Comparative Analysis of Differences of American Pharmaceutical Stocks Before and After the Epidemic

Shicheng Zhao[1,*]

[1] College of Information Science and Technology, Jinan University, Guang Zhou, 511495, China
[*]Corresponding author. Email: 250828zsc@stu2019.jnu.edu.cn

ABSTRACT

This research adopts the data of American medical stocks in the recent ten years from Yahoo Finance and compares and analyzes the data before and after the epidemic. Try to analyze whether the epidemic has had a great impact on the stock price trend of American medical stocks. The machine learning models predict the amount of price-rising days of stock in the epidemic stage, compared with the price trend before the epidemic. Judging whether the epidemic has significantly reduced the price-rising days of medicine stocks in the United States. By regressing the data of the price movement, we can judge whether the epidemic has had a great impact on the stock price trend. And try to pay attention to whether there is a significant gap between the stock prices of vaccine-producing companies and non-vaccine-producing companies under the influence of the epidemic.

Keywords: *Machine learning, Stock price, Epidemic impact, Feature classification*

1. INTRODUCTION

1.1. Background

Since the covid-19 pneumonia outbreak, the virus has spread rapidly all over the world. The epidemic has had a far-reaching impact on human society, and people's normal social activities have been hindered. The pneumonia epidemic also caused great disasters to the world economy. At the onset of the epidemic, the stock markets of major economies in the world suffered a tragic decline, and the US stock market triggered the circuit breaker mechanism many times. After the initial control of the epidemic, most countries in the world began to shift the focus toward the development of economic growth again.

The United States has implemented a flexible monetary policy and issued a substantial amount of money to promote funds to go into the stock market and then raise the stock index. Some countries have also promoted domestic investment by reducing bank deposit interest. Most countries have maintained the overall stock price level at a relatively stable state through many different efforts.

Pharmaceutical companies in the United States have undertaken the task of providing vaccines for COVID-19, so the relevant pharmaceutical enterprises face a large demand in this epidemic. Under the macro influence of US monetary policy, more investors are willing to invest in these pharmaceutical companies that produce vaccines this time. Therefore, reviewing all epidemic stages, the prices of pharmaceutical stocks in the U.S. stock market did not show a largely downward trend under the leadership of these two factors.

By analyzing the price trend of American Medical stocks before and after the epidemic, this article tries to judge whether the epidemic has had an impact on the original price trend of pharmaceutical stocks, or whether the reflections of the price trend considering the basic data of the stocks is different from the performance law before the epidemic.

1.2. Related research

There are a lot of researches have been done. For example, Altig considered economic uncertainty led by COVID-19. They believed that subjective uncertainty about future business growth and disagreement among professional forecasters about future GDP growth were the drivers. And these factors point to an enormous increase in uncertainty in response to the pandemic and its economic benefits. In contrast, the more general

Y. Jiang et al. (Eds.): ICEDBC 2022, AEBMR 225, pp. 238–244, 2022.
https://doi.org/10.2991/978-94-6463-036-7_35

measures of uncertainty peaked later and then peaked as job losses increased, highlighting the difference in measures of uncertainty between Wall Street and Main Street [1]. COVID-19 also leads to uncertainty, which includes basic hospital resources, medications and the high-profile provision of personal protective equipment. Ignorance extends to the place of care as well. Although the number of hospital beds devoted to COVID-19 has declined, there have been unforeseen increases in the level of nursing home care for frail seniors with complex needs in a context of limited staff and resources [2]. These factors may also give a huge impact on the demand side, which will also have the ability to enhance the related stock's price. The uncertainty surrounding COVID is widespread. He reviewed the time line and quantified the impact of COVID-19 uncertainty on regional market aggregate returns and volatility using the ARCH/GARCH models. Based on economic psychology, the uncertainty related to COVID-19 is measured by information searches mirrored by Google's search trends. Asian markets are more resilient than those in other parts of the world. In terms of returns and volatility, Latin American markets are the most hit [3].

