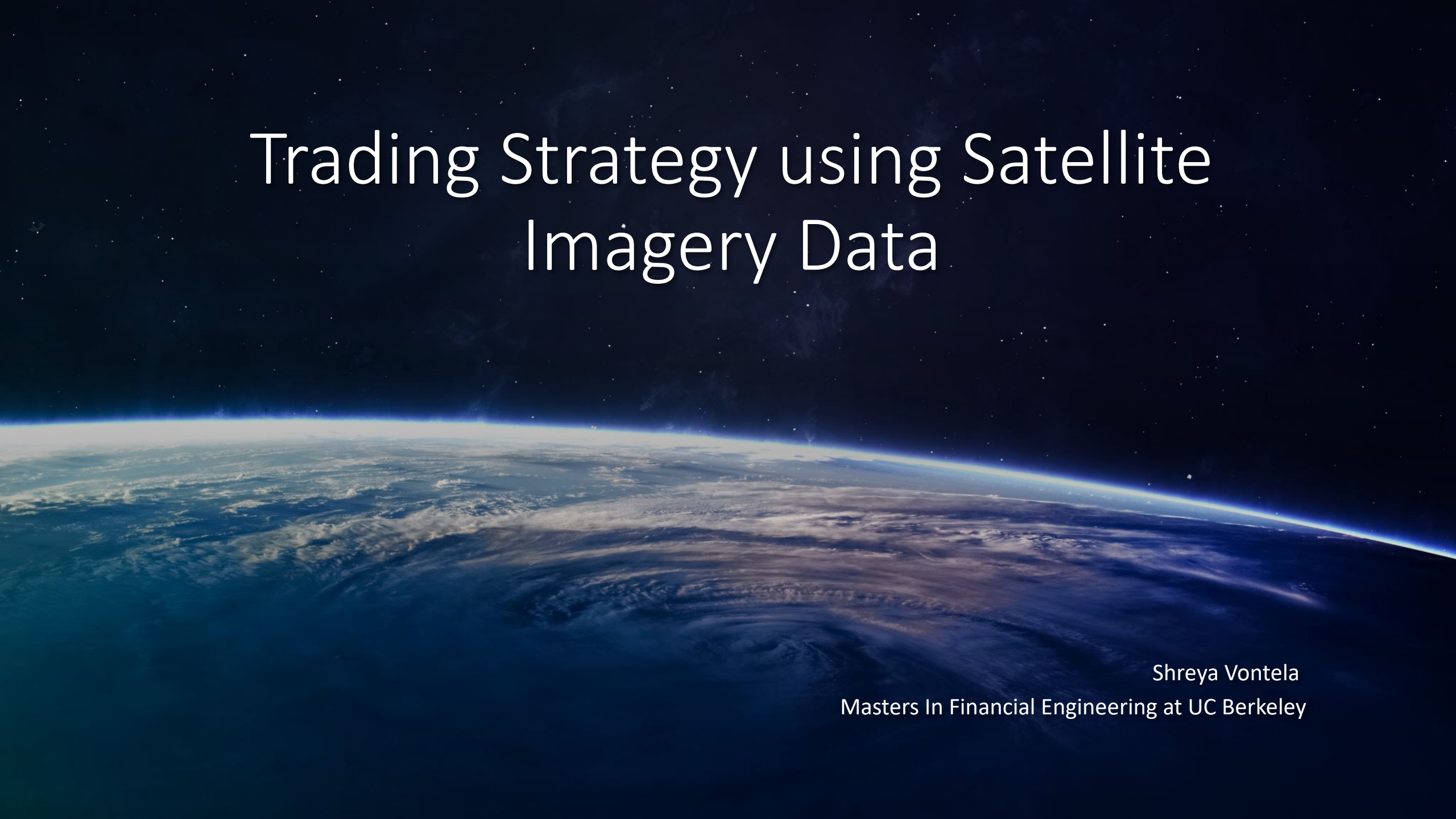
Trading Strategy using Satellite Imagery Data



Shreya Vontela
Masters In Financial Engineering at UC Berkeley

Problem Statement

Is it possible to use Satellite Imagery data to predict the stock price movement of a casino industry stock ?

Business Value

Geolocation data can help traders make key decisions and this increases the chances of getting positive returns

