

ĐẠI HỌC QUỐC GIA TP.HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KINH TẾ - LUẬT
KHOA TÀI CHÍNH-NGÂN HÀNG



BÁO CÁO CUỐI KÌ
MÔN: ỨNG DỤNG PYTHON TRONG TÀI CHÍNH I

**ỨNG DỤNG MÔ HÌNH GARCH VỚI DỮ
LIỆU CHUỖI THỜI GIAN VỀ GIÁ CỔ
PHIẾU SONY**

Sinh viên: Đỗ Thị Lan Phương

MSSV: K194141741

Thành phố Hồ Chí Minh, Tháng 1 năm 2022

Mục lục

1. GIỚI THIỆU	2
2. MỤC TIÊU, ĐỐI TƯỢNG VÀ PHẠM VI NGHIÊN CỨU	2
2.1 Mục tiêu nghiên cứu.....	2
2.2 Đối tượng nghiên cứu.....	2
2.3 Phạm vi nghiên cứu	2
3. PHƯƠNG PHÁP NGHIÊN CỨU	3
3.1 Nội dung của mô hình GARCH	3
3.2 Dữ liệu	3
4. NỘI DUNG VÀ KẾT QUẢ NGHIÊN CỨU.....	6
4.1 Thứ tự thực hiện mô hình	6
4.2 Kết quả phân tích.....	6
4.2.1 Diễn biến giá cổ phiếu SONY và tỷ suất sinh lời	6
4.2.2 Kiểm định tính dừng và hiệu ứng ARCH	8
4.2.3 Thiết lập và ước lượng mô hình GARCH.....	9
5. KẾT LUẬN	12

1. GIỚI THIỆU

Cuộc khủng hoảng tài chính toàn cầu (2008) bùng nổ và dịch bệnh Covid-19 ở hiện tại đã có phần nào các tác động tiêu cực đến thị trường tài chính của các nền kinh tế, kéo theo sự sụp đổ của hàng loạt tổ chức tài chính và tập đoàn đầu tư lớn và cũng kéo theo chỉ số chứng khoán toàn cầu giảm mạnh... Cho nên, việc nghiên cứu, phát triển và đề xuất các mô hình quản lý rủi ro phù hợp để lượng hóa và dự đoán các tổn thất tài chính là thực sự cần thiết.

Trong tài chính có nhiều mô hình đánh giá rủi ro như ARCH, GARCH, VAR,... Tuy nhiên, hai mô hình cơ bản và phổ biến được sử dụng mở rộng ra thành nhiều mô hình khác là ARCH và GARCH. Mô hình ARCH được sử dụng để mô phỏng các dữ liệu chuỗi thời gian trong tài chính có phương sai thay đổi theo thời gian. Còn mô hình GARCH là mô hình nâng cao, phát triển lên từ mô hình ARCH bằng cách thêm vào các yếu tố phương sai có điều kiện của độ trễ. Trong bài này, em sẽ sử dụng mô hình GARCH để ước lượng tỷ suất sinh lời của cổ phiếu của Sony.

Bài nghiên cứu sẽ gồm các phần cụ thể như sau:

Phần 1: Tổng quan mục tiêu, đối tượng và phạm vi nghiên cứu

Phần 2: Trình bày phương pháp nghiên cứu về mô hình được sử dụng trong bài và dữ liệu

Phần 3: Trình bày nội dung và kết quả ước lượng được từ mô hình ứng dụng. Từ đó, tiến hành phân tích các kết quả thu được

Phần 4: Tổng kết lại bài nghiên cứu

2. MỤC TIÊU, ĐỐI TƯỢNG VÀ PHẠM VI NGHIÊN CỨU

2.1 Mục tiêu nghiên cứu

Ứng dụng mô hình GARCH vào ước lượng tỷ suất sinh lời của cổ phiếu SONY, trong đó sẽ ứng dụng python xây dựng mô hình GARCH để kiểm chứng khả năng giải thích của nó cho sự thay đổi của tỷ suất sinh lời cổ phiếu SONY

2.2 Đối tượng nghiên cứu

Sử dụng mô hình GARCH cho việc ước lượng lợi suất cổ phiếu của Công ty công nghiệp Sony.