Firms that manufacture non-durable household goods such as toiletries, shampoo, toothpaste, and shaving cream are good examples of defensive firms because the public will continue to use these items during an epidemic [4]. In the face of the pandemic, investors put a larger amount of their portfolio in money market securities of deposit, such as certificates of deposit, Governments Treasury bills, and fixed instruments or money market instruments, which are less volatile than the equities and commodity markets [5]. The stock market forecast is open to many ML algorithms. At the same time, there are models with the high success that can be applied due to their time-series data. Results are very promising, it can say that ANN and ARIMA models are more successful in this regard. The evaluation among other ML algorithms by comparing the results can be suggested as another research subject [6].

Vignesh's project is a demonstration of how machine learning may be used to tackle stock prediction challenges. The stock market's historical data was used to train the model so that it could detect trends and patterns and, as a result, forecast future data. This project also proved that LSTM worked better compared to back propagation and SVM algorithms [7].

The stock market trend is influenced by a number of financial indexes. However, not all financial indicators are suitable for forecasting. Some financial indicators may have a higher correlation with stock market trends, while others may have a lower correlation. In Y. Lin's generally speaking, a good subset should not only contain the diversity of selected financial indexes but also have higher effects on the stock market trend.

Brzoza-Brzezina investigates this trade-off and its consequences for monetary policy. To this end, this research constructs a model that draws from the epidemic modeling literature and the macroeconomic business cycle literature. More precisely, we connect a SIR-type model with a standard new Keynesian framework. This allows to speak not only about the pandemic (and potential containment measures) but also about macroeconomic effects and monetary policy[9]. Jiang to deal with the economic issue caused by the epidemic, it is necessary to use existing experience for reference. Not only overseas measures but also domestic former experience is important. Facing the risk, it is significant to view the newborn opportunity rather than short-term interest [10].

1.3. Objective

Support Vector Machines (SVM) have earned popularity among Machine Learning (ML) algorithms used for predicting stock price[4]. This paper also uses SVM model for supervised learning. SVM is mainly used to classify the data, and SVR model and some other models are used to predict the data. An algorithm similar to the following method is used in feature selection. Some paper has done that after doing some clustering for the features of the stock, the financial indexes from the same cluster have similar characteristics. From each cluster, selecting one financial index as the delegate based on its influence quantities on the output obtained in the SVM-based feature ranking module[8]. With a better calculation of the similarities for the characteristics, the AP algorithm shows a better result than the model with the default parameter ‘reference’.

2. METHODOLOGY

2.1. Feature Selection

The affinity Propagation (AP) algorithm is a clustering algorithm based on "information transmission" between data points. Unlike the k-means algorithm or k-center point algorithm, the AP algorithm does not need to determine the number of clusters before running the algorithm. Using AP for features selection did a good job than finishing the same work in k-means.