Methodology

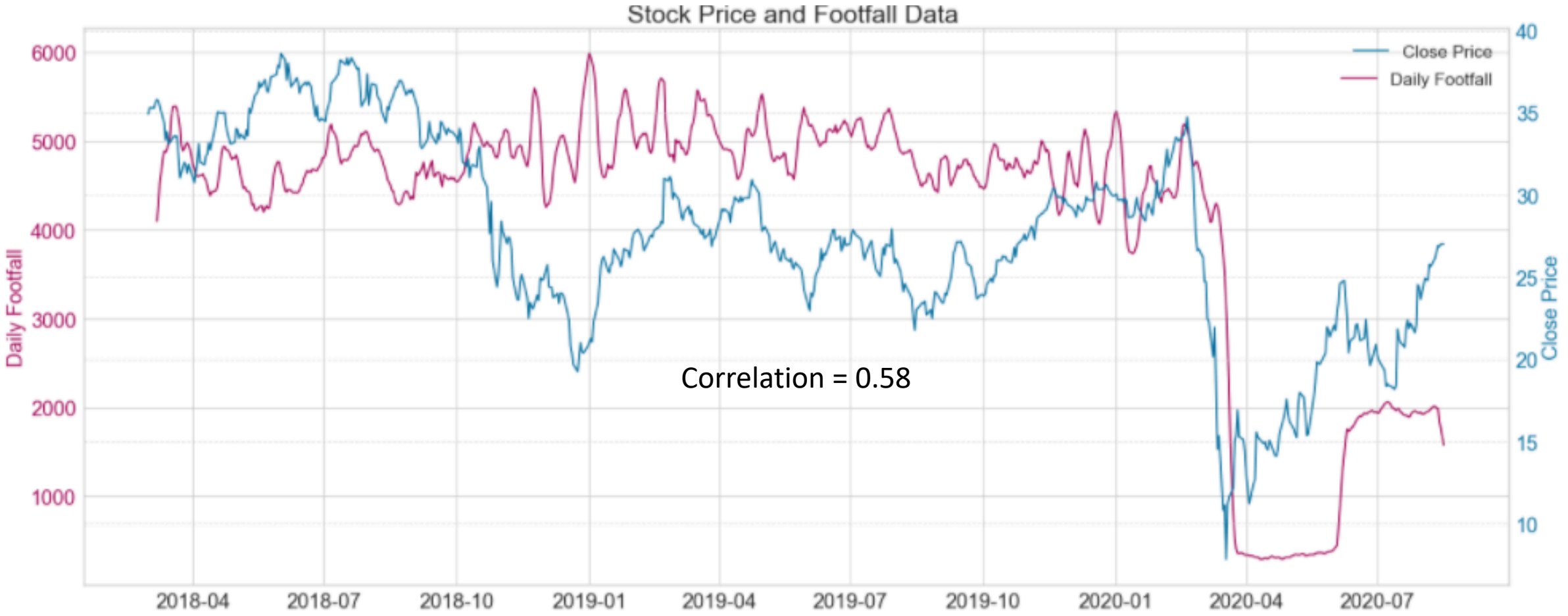
1. Data Collection
2. Exploratory Data Analysis
3. Feature Selection and Scaling
4. Modelling
    - Benchmark ARIMA model
    - Linear Regression model
    - Lasso Regression model
    - Neural Network model
5. Long only Trading strategy
6. Results
7. Conclusion

Data Collection

- Casino industry stock : Boyd Corp
- Google Earth to locate 9 casinos
- Collected footfall data from Kayrros platform

Footfall Data

	Casino1	Casino2	Casino3	Casino4	Casino5	Casino6	Casino7	Casino8	Casino9
2018-03-05	413	596	919	544	94	19	279	528	227
2018-03-06	527	749	1248	559	87	15	274	664	262
2018-03-07	514	843	985	615	108	21	326	696	246
2018-03-08	674	801	1125	653	79	22	342	769	284

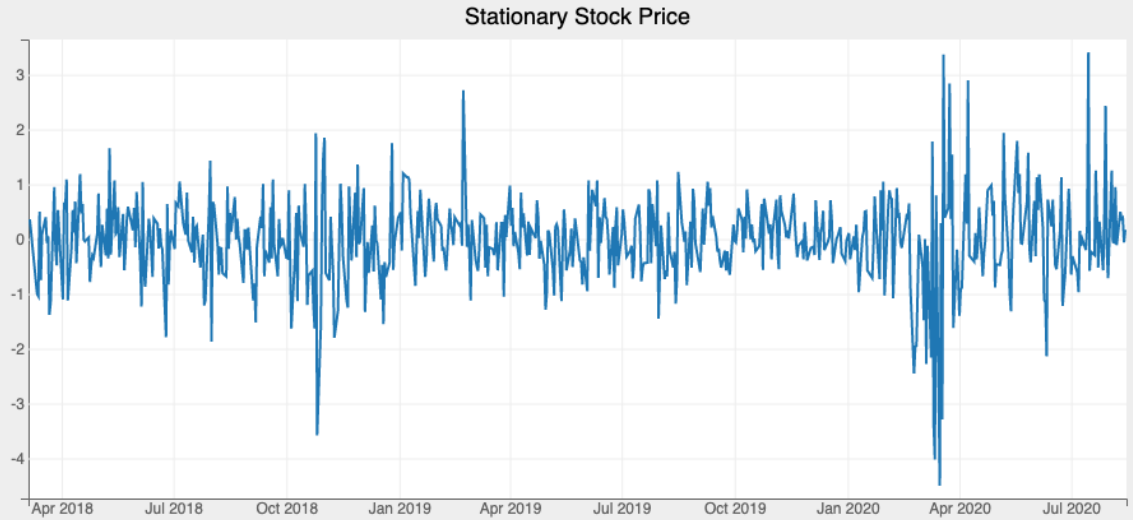


Data does not have high correlation but does follow similar trends.