2.3 Phạm vi nghiên cứu

Nghiên cứu dựa trên các số liệu thứ cấp được thu nhập trên thị trường chứng khoán chứng khoán thế giới, mẫu dữ liệu là giá đóng cửa (bởi vì lấy data API ở trang web twelvedata chỉ có giá đóng cửa chứ không có giá đóng cửa hiệu chỉnh), từ đó tính tỷ suất sinh lời theo thời gian của Công ty công nghiệp Sony.

3. PHƯƠNG PHÁP NGHIÊN CỨU

3.1 Nội dung của mô hình GARCH

Dựa vào giá đóng cửa của cổ phiếu, ta xác định được tỷ suất sinh lời của cổ phiếu theo công thức:

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}} \times 100$$

Trong đó:

- P_t là giá đóng cửa của SONY tại ngày t
- P_{t-1} là giá đóng cửa của SONY tại ngày $t-1$

Trong mô hình GARCH, phương sai có điều kiện tham gia trực tiếp vào phương trình trung bình. Tỷ suất sinh lời của cổ phiếu có thể phụ thuộc vào độ biến động của nó. Mô hình GARCH(p,q) tổng quát có dạng như sau:

Phương trình trung bình:

$$R_t = \beta_1 + \beta_2 X_t + u_t$$
$$u_t \sim N(0, h_t)$$

Phương trình phương sai:

$$h_t^2 = \gamma_0 + \sum_{i=1}^p \delta_i h_{t-i}^2 + \sum_{j=1}^q \gamma_j u_{t-j}^2$$

Trong đó: $\delta_i \geq 0, \gamma_j \geq 0$ để đảm bảo phương sai lớn hơn 0. Hệ số γ_j đo sự biến động có thể xảy ra ở thời kỳ tiếp theo. Nếu hệ số δ_i cao điều đó thể hiện có sự biến động trong thời gian dài. Độ lớn của các tham số δ, γ quyết định tác động ngắn hạn của dao động chuỗi thời gian. Nếu tổng hệ số hồi quy bằng 1, cú shock sẽ có tác động đến sự biến động của SONY trong dài hạn. Đó là cú shock với phương sai có điều kiện kiên lâu dài.

Điều kiện tồn tại mô hình:

Chuỗi dữ liệu theo thời gian phải có tính dừng, nghĩa là có các giá trị trung bình, Phương sai và hiệp phương sai tại các độ trễ khác nhau sẽ giống nhau, hay nói cách khác, các đại lượng này là không đổi theo thời gian. Bởi vì nếu chuỗi không dừng sẽ gây ra hiện tượng “hồi quy giả mạo”

- Tiêu chuẩn AIC và BIC: là tiêu chí để tìm ra mô hình GARCH phù hợp. Mô hình được chọn là mô hình có AIC hoặc BIC nhỏ nhất.

3.2 Dữ liệu

- Thu thập dữ liệu nghiên cứu

Dữ liệu nghiên cứu được thu thập qua cổng dữ liệu API của URL twelvedata bằng key riêng (api_key = 'dfb57946806a4f30985a037482f4c40b'). Dữ liệu được sử dụng lệnh request từ URL của website twelvedata ở dạng json. Sau đó, được chuyển về dạng dataframe để thuận tiện cho việc xử lý dữ liệu. Dữ liệu bao gồm giá mở cửa, giá đóng cửa, giá cao/thấp và khối lượng giao dịch trong ngày. Được lấy 15 năm từ ngày 01/01/2007 đến ngày 11/01/2022 gồm 3783 quan sát.

```
ticker = 'SONY'
api_key = 'dfb57946806a4f30985a037482f4c40b'
interval = '1day'

url = f'https://api.twelvedata.com/time_series?symbol={ticker}&start_date=2007-01-01&end_date=2022-01-11&interval={interval}&apikey={api_key}'
data = requests.get(url).json()

sony_data = pd.DataFrame(data['values'])
sony_data.head(10)
```