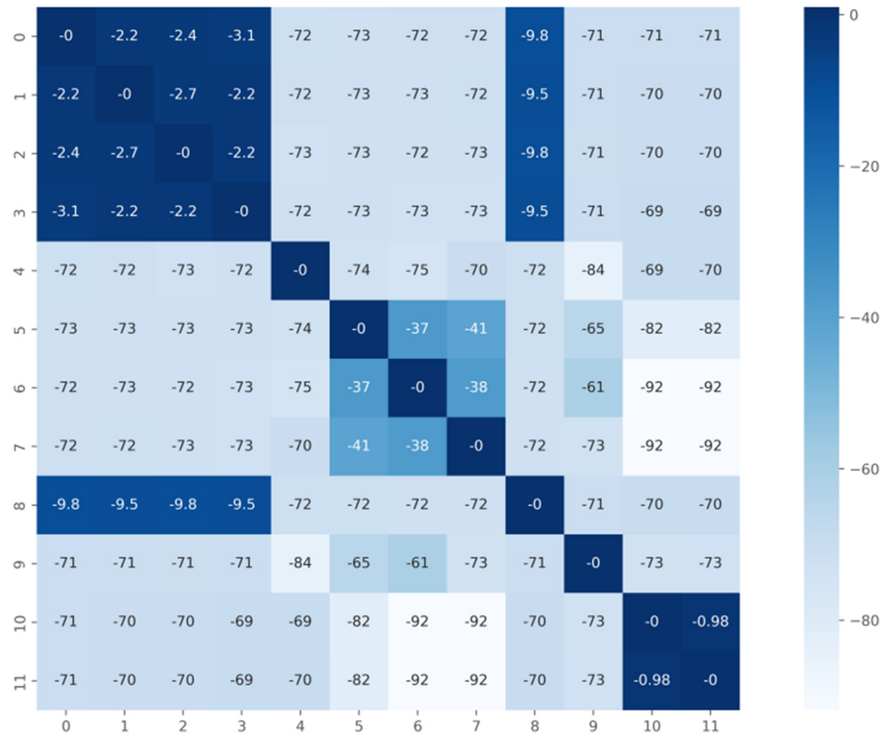

Figure 1 Similarities Matrix

The parameter named reference is the most important one for the clustering result. This research used a method to numeric the similarity of those features. For two features vector v1 and v2, the negative Euclidean distance is used as the formula to calculate the degree of similarity. Then there is the similarity matrix, as shown in Figure 1)

In this matrix, the darker the color, the smaller the calculation result. That is, the higher the similarity between the two vectors. It can be seen from Figure 2 that the four features in the upper left corner are highly correlated and are likely to be divided into one cluster in the next AP algorithm procession.

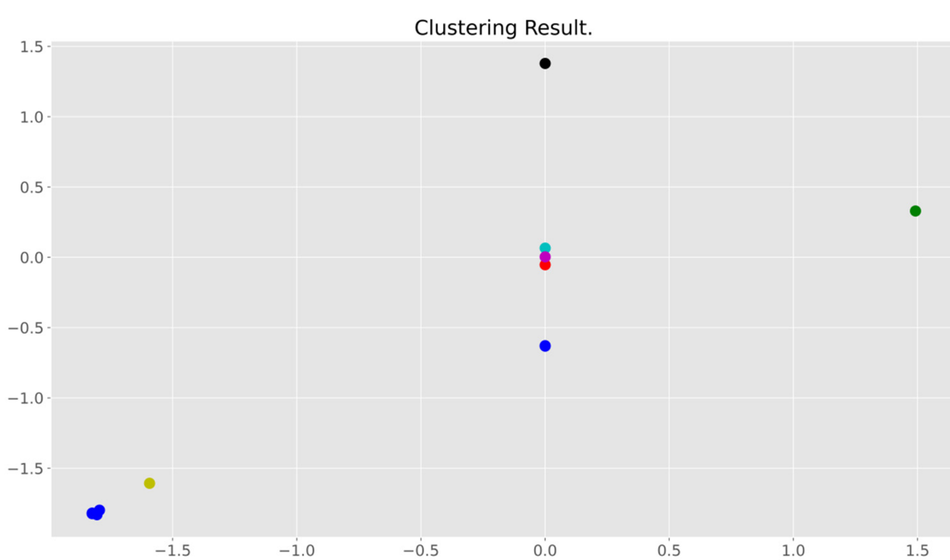

Figure 2 AP Results

Next, according to the results shown by this matrix, performing AP algorithm cluster analysis on 12 feature vectors, and performing the above operations of selecting features which are described at the front of this paper. Figure 2 Cluster Result shows the result of the AP algorithm.

According to Figure 2, the AP algorithm divides 12 feature vectors into 8 parts. According to the program results in output by python, selecting one feature vector in each part to form the final feature matrix according to the method mentioned earlier in the article.

According to the code, the output results of the handler are shown in the table below(Table 2.1.3.Feature Selection Result). In the following table, the same groups represent the same features, so select a vector in each group to enter the final feature matrix. If there is only one feature vector in classification, this vector is directly selected into the final feature matrix.

