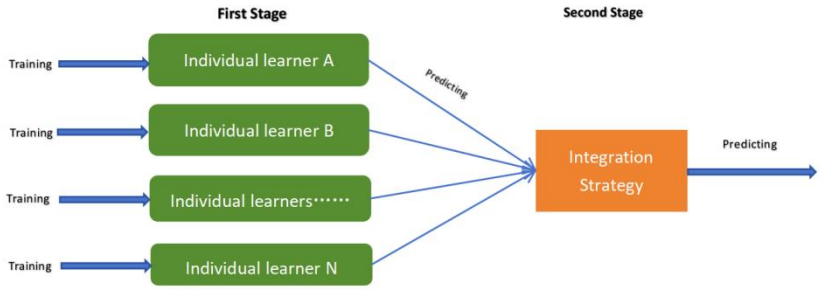

Figure 4 Ideas of integrated learning

Integrated learning is a machine learning method that improves model performance by combining multiple sub-learners. The core idea is to combine several sub-learners through a certain strategy to form a larger learner. Random forest is one of the important methods to realize ensemble learning.

The process of constructing the random forest model is as follows: First, random samples are selected and the selected samples are determined by the sampling method. Then, by randomly selecting the features, the features and the targets are separated, and then the decision tree is constructed.:

$$F(X)=\frac{1}{T}\sum_{i=1}^{T}T_i(X) \quad \text{Formula (1)}$$

The specific process is shown in Figure 5:

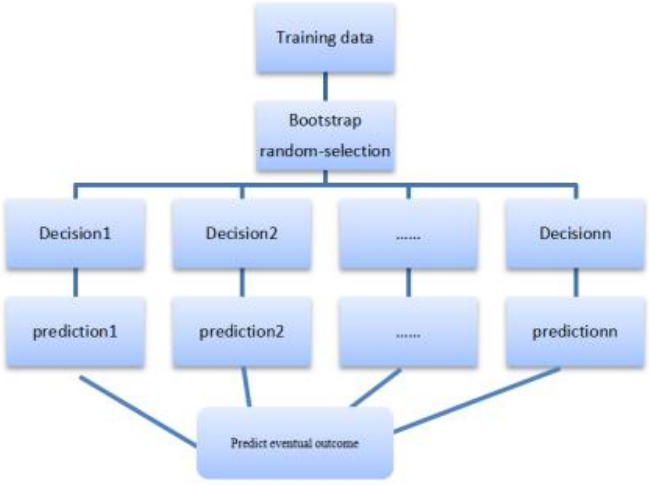

Figure Figure 5 General idea map of the random forest model

5. Stock price forecast

(1) Code implementation of using the random forest model for stock price prediction

Data preprocessing:

Data preprocessing includes data collation, data cleaning, and normalization processing.

The first step is the data collation. Collect the factors related to the target stock, and align the time span of the selected factor data into the factor data of the target stock.

The second step is the data cleaning. Clean duplicate, redundant, missing and errors into complete data that can be further processed and used.

The third step is the normalization process. To eliminate the dimensional influence between the factors, the data standardization needs to be processed to make the data comparable.

Divide the training set and the test set

X_test, X _ test, y_test, y _ test are divided under the test_size = 0.3 standard for subsequent model evaluation and prediction.

Model introduction

After the module installation import and data preprocessing, the model code of the random forest is imported. The code is shown in Figure 6.

```
rf = RandomForestRegressor(n_estimators=20
                           ,max_depth=3
                           ,random_state=42)

rf.fit(X_train,y_train)
```

```
RandomForestRegressor
RandomForestRegressor(max_depth=3, n_estimators=20, random_state=42)
```

Figure Figure 6 Random Forest model code import

Model evaluation

MAE, MSE, RMSE, MAPE and other indicators are calculated by the metrics in the machine learning module. The specific indicators are shown in Table 3.