	datetime	open	high	low	close	volume
0	2022-01-10	123.50000	123.70000	121.17000	123.17000	1028500
1	2022-01-07	125.50000	125.74000	123.30000	124.31000	574800
2	2022-01-06	125.40000	126.31000	122.37000	123.11000	1836500
3	2022-01-05	133.56000	133.75000	128.31000	128.59000	2594400
4	2022-01-04	128.53000	129.31000	127.83000	128.56000	708700
5	2022-01-03	126.40000	127.17000	125.65000	126.27000	373700
6	2021-12-31	125.80000	126.79000	125.50000	126.40000	275000
7	2021-12-30	126.25000	126.69000	125.76000	125.92000	232200
8	2021-12-29	127.00000	127.13000	126.30000	126.69000	259100
9	2021-12-28	128.00000	128.30000	127.28000	127.48000	415700

- Phương pháp xử lý và tính toán dữ liệu
 - Xử lý dữ liệu
- Đầu tiên, kiểm tra kiểu dữ liệu.

```
sony_data.dtypes
```

```
datetime    object
open        object
high        object
low         object
close       object
volume      object
dtype: object
```

Thì thấy rằng các cột “open”, “high”, “low”, “close”, “volume” sai kiểu dữ liệu. Cho nên sẽ chuyển về kiểu dữ liệu như sau: các cột “open”, “high”, “low”, “close” về kiểu float, “volume” về kiểu int. Cột “datetime” không xử lý bởi vì sẽ dùng để reset index.

- Tiếp theo, set index bằng cột “datetime” và sắp xếp theo đúng thứ tự.

```
sony_data = sony_data.sort_values(by=['datetime'])
sony_data.set_index('datetime', inplace=True)
sony_data
```

	open	high	low	close	volume
datetime					
2007-01-03	42.90	43.37	42.73	42.91	1200600
2007-01-04	43.18	43.88	43.12	43.80	1209600
2007-01-05	43.98	45.60	43.96	44.80	3197500
2007-01-08	44.81	45.33	44.43	44.81	2344300
2007-01-09	46.25	47.00	45.80	46.40	1731000
...
2022-01-04	128.53	129.31	127.83	128.56	708700
2022-01-05	133.56	133.75	128.31	128.59	2594400
2022-01-06	125.40	126.31	122.37	123.11	1836500
2022-01-07	125.50	125.74	123.30	124.31	574800
2022-01-10	123.50	123.70	121.17	123.17	1028500

3783 rows × 5 columns

- Kiểm tra null thì không có null trong dataframe.

```
sony_data.isnull().sum()
```

```
open      0
high      0
low       0
close     0
volume    0
dtype: int64
```

- Tính toán dữ liệu
- Tính toán tỷ suất sinh lời của cổ phiếu SONY theo giá đóng cửa. Như đã giải thích ở phần phạm vi nghiên cứu, bởi vì đọc dữ liệu API từ cổng website của twelvedata chỉ cho giá đóng cửa, không có giá đóng cửa hiệu chỉnh nên trong bài nghiên cứu sẽ tính toán tỷ suất sinh lời giản đơn dựa trên giá đóng cửa. Bởi khi tính tỷ suất sinh lời thì ngày đầu tiên sẽ không có kết quả nên sẽ dùng lệnh dropna() xóa luôn dòng đó.

```

sony_data = sony_data.loc[:, ['close']]
sony_data['returns'] = sony_data.close.pct_change()*100
sony_data = sony_data.dropna()
sony_data

```

	close	returns
datetime		
2007-01-04	43.80	2.074109
2007-01-05	44.80	2.283105
2007-01-08	44.81	0.022321
2007-01-09	46.40	3.548315
2007-01-10	45.78	-1.336207
...
2022-01-04	128.56	1.813574
2022-01-05	128.59	0.023335
2022-01-06	123.11	-4.261607
2022-01-07	124.31	0.974738
2022-01-10	123.17	-0.917062

3782 rows × 2 columns

4. NỘI DUNG VÀ KẾT QUẢ NGHIÊN CỨU

4.1 Thứ tự thực hiện mô hình

- Tính toán chuỗi tỷ suất lợi nhuận của cổ phiếu SONY theo ngày
- Kiểm định tính dừng của chuỗi tỷ suất lợi nhuận theo ngày bằng kiểm định Dickey-Fuller (ADF)
- Kiểm định hiện tượng phương sai có điều kiện của chuỗi tỷ suất lợi nhuận thay đổi theo thời gian bằng kiểm tra hiệu ứng ARCH
- Thiết lập và ước lượng mô hình GARCH

