

Báo cáo Phân tích Dữ liệu Cổ phiếu Tesla

Tạ Quang Tuấn SE190841 Phạm Hoàng Nam SE192423

Nguyễn Cao Thành Trung SE191057

Trần Minh Thọ SE190999

Ngày 23 tháng 3 năm 2025

1 Lý do chọn chủ đề và vấn đề nghiên cứu

Chủ đề nghiên cứu về giá cổ phiếu Tesla được chọn vì tính hấp dẫn và sự phát triển mạnh mẽ của công ty này trên thị trường chứng khoán. Tesla là một công ty nổi bật trong lĩnh vực xe điện và năng lượng tái tạo, có ảnh hưởng lớn đến nền kinh tế và xu hướng tiêu dùng toàn cầu. Phân tích giá cổ phiếu Tesla có thể giúp nhận diện xu hướng thị trường, rủi ro và cơ hội đầu tư, đưa ra các phương án đầu tư - kinh doanh dưới các điều kiện kinh tế khác nhau. Trong bài báo này, chúng ta sẽ tập trung vào việc phân tích các yếu tố tác động đến giá cổ phiếu Tesla, bao gồm các thống kê mô tả, kiểm định giả thuyết, Mô phỏng phân phối xác suất, Mô phỏng phân phối lợi nhuận và rủi ro và mô hình hồi quy tuyến tính.

2 Giải thích và phân tích kèm số liệu và biểu đồ

2.1 Giới thiệu dữ liệu

Dữ liệu: Giá cổ phiếu Tesla (TSLA) trên Kaggle

Dữ liệu được thu thập từ cổ phiếu Tesla trong khoảng thời gian từ năm 8/2016 đến 8/2021. Dữ liệu bao gồm các chỉ số quan trọng như:

- **Date:** Ngày giao dịch
- **Open:** Giá mở cửa trong ngày
- **High:** Giá cao nhất trong ngày
- **Low:** Giá thấp nhất trong ngày
- **Close:** Giá đóng cửa
- **Adj Close:** Giá đóng cửa điều chỉnh

- **Volume:** Khối lượng giao dịch trong ngày

Dưới đây là một ví dụ về các dòng dữ liệu:

	Open	High	Low	Close	Adj Close	Volume
Date						
2016-08-16	45.098000	45.438000	44.681999	44.722000	44.722000	11335500
2016-08-17	44.866001	44.966000	44.560001	44.647999	44.647999	8935500
2016-08-18	44.764000	45.132000	44.458000	44.702000	44.702000	8572500
2016-08-19	44.708000	45.034000	44.506001	45.000000	45.000000	8297500
2016-08-22	44.834000	45.021999	44.535999	44.585999	44.585999	10327500

Hình 1: Dữ liệu cổ phiếu Tesla (một phần)

2.2 Thống kê mô tả dữ liệu

Tổng quan thống kê mô tả về dữ liệu cổ phiếu Tesla cho thấy các đặc điểm quan trọng của dữ liệu, chẳng hạn như giá trị trung bình, độ lệch chuẩn, giá trị cực tiểu và cực đại, và các phần trăm phân vị. Cụ thể:

- **Mean (Trung bình):** Giá cổ phiếu Tesla trong khoảng thời gian nghiên cứu có giá trị trung bình là 183.35 USD.
- **Standard Deviation (Độ lệch chuẩn):** Độ lệch chuẩn là 227.99, cho thấy sự biến động lớn trong giá cổ phiếu.
- **Min (Giá trị tối thiểu):** Giá thấp nhất đạt 35.79 USD.
- **Max (Giá trị tối đa):** Giá cao nhất đạt 891.38 USD.
- **25%, 50%, 75% (Các phân vị):** Các giá trị phân vị cho thấy sự phân phối không đồng đều, với giá trị ở mức 25% là 54.99 USD, 50% (trung vị) là 66.76 USD, và 75% là 163.86 USD.