Table 4 Specific values of the model evaluation indicators

metric	numeric value

MAE	406319951.04974383
MSE	5.217201868350938e17
RMSE	722302005.2824814
MAPE	7.93

According to the above formula can calculate the accuracy of the model score, X_train through python code running, y_train result is 0.8297280529500222; X _ test, y_train result is 0.7698620179277678, according to the score can see the model accuracy is accurate, the prediction result is more credible, for Alibaba's future stock price forecast is more accurate.

Model prediction

The predict method in the module is used to calculate, so as to predict the future transaction times and the future price trend, and through the scatter chart to show the forecast difference and the accuracy of the model.

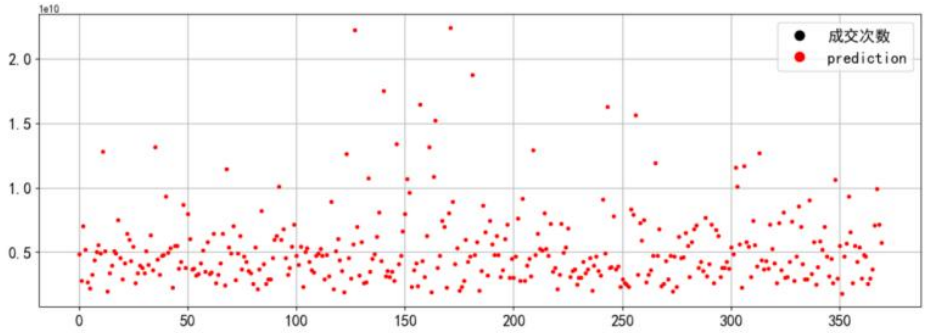

Figure 7 Forecast the future of Alibaba Group transaction times trend scatter chart

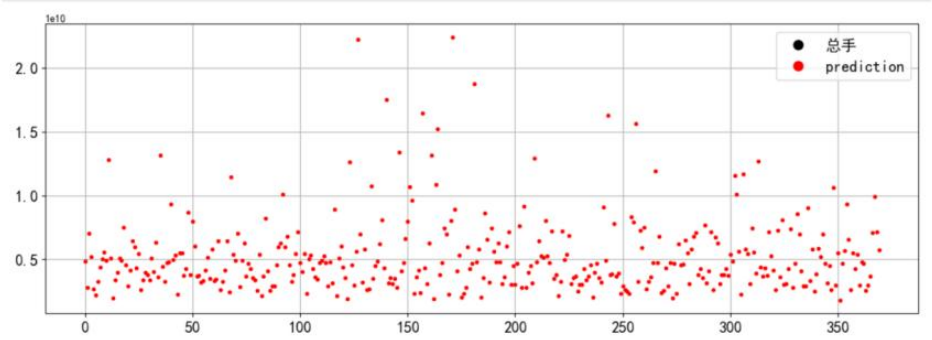

Figure 8 forecast the future trend of Alibaba Group general hand scatter chart

Model optimization

To improve the accuracy of the model prediction by adjusting the max _ depand

n_estimators parameters in the random forest model. The specific code is shown in Figure 9 and Figure 10.

```
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import GridSearchCV

param_grid = {
    "bootstrap":[True],
    "n_estimators":[5, 10, 20, 50, 100, 150, 200], #决策树的个数
    "max_depth":[3, 5, 7], #最大树深
    "max_features":[0.6, 0.7, 0.8, 1], #决策树划分时考虑的最大特征数
    "min_samples_leaf":[1, 2],
    "min_samples_split":[2, 3, 4]
}
#实例化随机森林回归器
rf = RandomForestRegressor(random_state=42)
#以随机森林回归器为基础构造网络搜索回归器
grid = GridSearchCV(rf, param_grid=param_grid, cv=3)
grid.fit(X_train, y_train)
```

```
57]:  GridSearchCV
      estimator: RandomForestRegressor
          RandomForestRegressor
```