Exploratory Data Analysis

Benchmark ARIMA Model

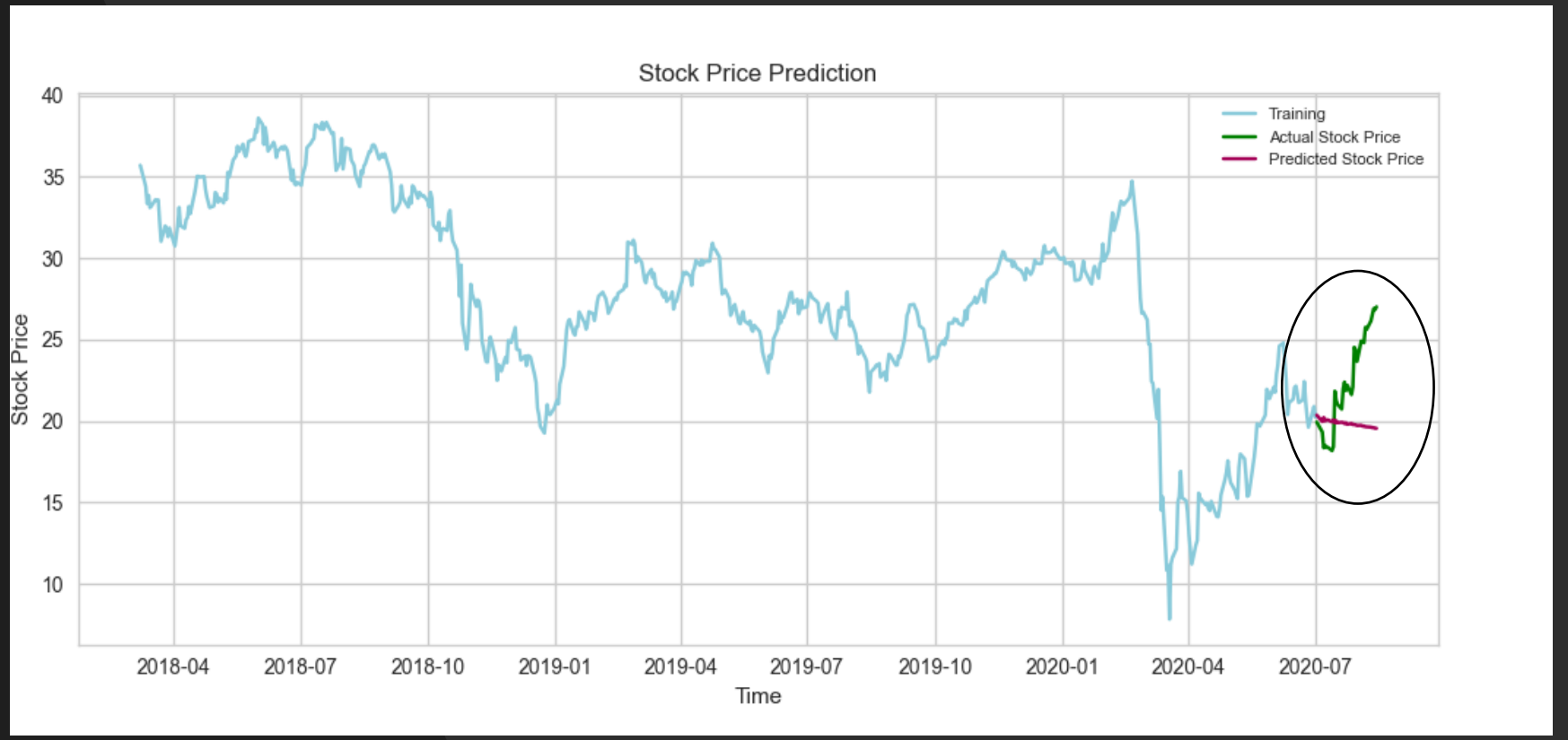
STEP 1) Difference the Series to achieve Stationarity




```
Results of dickey fuller test
Test Statistics                  -2.281535
p-value                           0.177992
No. of lags used                 10.000000
```

```
Results of dickey fuller test
Test Statistics                -7.675045e+00
p-value                         1.556873e-11
No. of lags used                8.000000e+00
```

Benchmark ARIMA model



STEP 2) Split the data into train and test set (90:10)

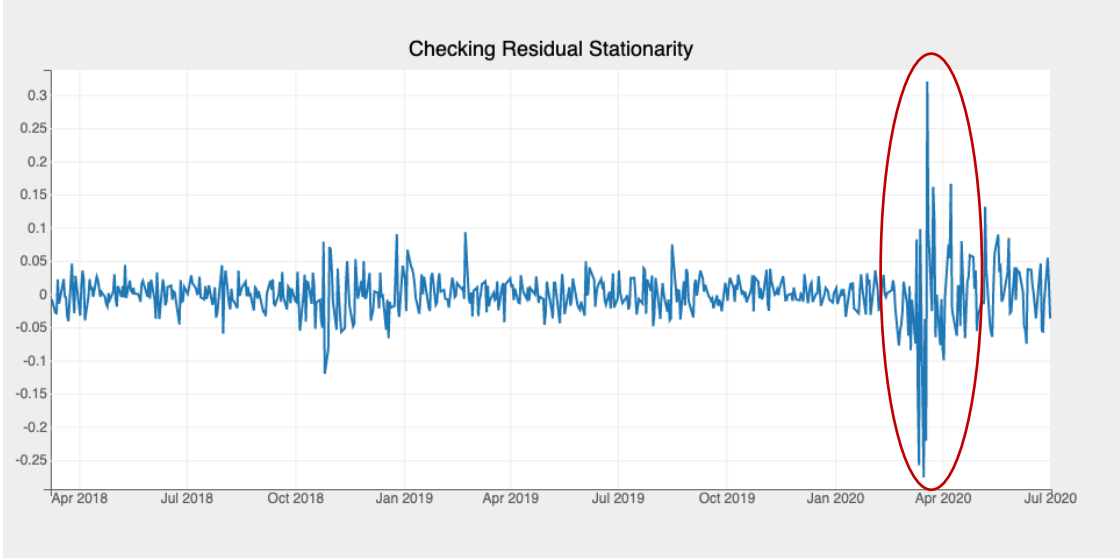
STEP 3) Implementing the appropriate ARIMA(3,1,3) model