Bảng thống kê mô tả dưới đây tóm tắt các chỉ số trên:

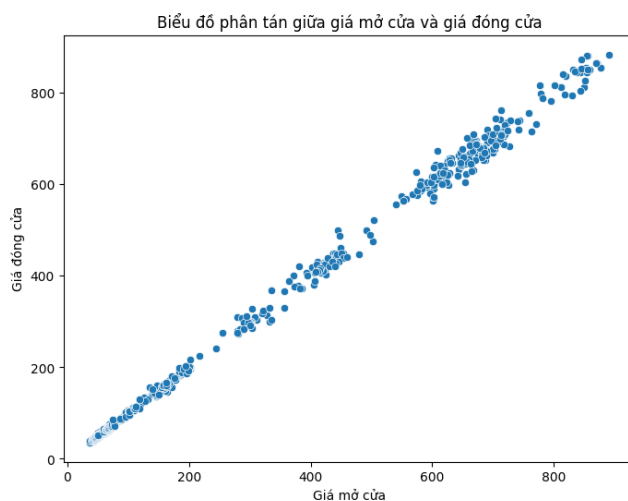
	Open	High	Low	Close	Adj Close	Volume
count	1258.000000	1258.000000	1258.000000	1258.000000	1258.000000	1.258000e+03
mean	183.274022	187.265251	178.941224	183.351102	183.351102	4.441050e+07
std	227.979867	232.845126	222.443374	227.987027	227.987027	3.035737e+07
min	36.220001	36.945999	35.397999	35.793999	35.793999	8.297500e+06
25%	54.985001	55.930001	53.929001	54.994501	54.994501	2.478255e+07
50%	66.602002	67.950001	65.445999	66.756000	66.756000	3.498150e+07
75%	165.516998	168.243504	161.504005	163.856503	163.856503	5.211238e+07
max	891.380005	900.400024	871.599976	883.090027	883.090027	3.046940e+08

Hình 2: Thống kê mô tả về giá cổ phiếu Tesla

Các thống kê này giúp ta hiểu rõ hơn về sự phân phối và mức độ biến động của giá cổ phiếu Tesla trong suốt thời gian nghiên cứu.

2.3 Biểu đồ phân tán

Biểu đồ phân tán giữa giá mở cửa và giá đóng cửa cho thấy mối quan hệ chặt chẽ giữa hai yếu tố này. Khi giá mở cửa tăng, giá đóng cửa cũng có xu hướng tăng theo.



Hình 3: Biểu đồ phân tán giữa giá mở cửa và giá đóng cửa

2.4 Ma trận tương quan giữa các yếu tố

Ma trận tương quan giữa các yếu tố cho thấy giá mở cửa, giá cao nhất, giá thấp nhất và khối lượng giao dịch có mối quan hệ tương đối thấp với giá đóng cửa. Điều này chỉ ra rằng các yếu tố như giá mở cửa và khối lượng giao dịch không phải là yếu tố quyết định duy nhất cho giá cổ phiếu, mà còn có thể bị ảnh hưởng bởi các yếu tố khác như tâm lý thị trường, tin tức kinh tế và các yếu tố vĩ mô.



Hình 4: Ma trận tương quan giữa các yếu tố

3 Ước lượng trung bình tổng thể và tỷ lệ

Dựa trên dữ liệu của cổ phiếu Tesla, ta tính toán ước lượng trung bình tổng thể và khoảng tin cậy cho trung bình, cũng như tỷ lệ ngày có thay đổi giá dương.

3.0.1 Ước lượng trung bình tổng thể

```
Ước lượng trung bình tổng thể: 183.3511015317965
Khoảng tin cậy 95% cho trung bình: (170.75262866205333, 195.94957440153968)
```

Hình 5: kết quả

Trung bình mẫu của giá cổ phiếu Tesla được tính toán là 183.35 USD. Khoảng tin cậy 95% cho trung bình tổng thể được tính theo công thức:

$$CI = \bar{X} \pm Z_{\alpha/2} \times \frac{s}{\sqrt{n}}$$

Trong đó:

- \bar{X} là trung bình mẫu,
- $Z_{\alpha/2}$ là giá trị Z cho mức độ tin cậy 95%,
- s là độ lệch chuẩn mẫu,
- n là kích thước mẫu.

```
Tỷ lệ ngày có thay đổi giá dương: 0.5174880763116058
Khoảng tin cậy 95% cho tỷ lệ: (0.48987517954976023, 0.5451009730734513)
```

Hình 6: Kết quả

Kết quả cho thấy khoảng tin cậy 95% cho trung bình là:

(170.75, 195.95)

3.0.2 Tỷ lệ ngày có thay đổi giá dương

Tỷ lệ ngày có thay đổi giá dương được tính là 0.5175, nghĩa là khoảng 51.75% số ngày có giá đóng cửa cao hơn giá mở cửa. Khoảng tin cậy 95% cho tỷ lệ này được tính bằng công thức:

$$CI = p_{\text{observed}} \pm Z_{\alpha/2} \times \sqrt{\frac{p_{\text{observed}}(1 - p_{\text{observed}})}{n}}$$