Figure 9. Model optimization code modification section

```
#制定效果最好参数对应的模型
rf_reg = RandomForestRegressor(bootstrap=True
                               ,n_estimators=50
                               ,max_depth=7
                               ,max_features=0.6
                               ,min_samples_leaf=1
                               ,min_samples_split=3
                               ,random_state=42)
rf_reg.fit(X_train, y_train)
```

```
]:  RandomForestRegressor
    RandomForestRegressor(max_depth=7, max_features=0.6, min_samples_split=3,
                          n_estimators=50, random_state=42)
```

Figure 10 Specific parameter changes of model optimization

(2) General description

Today, cloud computing, big data, block chain, artificial intelligence, such as advanced electronic technology is rapid development, e-commerce momentum is fierce, business sector and business trend is also experiencing reshaping, cloud computing by providing on-demand allocation of computing resources and services, greatly reduce the operating costs of enterprises, improve the operational efficiency, provides a strong support for the digital transformation for the enterprise. And e-commerce also by virtue of its own convenience, efficiency and global, received warm attention. Alibaba group development, in the "1 + 6 + N" and a series of financial strategy, operation and management reform, with the company business sector (cloud computing, electricity, etc.) of efficient operation and effective jurisdiction, alibaba shares rose steadily in

recent years, according to the flush APP derived stock data, using the random forest model calculation, as shown in figure 11.

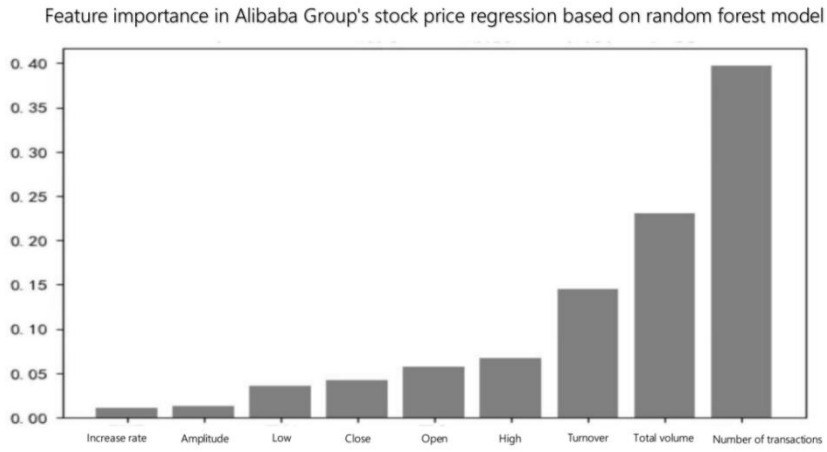

Figure 11 Characteristic importance in Alibaba stock price regression based on random forest model

As can be seen from Figure 11, Alibaba's share price is mainly affected by the daily number of transactions. According to the survey and study, we can see the proportion and profitability of the business segments of Alibaba Group as shown in Figure 12.