4.2 Kết quả phân tích

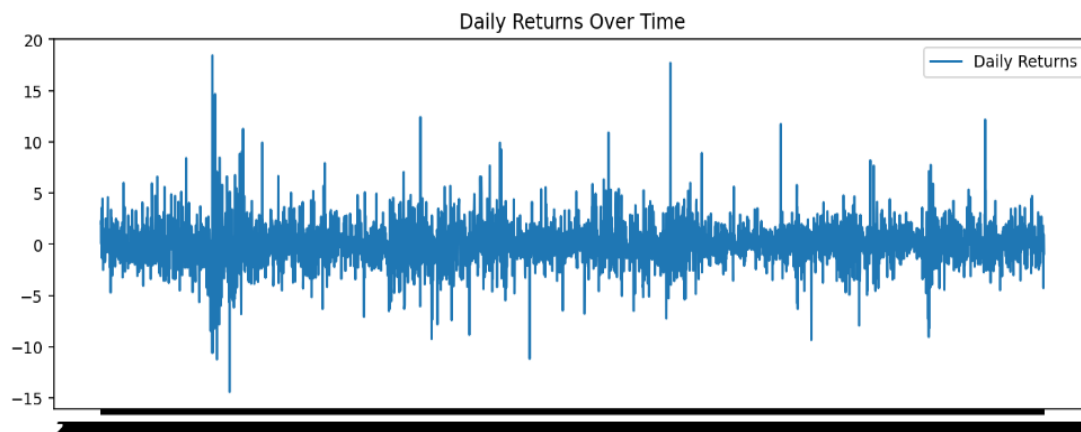
4.2.1 Diễn biến giá cổ phiếu SONY và tỷ suất sinh lời

- Diễn biến thực tế giá cổ phiếu SONY



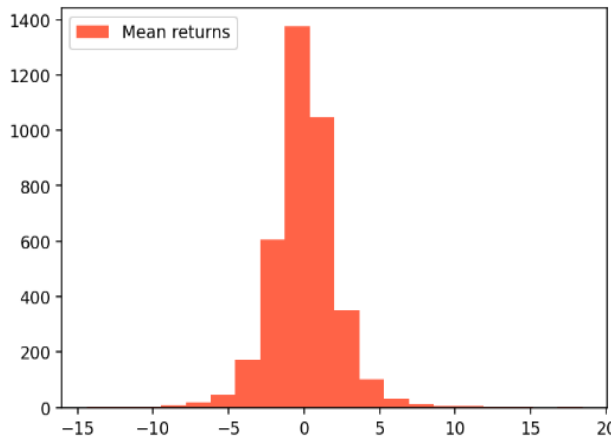
Biểu đồ cho thấy giá cổ phiếu SONY nằm trong khoảng từ 9.63 đến 128.59 USD. Vào năm 2008, bị ảnh hưởng bởi cuộc khủng hoảng kinh tế thế giới nên giá cổ phiếu của SONY giảm sút. Trong giai đoạn từ 2012-2018, giá cổ phiếu SONY vẫn chưa thể vượt được mức đỉnh trong quá khứ. Mặc dù dịch bệnh Covid-19 ảnh hưởng nhiều đến thị trường chứng khoán, giá cổ phiếu SONY lại khá khả quan. Từ năm 2022 đến hiện tại, giá có xu hướng tăng. Để thấy rõ độ biến động của giá cổ phiếu thì ta sẽ quan sát ở biểu đồ tỷ suất lợi nhuận.