Kết quả cho thấy khoảng tin cậy 95% cho tỷ lệ ngày có thay đổi giá dương là:

(0.4899, 0.5451)

4 Kiểm định giả thuyết

4.1 Kiểm định t-test cho giá cổ phiếu

```
T-statistic: 20.745637904159203, P-value: 1.8903356478184406e-82
Reject H0: The mean close price is significantly different from 50 (H1 is true).
```

Hình 7: Kết quả

Kiểm định t-test được sử dụng để kiểm tra sự khác biệt giữa giá trị trung bình thực tế của giá cổ phiếu và giá trị giả thuyết. Cụ thể, ta kiểm tra giả thuyết sau:

$H_0 : \mu = 50$ (Giá trị trung bình cổ phiếu bằng 50)

$H_1 : \mu \neq 50$ (Giá trị trung bình cổ phiếu khác 50)

Giá trị t-statistic được tính bằng công thức sau:

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Trong đó:

```
T-statistic for difference in means: -46.208063422364285, P-value: 3.6088293968693175e-273
Reject H0: The mean close price is significantly different between the two periods (H1 is t
```

Hình 9: Kết Quả

- \bar{X} là giá trị trung bình mẫu,
- μ_0 là giá trị giả thuyết về trung bình (50),
- s là độ lệch chuẩn mẫu,
- n là kích thước mẫu.

Kết quả kiểm định t-test cho thấy P -value rất nhỏ ($P = 1.89 \times 10^{-82}$), do đó ta bác bỏ giả thuyết H_0 và kết luận rằng giá trị trung bình của giá cổ phiếu Tesla là khác biệt đáng kể so với 50.

4.2 Kiểm định tỷ lệ thay đổi giá dương

```
Z-statistic: 1.2405445237561246, P-value: 0.10738702832675495
Fail to reject H0: The proportion of days with positive price change
```

Hình 8: Kết Quả

Kiểm định tỷ lệ thay đổi giá dương cho thấy tỷ lệ thay đổi giá dương của cổ phiếu Tesla là 51.75%. Tuy nhiên, kiểm định Z cho thấy tỷ lệ này không có sự khác biệt đáng kể so với 50%, với P -value = 0.107.

Kiểm định Z được tính theo công thức:

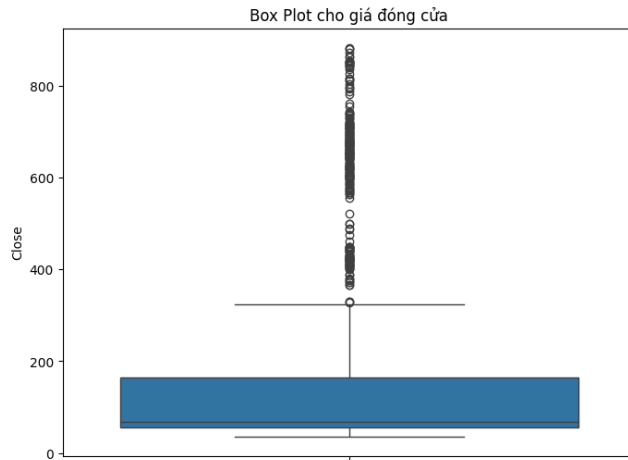
$$z = \frac{p_{\text{observed}} - p_{\text{expected}}}{\sqrt{\frac{p_{\text{expected}}(1-p_{\text{expected}})}{n}}}$$

Trong đó:

- p_{observed} là tỷ lệ ngày có thay đổi giá dương,
- p_{expected} là tỷ lệ kỳ vọng (50%),
- n là tổng số ngày quan sát.

4.3 Kiểm định sự khác biệt giữa các nhóm

Kết quả kiểm định t-test giữa hai giai đoạn 2016-2020 và 2020-2022 cho thấy sự khác biệt rất lớn về giá trị trung bình của giá cổ phiếu trong hai giai đoạn này với P -value cực kỳ nhỏ ($P = 3.61 \times 10^{-273}$). Điều này chỉ ra rằng có sự thay đổi mạnh mẽ trong giá cổ phiếu Tesla qua các năm.



Hình 10: Box Plot cho giá đóng cửa

Biểu đồ box plot cho thấy sự phân phối không đối xứng của giá đóng cửa, với một số lượng lớn các giá trị nhỏ, đồng thời có một số ngoại lệ (outliers) trong phạm vi giá cao.