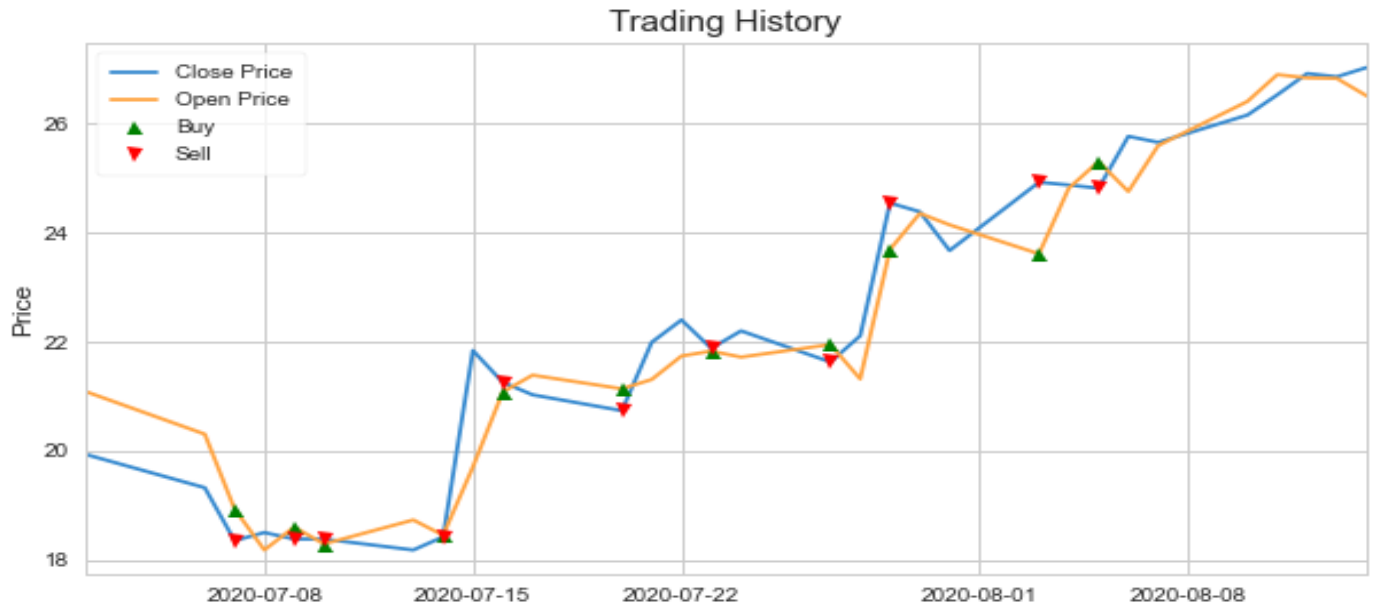
Performance Metrics : MSE is 16.7 and MAE is 3.4

Benchmark ARIMA model

STEP 4) Residual Analysis

STEP 5) Long Only Trading Strategy





Machine Learning Models Built using Footfall Data

1) Linear Regression Model
2) Lasso Regression Model
3) Neural Networks Model

Feature Selection

1) Technical Indicators:

RSI, Volume , MACD

2) Footfall Data :

Summation of footfall of all nine casinos

	Casino1	Casino2	Casino3	Casino4	Casino5	Casino6	Casino7	Casino8	Casino9	Total_Footfall
2018-03-05	413	596	919	544	94	19	279	528	227	3619
2018-03-06	527	749	1248	559	87	15	274	664	262	4385
2018-03-07	514	843	985	615	108	21	326	696	246	4354
2018-03-08	674	801	1125	653	79	22	342	769	284	4749

Feature Engineering

5 days technical indicators (5 X 3)

7 days Footfall Values (7 X 1)

Input vector to predict price of 8th Day (22 X 1)

	trend_macd	momentum_rsi	Volume
2018-03-12	-0.018297	22.116384	2354100.0
2018-03-13	-0.170868	13.495214	2500100.0
2018-03-14	-0.248571	28.022092	1757100.0
2018-03-15	-0.366446	22.042528	1278000.0
2018-03-16	-0.447442	24.134613	1432200.0

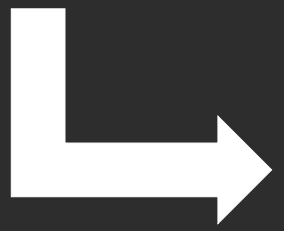
	total_foot
2018-03-12	4283.0
2018-03-13	4286.0
2018-03-14	4563.0
2018-03-15	5271.0
2018-03-16	6123.0
2018-03-17	7275.0
2018-03-18	5922.0