- Tỷ suất sinh lời giá cổ phiếu SONY



Biểu đồ trên cho thấy, trong một số giai đoạn, tỷ suất sinh lời của SONY biến động cao hơn (cho nên rủi ro cũng sẽ cao hơn) so với các giai đoạn khác. Cụ thể là năm 2008, như đã nói ở trên, bị ảnh hưởng bởi cuộc khủng hoảng kinh tế thế giới nên lúc này tỷ suất sinh lời biến động khá mạnh và có tính tập trung mà không kéo dài mãi mãi. Có nghĩa là trong giai đoạn này, rủi ro

rất cao. Còn ở hiện tại, mặc dù có ảnh hưởng của dịch bệnh Covid-19, độ biến động khá nhỏ, chỉ có ở giữa năm 2020 là có biến động nhưng không lớn lắm và chỉ tập trung trong một thời gian ngắn. Các thay đổi lớn trong tỷ suất sinh lời được theo sau bởi những thay đổi lớn khác trước khi có xu hướng giảm xuống và ổn định trong một thời gian nhất định.



Daily volatility: 2.21%
 Monthly volatility: 9.87%
 Annual volatility: 34.88%

Ở biểu đồ phân phối trung bình tỷ suất lợi nhuận, có thể thấy rằng chuỗi phân phối tỷ suất sinh lời có dạng phân phối chuẩn và có kì vọng bằng 0.

Trong đó:

- Độ biến động hàng ngày là 2.21%
- Độ biến động hàng tháng là 9.87%
- Độ biến động hàng năm là 34.88%

Có thể thấy mức biến động hàng năm của cổ phiếu này là khá cao.

4.2.2 Kiểm định tính dừng và hiệu ứng ARCH

- Kiểm định tính dừng: sử dụng kiểm định Dickey-Fuller (ADF)

{ H0: Chuỗi tỷ suất sinh lời là chuỗi không dừng
 { H1: Chuỗi tỷ suất sinh lời là chuỗi dừng

```
def adf_test(x):
    indices = ['Test Statistic', 'p-value', '# of Lags Used', '# of Observations Used']
    adf_test = adfuller(x)
    results = pd.Series(adf_test[0:4], index=indices)
    for key, value in adf_test[4].items():
        results[f'Critical Value ({key})'] = value
    return results
adf_test(sony_data['returns'])
```

```
Test Statistic          -22.459831
p-value                  0.000000
# of Lags Used           7.000000
# of Observations Used   3774.000000
Critical Value (1%)      -3.432084
Critical Value (5%)      -2.862306
Critical Value (10%)     -2.567178
dtype: float64
```

Dựa vào bảng kết quả kiểm định, ta thấy:

- $p\text{-value} = 0.00000 < 0.05$
- Test Statistic có giá trị nhỏ hơn tất cả Critical Value ở cả 3 mức độ tin cậy 1%, 5%, 10%

Cho nên, bác bỏ H_0 , chấp nhận H_1 . Khi đó, ta khẳng định rằng chuỗi tỷ suất sinh lời có tính dừng.

- Kiểm định hiệu ứng ARCH: sử dụng kiểm định Lagrange Multiplier (LM)

{ H_0 : Chuỗi tỷ suất sinh lời không có hiệu ứng ARCH

{ H_1 : Chuỗi tỷ suất sinh lời có hiệu ứng ARCH

```
# Test by LM
arch_test = het_arch(sony_data['returns'])
print('LM: ', arch_test[0])
print('p_value: ', arch_test[1])
```