5 Mô hình hồi quy tuyến tính

5.1 Mean Squared Error (MSE)

Công thức tính MSE (Sai số bình phương trung bình) là:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Trong đó:

- y_i là giá trị thực tế,
- \hat{y}_i là giá trị dự đoán,
- n là số lượng mẫu.

MSE cho biết mức độ sai lệch trung bình giữa các giá trị thực tế và giá trị dự đoán. MSE càng nhỏ, mô hình càng chính xác.

5.2 R-Square (R^2 Score)

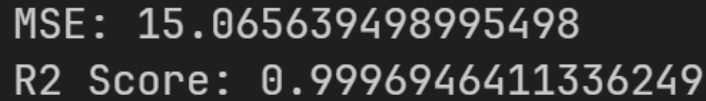
Công thức tính R^2 (Hệ số xác định) là:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Trong đó:

- \bar{y} là giá trị trung bình của y_i ,

R^2 cho biết mức độ phù hợp của mô hình với dữ liệu (giá trị từ 0 đến 1, càng gần 1 mô hình càng phù hợp).



MSE: 15.065639498995498
R2 Score: 0.9996946411336249

Hình 11: MSE và R^2

Mô hình hồi quy tuyến tính giữa giá đóng cửa và các yếu tố khác như giá mở cửa, giá cao, giá thấp và khối lượng giao dịch cho thấy hệ số hồi quy rất mạnh, với $R^2 = 0.9997$, chứng tỏ rằng mô hình có thể giải thích hầu hết sự biến động của giá cổ phiếu Tesla.

Mô hình hồi quy tuyến tính có dạng:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \epsilon$$

Trong đó:

- y là giá đóng cửa,
- X_1, X_2, X_3, X_4 là các yếu tố đầu vào (giá mở cửa, giá cao, giá thấp, khối lượng giao dịch),
- β_0 là intercept,
- $\beta_1, \beta_2, \beta_3, \beta_4$ là các hệ số hồi quy,
- ϵ là sai số ngẫu nhiên.

Hệ số góc β_1 thể hiện độ dốc của đường hồi quy. Công thức tính β_1 như sau:

$$\beta_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

Trong đó:

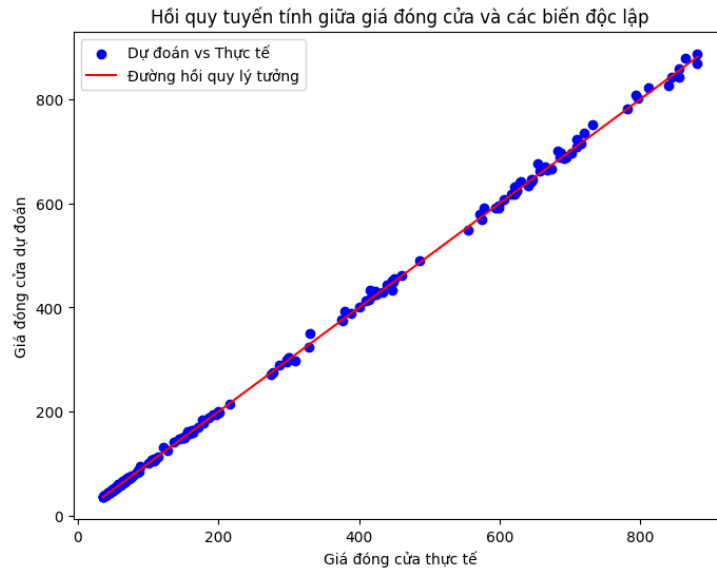
- X_i và Y_i là các giá trị dữ liệu quan sát được,
- \bar{X} và \bar{Y} là các giá trị trung bình của X và Y ,
- n là số lượng quan sát.

Hệ số chặn β_0 cho biết giá trị của Y khi $X = 0$. Công thức tính β_0 như sau:

$$\beta_0 = \bar{Y} - \beta_1 \bar{X}$$

Trong đó:

- \bar{Y} là giá trị trung bình của Y ,
- β_1 là hệ số góc đã tính ở trên,
- \bar{X} là giá trị trung bình của X .