Table 1. Feature Selection Result

Features	Clustering Group	Features	Clustering Group
Open	Group0	Adj.Open	Group2
High	Group0	Adj.Open	Group3
Low	Group0	Adj.Low	Group4
Close	Group0	Adj.Close	Group5
Volume	Group1	Adj.Volume	Group6
Simple Return	Group7	Log Return	Group7

According to this method, the vectors of the final matrix represent "closing price, stock volume, changing prices and changing volume and the natural logarithm of return”.

2.2. Data Preparation and Procession

Taking Pfizer as an example, Pfizer is famous for the vaccines he produced during the epidemic. Getting the relative data by using the yahoo finance API in python, and there are the "close price, low price, high price, open price, adjustive close price and volume". Setting the "adjustive close price " as the label of this research.

Using “Simple Imputer” to impute the mean value into the Nan items, and using the “Standard Scaler” to remove Remove dimensional problems.

2.3. Price-rising Dates and Price trend Prediction

The epidemic has affected the general trend and specific prices of stocks. Therefore, this paper uses the characteristic matrix before the epidemic to build a model, calculate the price rise date and specific price of stocks after the epidemic, and observe the gap between stock prices before and after the epidemic. Based on the data obtained by the above methods. This paper attempts to analyze whether the price impact of the epidemic on American pharmaceutical stocks is significantly related to the performance of the company in the epidemic.

3. NUMERICAL EXPERIMENT

3.1. Classification training based on the SVM model

Using the SVM algorithm to do the supervised learning, the result is in Table 1 which shows the accuracy with standard variation in brackets. All the classifications use a train set with the size of it is 1362. Using the crossed validation to evaluate the results. The result shows that SVM with the Linear kernel has the best accuracy among these three different kernels. And all the accuracy is above 95%, the results have high reliability.

Although the variance of linear classification may not be the smallest, that is to say, the fluctuation of the results of linear classification is greater than that of the other two classifications. However, because the accuracy of the linear classification is high enough and the absolute value of variance is within an acceptable range, the linear kernel classifier is still used in the later prediction procession.

3.2. Price trend training based on multiple regression models

Using Linear regression, polynomial regression (the highest quadratic square term), and SVR algorithm model with different kernels to predict the specific price of all 9 stocks. In Table 3 the prediction errors can be found. Among these five models, the polynomial regression with a quadratic square term always shows the smallest error, all the errors are only about 1% order of magnitude. So, the polynomial regression is used in the following price prediction.

Table 2. Train Result of quadratic square term

Stock ID	Stock Group	Price-rising Day			Price Prediction				
		SVM			Linear	Poly	SVR		
		Linear	Poly	RBF	-	-	Poly	Linear	RBF
PFE	Vaccine	0.954 (0.023)	0.913 (0.029)	0.935 (0.002)	2.43%	1.05%	167.14%	2.41%	2.43%
JNJ	Vaccine	0.983 (0.002)	0.942 (0.010)	0.952 (0.006)	1.76%	0.95%	580.08%	1.79%	2.07%
MRK	Vaccine	0.988 (0.010)	0.942 (0.013)	0.942 (0.003)	2.78%	0.82%	185.56%	2.77%	3.66%
REGN	Vaccine	0.993 (0.01)	0.954 (0.019)	0.966 (0.006)	5.96%	2.61%	81299.58%	6.18%	3.48%
ABT	Test and Cure	0.974 (0.004)	0.92 (0.013)	0.937 (0.02)	2.36%	0.61%	1106.19%	2.42%	2.22%
ABBV	No Relation	0.985 (0.003)	0.887 (0.034)	0.96 (0.012)	2.73%	0.45%	2334.88%	2.73%	2.46%
BMY	No Relation	0.976 (0.021)	0.94 (0.034)	0.962 (0.007)	2.56%	1.19%	153.88%	2.55%	1.85%
LLY	No Relation	0.99 (0.007)	0.892 (0.026)	0.933 (0.015)	3.23%	0.83%	114.79%	3.19%	2.98%
SYK	No Relation	0.988 (0.002)	0.954 (0.017)	0.962 (0.01)	2.35%	0.60%	126.36%	2.37%	1.72%

3.3. Result Discussion

After training all the required models, these models are used to calculate the characteristic data of the stocks used on the stock trading day after the epidemic. And predicting the stock price trend that these stocks should behave in theory according to the model trained with the data before the epidemic, that is, if there is no impact of the epidemic, what should their price look like, and how many days their price should increase. The actual data and the prediction are shown in the table 2.