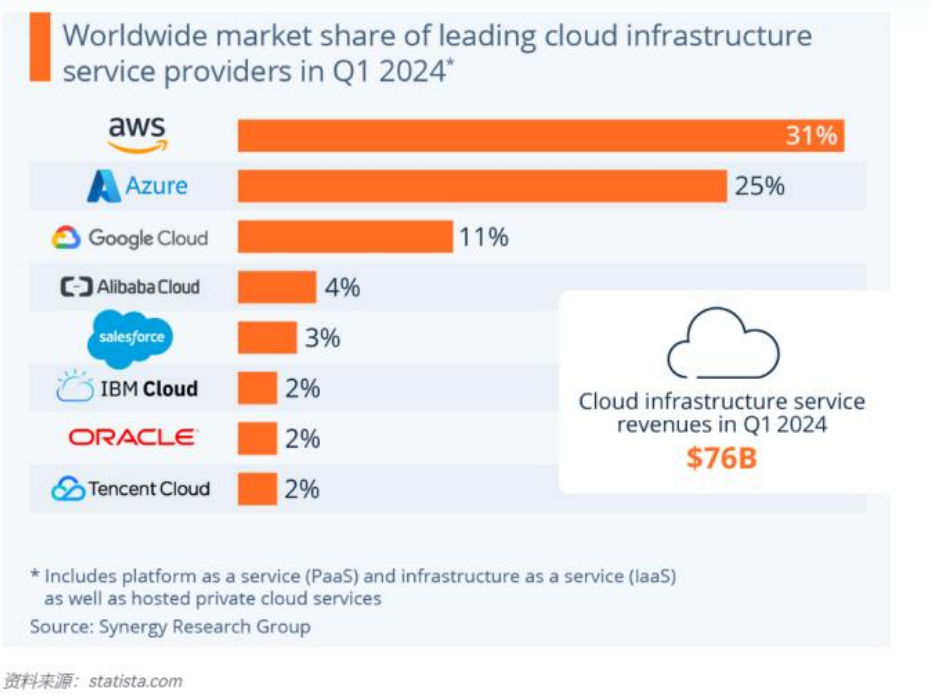

Figure 12 Business segments of Alibaba Group

Target	Low	Average	Median	High
Price	$85	$110.26	$107	$140
Change	-4.00%	+24.53%	+20.85%	+58.12%

资料来源：stockanalysis.com

Figure 13 Alibaba stock price forecast trend chart

Figure 13 shows the forecast trend of Alibaba Group in the future stock and the line chart. By May 2025, the overall stock price of Alibaba Group showed a steady upward trend. Under the lowest, highest and average forecast, it can be seen that it is a stable and small fluctuation trend. Relative strength index (RSI) is 72.4, higher than the bullish level 31.67, and significantly higher than the bearish level of 48.82, which can see alibaba has a strong bullish momentum, RSI line upward trend, ali group income has diversified characteristics, focus on expanding international influence and enhance the consumer experience, so the company revenue and profits will continue to grow. With growing demand for cloud services and AI products, Aliyun's revenue will be a bullish indicator. Alibaba is conducting massive stock buybacks and dividends, trying to return value for shareholders, strengthen shareholder confidence and stabilize stocks in the face of volatile earnings. Currently, Taobao and Tmall Group (TTG) are strengthening their user-first strategy by improving product supply, competitive prices and service quality. This will drive double-digit growth in GMV, while increasing the number of members in the 88VIP program. Aliyun is also adjusting its product strategy in the era of artificial intelligence, with triple-digit growth in AI-related revenue. According to the AIDC (Alibaba International Digital Commerce) report, Alibaba International Digital Commerce revenue rose 45 percent and orders rose 20 percent, driven by AliExpress's shift to a supply chain efficiency-driven model and its strategic expansion in key markets. Combined with the data and model implementation analysis, Alibaba Group's stock price will continue to develop in a stable trend, continue to rise and fluctuate with a small fluctuation in the future.

6. References

[1] Zou Jie, Li Lu. Based on SA-BiGRU model

[2] Li Bin. Research on the application of deep learning in the stock price prediction model

[3] Luckyson Khaidem,Snehanshu Saha,Sudeepa RoyDey.Predicting the direction of stock market prices using random forest

[4] Cao Yu. Analysis of stock rise and fall prediction model based on dynamic parameter adjustment KNN classification algorithm

[5] Liu Ye. Building a new business ecology: Alibaba's "1 + 6 + N" organizational reform strategy analysis

[6] Qian Qimiao, Zhang Duo, Wang Ying, Liu Rui, Cai Feijun. The application of machine learning in stock price forecasting

[7] Tan Qi, Zhang Xuemei. Research on the path of corporate management change under the background of digital economy

[8] Chen Zi Zhihai. Stock price forecasting research based on machine learning algorithms

[9] Han Dongmei, Wang Jin, Gu Yuping. Comparative analysis of the financial strategy matrix under the digital platform business model- -take Jingdong and Alibaba as examples