Hình 12: Hồi quy tuyến tính giữa giá đóng cửa và các biến độc lập

	Date	Predicted Price	Actual Price
0	2018-11-06	69.004981	68.211998
1	2017-01-10	45.665588	45.973999
2	2016-10-27	41.025877	40.801998
3	2016-11-14	36.284609	36.290001
4	2020-11-18	490.702069	486.640015

Hình 13: Kết quả dự đoán

6 Phân phối lợi nhuận đầu tư

Công thức tính lợi nhuận kỳ vọng ($E(X)$) và độ lệch chuẩn (σ) như sau:

$$E(X) = \frac{1}{n} \sum_{i=1}^n X_i$$

Trong đó:

- X_i : Giá trị lợi nhuận của từng ngày.
- n : Tổng số ngày quan sát.

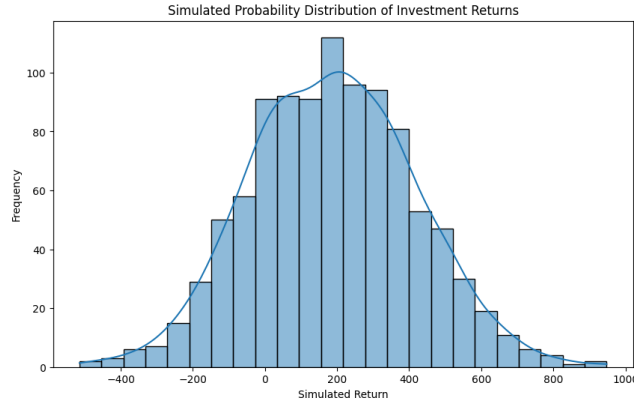
Độ lệch chuẩn được tính bằng công thức:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - E(X))^2}$$

Trong đó:

- $E(X)$: Lợi nhuận kỳ vọng.

Simulasi phân phối lợi nhuận đầu tư từ giá cổ phiếu Tesla cho thấy một phân phối chuẩn với giá trị kỳ vọng là 183.35 và độ lệch chuẩn là 227.99.



Hình 14: Simulated Probability Distribution of Investment Returns

7 Phân phối nhị thức và Poisson

7.1 Phân phối Nhị Thức

Phân phối nhị thức là một phân phối xác suất rời rạc dùng để mô tả số lần thành công trong một số lượng thử nghiệm cố định, mỗi thử nghiệm có xác suất thành công không đổi. Công thức của phân phối nhị thức được tính như sau:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Trong đó:

- n : Số lần thử nghiệm (số dự án đầu tư),
- k : Số lần thành công (số dự án thành công),
- p : Xác suất thành công trong một lần thử nghiệm,
- $\binom{n}{k}$: Hệ số nhị thức, tính số cách chọn k thành công từ n thử nghiệm.

Phân phối này có thể được sử dụng để mô phỏng số dự án thành công trong các tình huống có xác suất thành công cố định.

7.2 Phân phối Poisson

Phân phối Poisson là một phân phối xác suất dùng để mô tả số lượng sự kiện xảy ra trong một khoảng thời gian cố định hoặc không gian, với điều kiện rằng các sự kiện xảy ra độc lập và với tỷ lệ không đổi. Công thức của phân phối Poisson được tính như sau:

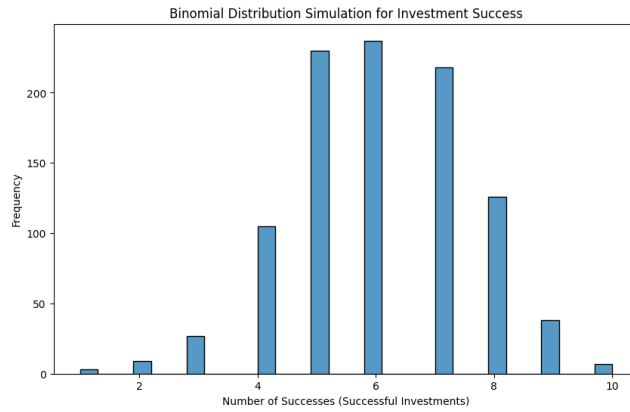
$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Trong đó:

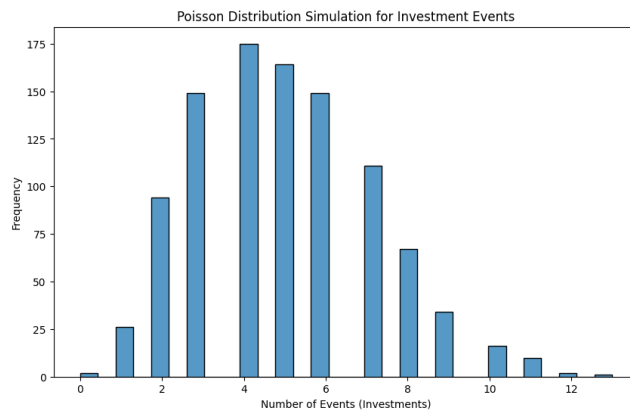
- k : Số lượng sự kiện xảy ra,
- λ : Tỷ lệ xảy ra sự kiện trong một khoảng thời gian hoặc không gian (tức là kỳ vọng số sự kiện),
- e : Cơ số của logarithm tự nhiên.