	22 vector Input
0	4.283000e+03
1	4.286000e+03
2	4.563000e+03
3	5.271000e+03
4	6.123000e+03
5	7.275000e+03
6	5.922000e+03
7	-1.829651e-02
8	-1.708682e-01
9	-2.485712e-01
10	-3.664460e-01
11	-4.474425e-01
12	2.211638e+01
13	1.349521e+01
14	2.802209e+01
15	2.204253e+01
16	2.413461e+01
17	2.354100e+06
18	2.500100e+06
19	1.757100e+06
20	1.278000e+06
21	1.432200e+06

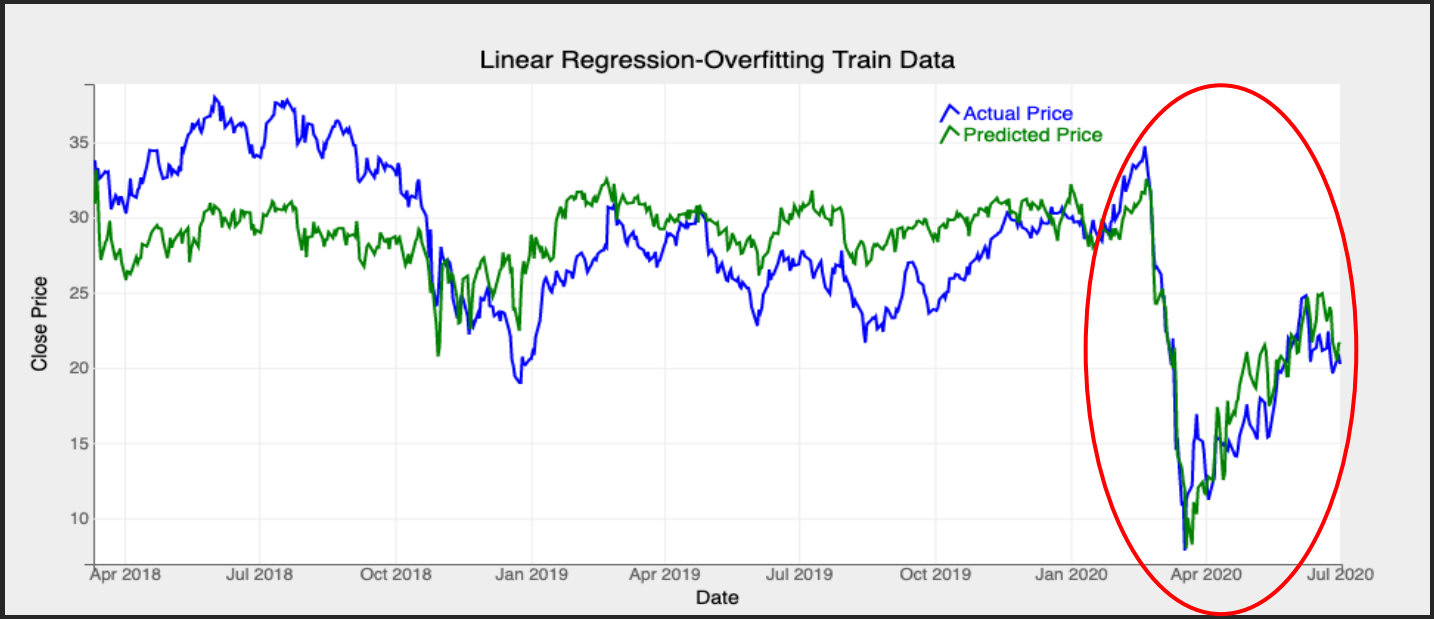
Handling Holidays

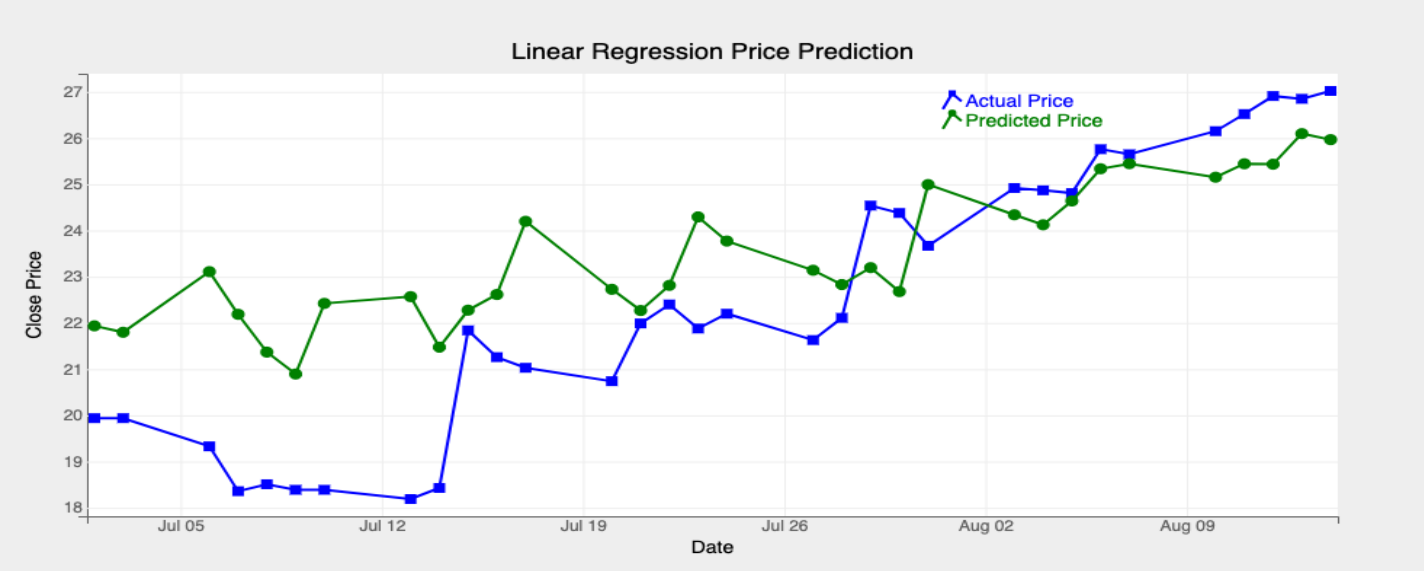
	trend_macd	momentum_rsi	Volume	Adj Close	total_foot
2018-03-28	-0.938731	27.522854	1087100.0	30.826263	4448
2018-03-29	-0.934802	34.235127	3020100.0	31.347742	4361
2018-03-30	NaN	NaN	NaN	NaN	5102
2018-03-31	NaN	NaN	NaN	NaN	5335
2018-04-01	NaN	NaN	NaN	NaN	4768
2018-04-02	-1.008820	28.363825	1128500.0	30.265427	4157

Footfall increased during Holidays



	trend_macd	momentum_rsi	Volume	Adj Close	total_foot
2018-03-28	-0.938731	27.522854	1087100.0	30.826263	4448.0
2018-03-29	-0.934802	34.235127	3020100.0	31.347742	4361.0
2018-03-30	-0.934802	34.235127	3020100.0	31.347742	5102.0
2018-03-31	NaN	NaN	NaN	NaN	5335.0
2018-04-01	NaN	NaN	NaN	NaN	4768.0
2018-04-02	-1.008820	28.363825	1128500.0	30.265427	4157.0