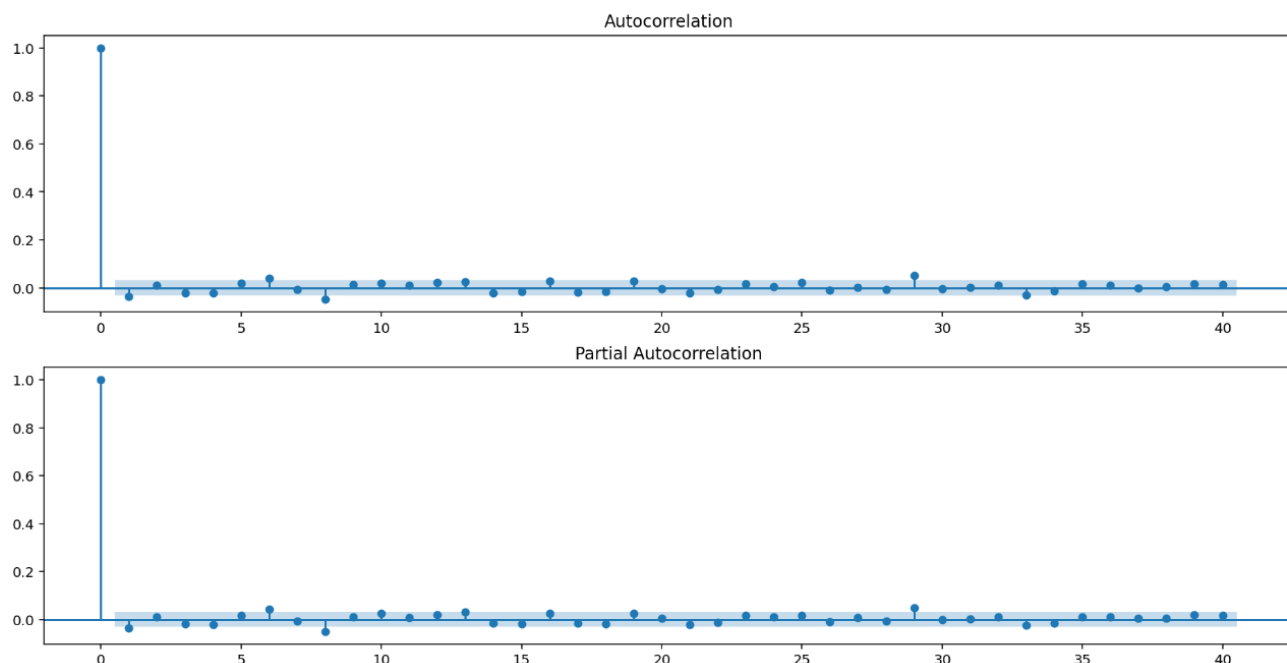
```
LM: 413.04973776718464
```

```
p_value: 6.410870403584773e-69
```

Dựa vào kết quả kiểm định trên, ta thấy p-value có giá trị rất thấp và nhỏ hơn 0.05. Cho nên, bác bỏ H_0 , chấp nhận H_1 . Khi đó, ta khẳng định rằng có hiệu ứng ARCH trong chuỗi dữ liệu tỷ suất sinh lời.

4.2.3 Thiết lập và ước lượng mô hình GARCH

- Lựa chọn tham số (p,q)



Dựa vào biểu đồ ACF và PACF, ta thấy p có thể nhận hai giá trị là 1, 2, 3, q cũng có thể nhận hai giá trị là 1, 2, 3. Cho nên, có thể chọn ra được 3 mô hình GARCH với tham số p, q : GARCH(1,1), GARCH(2,2) và GARCH(3,3).

- Thiết lập, ước lượng mô hình GARCH và chọn ra mô hình phù hợp
- Ước lượng mô hình GARCH(1,1):

Zero Mean - GARCH Model Results

```

=====
Dep. Variable:          returns      R-squared:          0.000
Mean Model:            Zero Mean    Adj. R-squared:     0.000
Vol Model:             GARCH        Log-Likelihood:     -8091.02
Distribution:          Normal       AIC:               16188.0
Method:               Maximum Likelihood BIC:             16206.8
                               No. Observations:         3782
Date:                 Mon, Jan 24 2022 Df Residuals:        3782
Time:                 20:08:46      Df Model:           0
                               Volatility Model
=====

```

```

=====
              coef      std err          t      P>|t|     95.0% Conf. Int.
-----
omega         0.3031      0.152       1.988  4.677e-02  [4.326e-03, 0.602]
alpha[1]      0.0953   3.305e-02       2.883  3.934e-03  [3.052e-02, 0.160]
beta[1]       0.8415   5.976e-02      14.082  4.880e-45  [ 0.724, 0.959]
=====