7.3 Phân tích kết quả phân phối nhị thức và phân phối Poisson

Phân phối Poisson thích hợp trong các tình huống mô tả số lần xảy ra của một sự kiện trong một khoảng thời gian hoặc không gian nhất định, chẳng hạn như số lượng đầu tư trong một thời gian cụ thể. Phân phối nhị thức mô phỏng thành công các dự án đầu tư cho thấy rằng số lần thành công dao động trong khoảng từ 2 đến 8, trong khi phân phối Poisson mô phỏng số sự kiện đầu tư cho thấy sự phân bố các sự kiện với tỷ lệ vừa phải.



Hình 15: Binomial Distribution Simulation for Investment Success



Hình 16: Poisson Distribution Simulation for Investment Events

8 Kết luận và hướng mở rộng

Từ các phân tích trên, ta có thể kết luận rằng giá cổ phiếu Tesla có sự biến động mạnh và phụ thuộc vào nhiều yếu tố khác nhau, bao gồm các yếu tố vĩ mô và tin tức thị trường. Mô hình hồi quy tuyến tính cho thấy rằng các yếu tố như giá mở cửa và khối lượng giao dịch có tác động nhỏ đến giá đóng cửa, nhưng vẫn có mối quan hệ chặt chẽ giữa chúng. Để đưa ra quyết định đầu tư chính xác, các nhà đầu tư cần phải đánh giá lợi nhuận kỳ vọng và rủi ro của các dự án, đặc biệt trong các điều kiện kinh tế thay đổi.

Dựa trên kết quả phân tích lợi nhuận kỳ vọng (183.35 USD) và độ lệch chuẩn (227.99 USD), ta có thể thấy rằng việc đầu tư vào cổ phiếu Tesla có mức độ rủi ro khá cao. Tuy nhiên, với mức lợi nhuận kỳ vọng lớn, nhà đầu tư có thể cân

nhắc việc tham gia đầu tư trong dài hạn, đặc biệt là khi nền kinh tế ổn định và công ty Tesla tiếp tục phát triển.

Hướng mở rộng nghiên cứu có thể bao gồm việc áp dụng các mô hình phức tạp hơn như mô hình ARIMA hoặc mạng nơ-ron nhân tạo để dự đoán chính xác hơn về giá cổ phiếu Tesla. Bên cạnh đó, việc kết hợp các yếu tố vĩ mô như chính sách của chính phủ, giá dầu, và các yếu tố toàn cầu sẽ giúp tăng cường độ chính xác trong dự đoán. Các nhà đầu tư cũng có thể sử dụng phương pháp đánh giá dòng tiền (DCF - Discounted Cash Flow) để ước lượng giá trị hiện tại của các dự án và đưa ra quyết định đầu tư dựa trên các kịch bản kinh tế khác nhau.

Cuối cùng, việc phân tích các phân phối xác suất của lợi nhuận đầu tư từ cổ phiếu Tesla giúp nhà đầu tư hiểu rõ hơn về các rủi ro tiềm ẩn và đưa ra các phương án đầu tư phù hợp với mục tiêu tài chính cá nhân, đồng thời tối ưu hóa danh mục đầu tư của mình.

Tài liệu

- [1] <https://machinelearningcoban.com/2016/12/28/linearregression/>
- [2] <https://www.kaggle.com/code/serkanp/tesla-stock-price-prediction/input>
- [3] <https://khoakinhteluat.ufm.edu.vn/Resources/Docs/SubDomain/khoakinhteluat/Gi%C3%A1o%20tr%C3%ACnh%201%C3%BD%20thuy%E1%BA%BFt%20x%C3%A1c%20su%E1%BA%A5t%20v%C3%A0%20th%E1%BB%91ng%20k%C3%AA%20%E1%BB%A9ng%20d%E1%BB%A5ng.pdf>