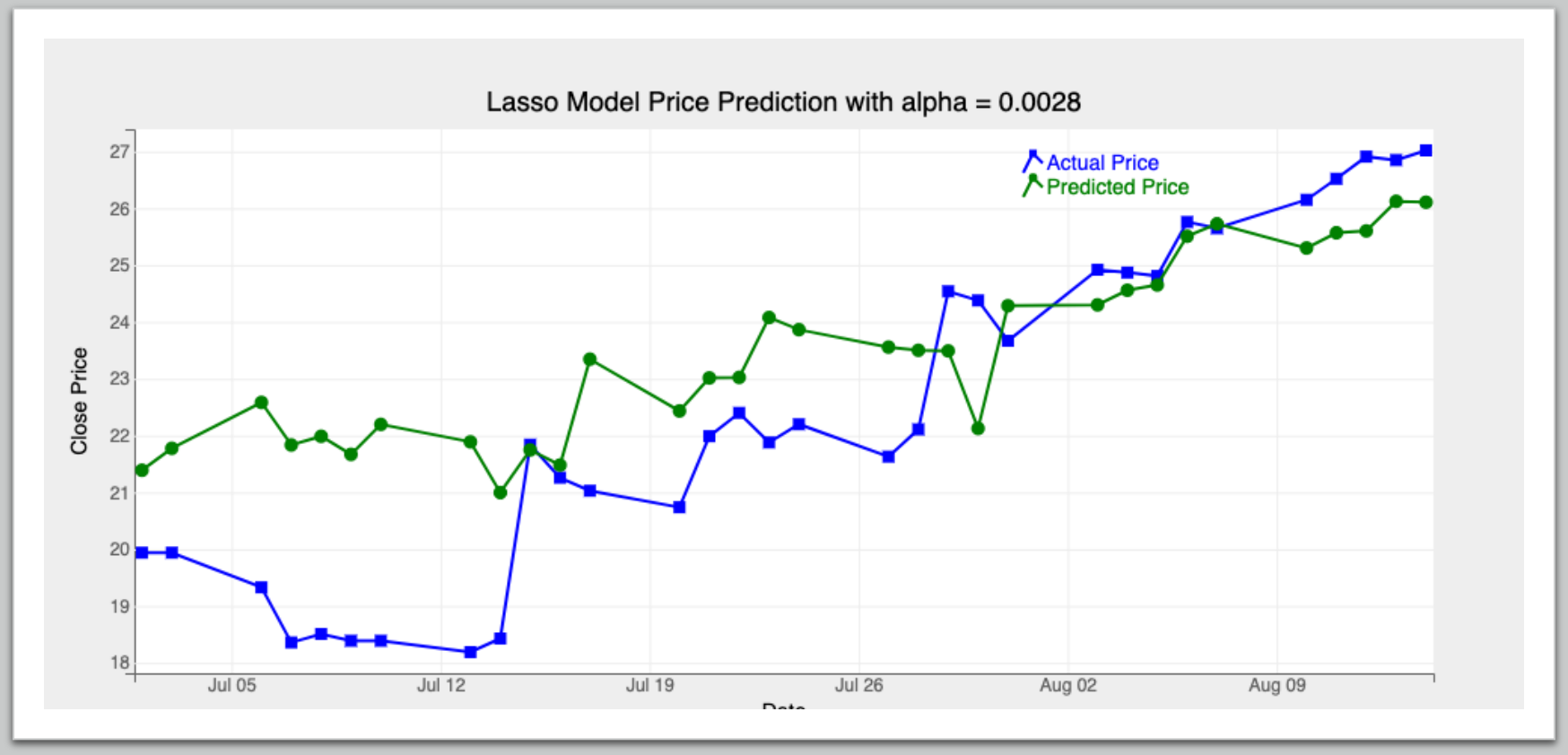




Linear Regression model

Train Data gives a MSE of 0.089 and Test data gives a MSE of 0.35.

So, the linear regression model is overfitting the train data.



Lasso Model

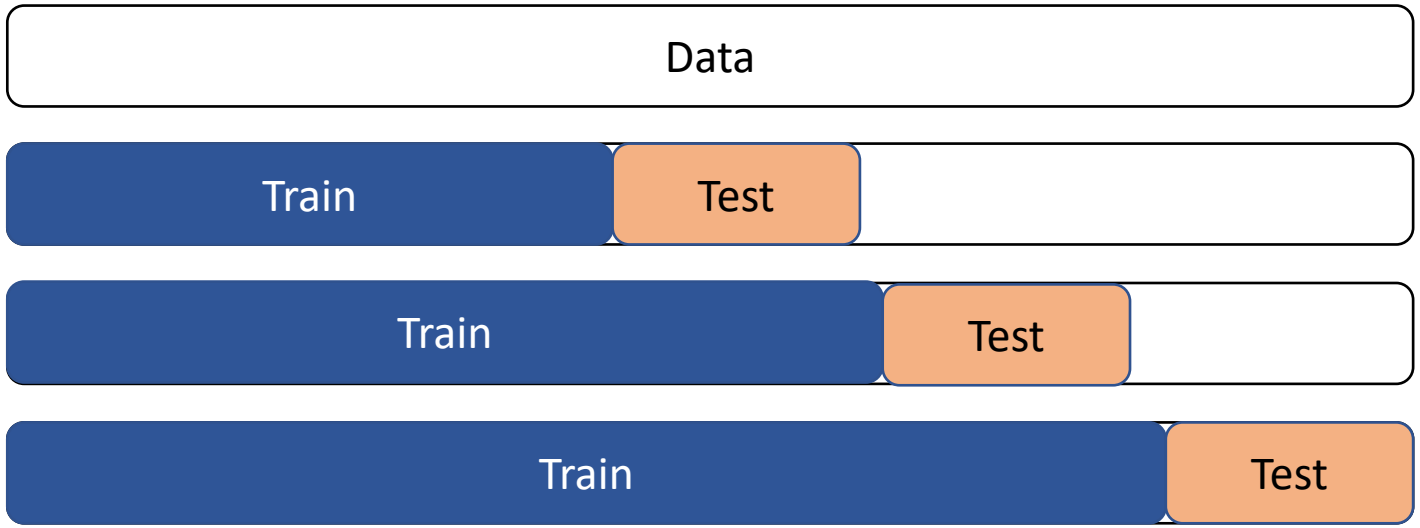
It has a marginal improvement over LR model.

MSE (Train) = 0.45 R squared (Train) = 0.53
MSE (Test) = 0.11. R squared (Test) = 0.55