In terms of the number of days of price-rising days, 66.7% of the stocks showed that the actual price-rising days in the epidemic were lower than those predicted based on the trend before the outbreak, but all the declines are within 10%. Surprisingly, the number of price-rising days of 75% of the companies responsible for producing vaccines in the epidemic fell, while there are increases in the amount days of those companies related to the demand for the epidemic increased. It can be simply understood as the demand impact of the epidemic on relevant companies, after the vaccine company develops the vaccine in a short time and effectively puts it on society, it weakens investors' confidence in the long-term appreciation of the company's stock. More money has been transferred to companies with more long-term investment value.

In terms of price prediction, companies dedicated to the detection of symptoms of virus-infected persons during the epidemic have become the biggest winner, but there is no data on the stocks of other similar companies, so there is not too much research. The decline in the prices of companies that produced vaccines was much smaller than the decline in the shares of companies unrelated to the epidemic. This confirms the huge demand for the vaccine from pharmaceutical companies caused by COVID-19, effectively preventing the impact of macroeconomic backwardness, which is caused by the epidemic on the stock prices of these companies.

Table 3. Final Result

Stock ID	Stock Group	Price-rising Day			Price Prediction			Group Gap Analysis	
		Actual	Prediction	Falling	Actual	Prediction	Gap	Mean	Variance
PFE	Vaccine	274	275	0.36%	3.299	3.616	-0.014	-0.006	0.002
JNJ		287	300	4.53%	4.624	5.018	-0.005		
MRK		275	296	7.64%	3.921	4.32	0.023		
REGN		295	290	-1.69%	5.909	6.295	-0.01		
ABT	Test Cure	299	311	4.01%	3.948	4.703	0.027	0.027	0.002
ABBV	No Relation	313	304	-2.88%	4.09	4.576	-0.025	-0.089	0.008
BMY		291	291	0.00%	3.854	4.096	-0.002		
LLY		286	285	-0.45%	4.396	5.129	-0.087		
SYK		297	320	7.74%	4.787	5.447	0.025		

4. CONCLUSION

The sudden outbreak of coronavirus pneumonia has brought huge uncertainties to world economic development. The price of American pharmaceutical stocks is also affected by the epidemic, which is different from that before the epidemic in terms of the number of price-rising days and price trends. According to the conclusion of this experiment, the epidemic has led to a lower than expected increase in the overall price of pharmaceutical stocks, but the price remains stable on the whole. Under the influence of the epidemic, the price of American pharmaceutical stocks related to the treatment of pneumonia received the least impact, followed by those companies that can provide a large number of vaccines after the outbreak of the epidemic. Compared with the price before the epidemic, the price of pharmaceutical stocks that have little to do with the epidemic has the most obvious downward trend, and the impact of the epidemic is the greatest.

The above conclusion can be seen more intuitively in Figure 3. The vertical axis of the picture shows the gap between the actual price and the expected price. Three different colors represent companies with different performances in the epidemic.