```

- Ước lượng mô hình GARCH(2,2)

```

                                Zero Mean - GARCH Model Results
=====
Dep. Variable:                returns    R-squared:                0.000
Mean Model:                  Zero Mean  Adj. R-squared:          0.000
Vol Model:                   GARCH      Log-Likelihood:         -8085.15
Distribution:                Normal     AIC:                   16180.3
Method:                      Maximum Likelihood BIC:                 16211.5
                                No. Observations:                3782
Date:                        Mon, Jan 24 2022 Df Residuals:        3782
Time:                        20:08:49      Df Model:                0
                                Volatility Model
=====
                                coef      std err          t      P>|t|      95.0% Conf. Int.
-----
omega                0.4068        1.059        0.384      0.701      [ -1.669,  2.483]
alpha[1]             0.1356     8.180e-02        1.658     9.739e-02 [-2.473e-02,  0.296]
alpha[2]             0.0000        0.369        0.000      1.000      [ -0.724,  0.724]
beta[1]              0.2825        2.404        0.118      0.906      [ -4.430,  4.995]
beta[2]              0.4975        1.892        0.263      0.793      [ -3.210,  4.205]
=====

```

- Ước lượng mô hình GARCH(3,3)

Zero Mean - GARCH Model Results

Dep. Variable:	returns	R-squared:	0.000
Mean Model:	Zero Mean	Adj. R-squared:	0.000
Vol Model:	GARCH	Log-Likelihood:	-8085.14
Distribution:	Normal	AIC:	16184.3
Method:	Maximum Likelihood	BIC:	16228.0
		No. Observations:	3782
Date:	Mon, Jan 24 2022	Df Residuals:	3782
Time:	20:08:50	Df Model:	0
Volatility Model			

	coef	std err	t	P> t	95.0% Conf. Int.
omega	0.5218	0.345	1.514	0.130	[-0.154, 1.197]
alpha[1]	0.1346	4.519e-02	2.979	2.895e-03	[4.604e-02, 0.223]
alpha[2]	0.0389	0.122	0.319	0.749	[-0.200, 0.277]
alpha[3]	8.0071e-13	4.483e-02	1.786e-11	1.000	[-8.787e-02, 8.787e-02]
beta[1]	0.0000	0.924	0.000	1.000	[-1.812, 1.812]
beta[2]	0.5884	0.528	1.114	0.265	[-0.447, 1.624]
beta[3]	0.1298	0.480	0.271	0.787	[-0.810, 1.070]

Như vậy, qua 3 bảng kết quả của 3 mô hình GARCH thì mô hình GARCH(2,2) có chỉ số AIC nhỏ nhất và mô hình GARCH(1,1) có chỉ số BIC nhỏ nhất. Vì mô hình đơn giản hơn sẽ được ưu tiên hơn nên mô hình GARCH(1,1) phù hợp để ước tính và đo lường biến động của chuỗi tỷ suất sinh lời cổ phiếu SONY. Kết quả cũng cho thấy các hệ số của mô hình đều nhận giá trị dương (đảm bảo yêu cầu của mô hình GARCH) và p-value có giá trị nhỏ hơn 0.05 thể hiện rằng các chỉ số trên có giá trị thống kê. Trong đó, omega là nhiễu trắng, alpha và beta là các tham số của mô hình. Từ bảng kết quả, ta thấy rằng $\alpha[1] + \beta[1] < 1$ biểu thị đây là một mô hình ổn định. Giá trị $\alpha[1] = 0.0953$ khá nhỏ có nghĩa là tác động biến động tức thời của cổ phiếu nhỏ. Trong khi đó, $\beta[1] = 0.8415$ khá lớn cho thấy thời gian tác động của biến động kéo dài hơn.

5. KẾT LUẬN

Dựa vào giá đóng cửa của cổ phiếu SONY được thu nhập từ ngày 1/1/2007 đến ngày 11/1/2022, kết hợp kiểm định các giả thiết và tiêu chí AIC, BIC để chọn ra mô hình phù hợp thì trong bài đã chỉ ra rằng mô hình GARCH(1,1) là phù hợp trong việc giải thích sự thay đổi của tỷ suất lợi nhuận của cổ phiếu Công ty công nghiệp Sony. Để đảm bảo điều kiện áp dụng mô hình GARCH, các

kiểm định về tính dừng và hiệu ứng ARCH đã được thực hiện. Mô hình nghiên cứu chỉ ra rằng tỷ suất lợi nhuận trong quá khứ cũng ảnh hưởng đến hiện tại và có thể là tương lai. Từ mô hình, có thể dự báo tỷ suất sinh lời trong tương lai để một phần nào đó giúp các nhà đầu tư lập kế hoạch đầu tư tránh các kết quả không mong muốn.