Neural Networks Model

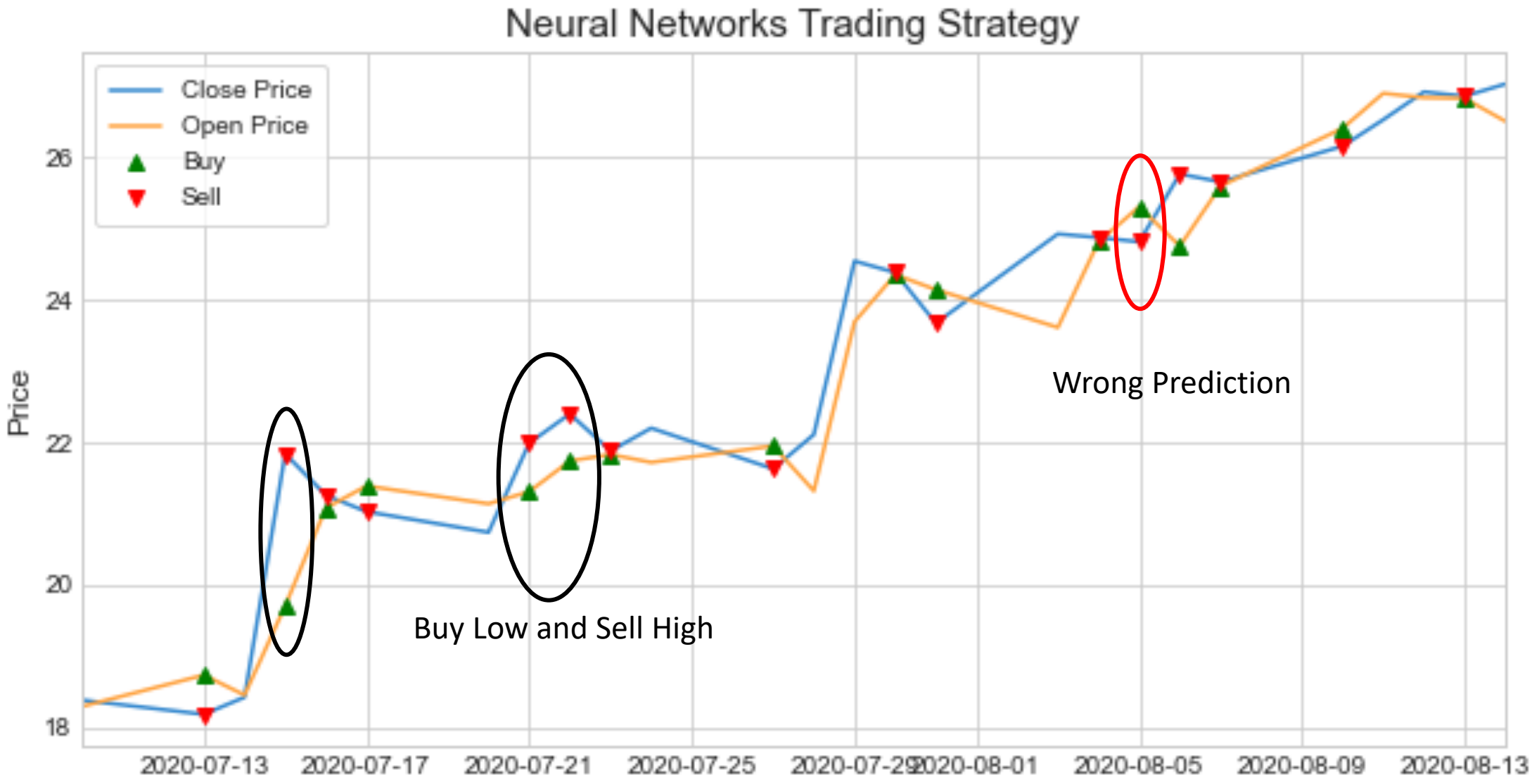
STEP 1)	STEP 2)	STEP 3)
Standardized data by making mean zero and variance one.	Run basic MLPRegressor model from Sklearn	Performed Hyper Parameter Tuning through Grid Search CV (3 fold Time Series Cross Validation Split)

Time Series CV in Tuning Hyper Parameters




```
model = MLPRegressor(random_state=1,
                     max_iter=500,
                     alpha=1,
                     hidden_layer_sizes = (100, 20, 10),
                     solver = 'sgd',
                     learning_rate = 'adaptive')
```

STEP 4)
Averaged the results by varying the random seed value


Neural Networks Trading Strategy
Close Price
Open Price
Buy
Sell
Price
26
24
22
20
18
2020-07-13
2020-07-17
2020-07-21
2020-07-25
2020-07-29
2020-08-01
2020-08-05
2020-08-09
2020-08-13
Buy Low and Sell High
Wrong Prediction

Results

Price Prediction model metrics

Mean Squared Error	Linear Regression	Lasso	Neural Networks
Technical Indicators model	0.35	0.34	0.25
Technical Indicators + Footfall	0.12	0.11	0.05

R squared score	Linear Regression	Lasso	Neural Networks
Technical Indicators model	-0.44	-0.37	-0.25
Technical Indicators + Footfall	0.49	0.55	0.76

Trading Strategy Metrics

Trading Strategy Returns	Linear Regression	Lasso	Neural Networks
Technical Indicators model	7.50%	7.50%	7.80%
Technical Indicators + Footfall	0.42%	9.10%	9.60%

Sharpe Ratio	Linear Regression	Lasso	Neural Networks
Technical Indicators model	1.10	1.15	1.20
Technical Indicators + Footfall	1.10	1.30	2.07

Benchmark Arima Model	
MSE	16.70
Trading Returns	1.70%

Results

Price Prediction model metrics

Mean Squared Error	Linear Regression	Lasso	Neural Networks
Technical Indicators model	0.35	0.34	0.25
Technical Indicators + Footfall	0.12	0.11	0.05

R squared score	Linear Regression	Lasso	Neural Networks
Technical Indicators model	-0.44	-0.37	-0.25
Technical Indicators + Footfall	0.49	0.55	0.76

Trading Strategy Metrics

Trading Strategy Returns	Linear Regression	Lasso	Neural Networks
Technical Indicators model	7.50%	7.50%	7.80%
Technical Indicators + Footfall	0.42%	9.10%	9.60%

Sharpe Ratio	Linear Regression	Lasso	Neural Networks
Technical Indicators model	1.10	1.15	1.20
Technical Indicators + Footfall	1.10	1.30	2.07

Benchmark Arima Model	
MSE	16.70
Trading Returns	1.70%

Conclusion

- Traders can benefit from having Footfall Data as an additional source, but the magnitude of the effect can be confirmed with additional data.
- The Lasso model and Neural Networks model run well

THANK YOU FOR YOUR TIME!