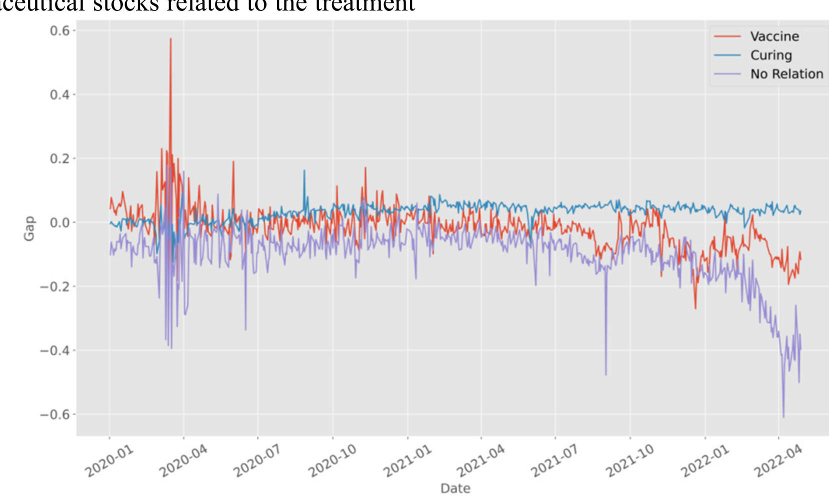

Figure 3 Three kinds companies' price difference

The blue line represents the prices of stocks responsible for detection and treatment in the epidemic. It has the smallest fluctuation and the smallest price difference. Next to and below the blue line is the stock price difference of vaccine manufacturers, which has greater volatility. And the overall price difference is greater than the blue line. This means that the stock prices of companies producing vaccines have fallen more significantly than the situation expected by algorithm, and the volatility caused by the uncertainty of the epidemic is more significant.

The purple line at the bottom represents the stock price difference trend of pharmaceutical companies that have made little contribution to the epidemic. It can be seen that its absolute value is the largest, that is to say, the epidemic has caused these stock prices to decline to a great extent than expected, and the closer it is to the post epidemic era, the more obvious the gap is. Based on the above analysis, the following conclusions can be drawn: The uncertainty of the epidemic has harmed stock prices, and the prices of all pharmaceutical stocks have decreased to varying degrees compared with expectations. However, companies related to the epidemic received huge demand caused by the epidemic, resulting in the expected difference in stock prices is not obvious.

REFERENCES

[1] D.Altig, S.Baker, J. M.Barrero, N.Bloom, P.Bunn, S.Chen, G.Thwaites, Economic uncertainty before and during the COVID-19 pandemic. Journal of Public Economics, vol.191, no.104274, 2020.

[2] J.Koffman, J.Gross, S. N.Etkind, L.Selman, Uncertainty and COVID-19: how are we to respond?. Journal of the Royal Society of Medicine, vol.113, no.6, 2020, pp.211-216.

[3] J. J.Szczygielski, P. R.Bwanya, A.Charteris, J.Brzeszczyński, The only certainty is uncertainty: An analysis of the impact of COVID-19 uncertainty on regional stock markets. Finance research letters, vol.43, no.101945, 2021.

[4] I.Nti, A.Adekoya, B.Weyori, Efficient Stock-Market Prediction Using Ensemble Support Vector Machine. Open Computer Science, vol.10, no.1, 2020, pp.153-163. https://doi.org/10.1515/comp-2020-0199

[5] A. F.Adekoya, I.Kofi Nti, The COVID-19 outbreak and effects on major stock market indices across the globe: A machine learning approach. Indian Journal of Science and Technology, vol.13, no.35, 2020, pp.3695-3706.

[6] S.Kalyoncu, J.Rasheed, A.Jamil, M.Yesiltepe, Machine Learning Methods for Stock Market Analysis. In 3rd International Conference on Data Science and Applications (ICONDATA), 2020.

[7] C. K.Vignesh, Applying machine learning models in stock market prediction. EPRA International Journal of Research & Development (IJRD), 2020, pp395-398.

[8] Y. Lin, H. Guo and J. Hu, "An SVM-based approach for stock market trend prediction," The 2013 International Joint Conference on Neural Networks (IJCNN), 2013, pp.1-7, doi:10.1109/IJCNN.2013.6706743.

[9] M.Brzoza-Brzezina, M.Kolasa, K.Makarski, Monetary policy and COVID-19. IMF Working Papers, vol.2021, no.274, 2021.

[10] Z.Jiang, X.Wang, H.Wang, G.Lu, G.Wang, X.Gong, The Comparison Analysis of China and Foreign Countries to Deal with the Economic Risk by Epidemic, in Terms of Monetary and Fiscal Policy. American Journal of Industrial and Business Management,vol.10, no.09, 2020. pp.1605.