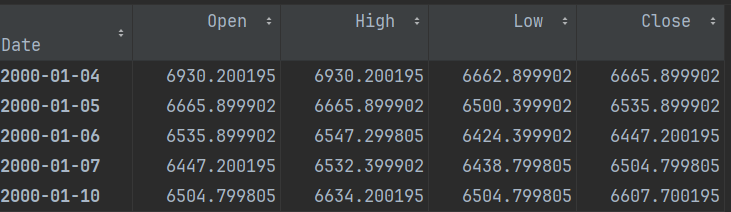
In this task I will use a logistic regression model in machine learning for trading and will try to predict stock price movement.

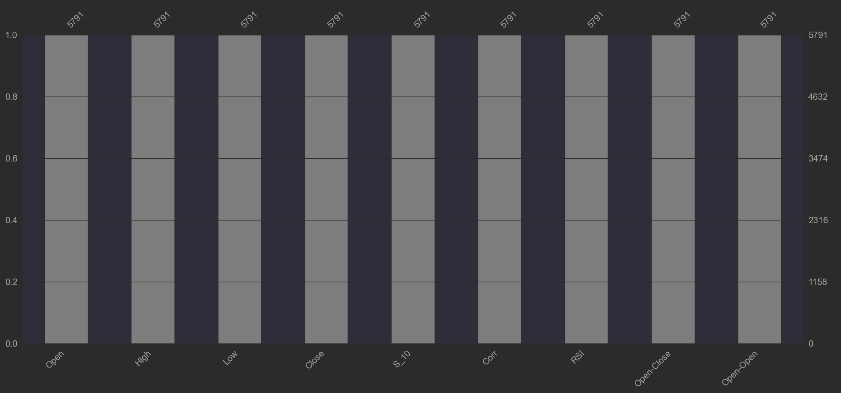
The Yahoo Source: FTSE 100 index for 2000 -2023 period.

Source Data:



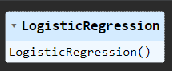
Manipulated Data:

We will use 10-days moving average, correlation, relative strength index (RSI), the difference between the open price of yesterday and today, difference close price of yesterday and the open price of today, open, high, low, and close price as indicators to make the prediction.



Model:

Logistic regression falls under the category of supervised learning; it measures the relationship between the categorical dependent variable and one or more independent variables by estimating probabilities using a logistic/sigmoid function. It is a classification problem which is used to predict a binary outcome (1/0, -1/1, True/False) given a set of independent variables.



Business task: predicting the stock price movement.

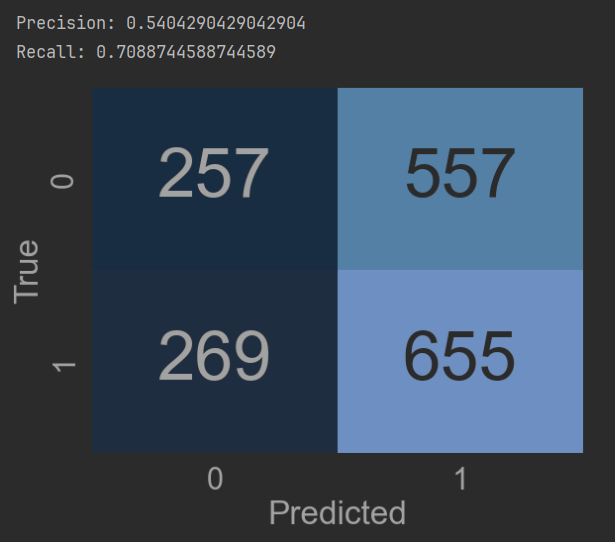
If tomorrow’s closing price is higher than todays closing price, then will buy stock, else we will sell it (-1).

If the output is 0.7, then we say there is a 70% chance that tomorrow’s closing price is higher than today’s closing price and classify it as 1.

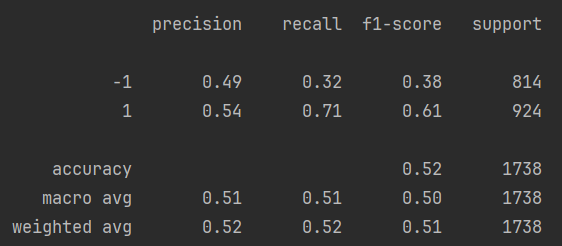
Business decision: buy or sell the stock

Evaluate The Model:

The Confusion matrix is used to describe the performance of the classification model on a set of test dataset for which the true values are known

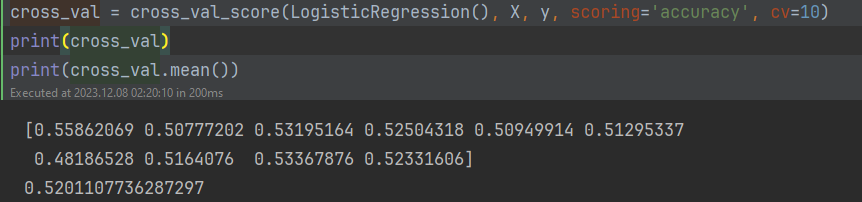


Classification Report is another method to examine the performance of the classification model.



The f1-score tells you the accuracy of the classifier in classifying the data points in that particular class compared to all other class. It is calculated by taking the harmonic mean of precision and recall. The support is the number of samples of the true response that lies in that class.

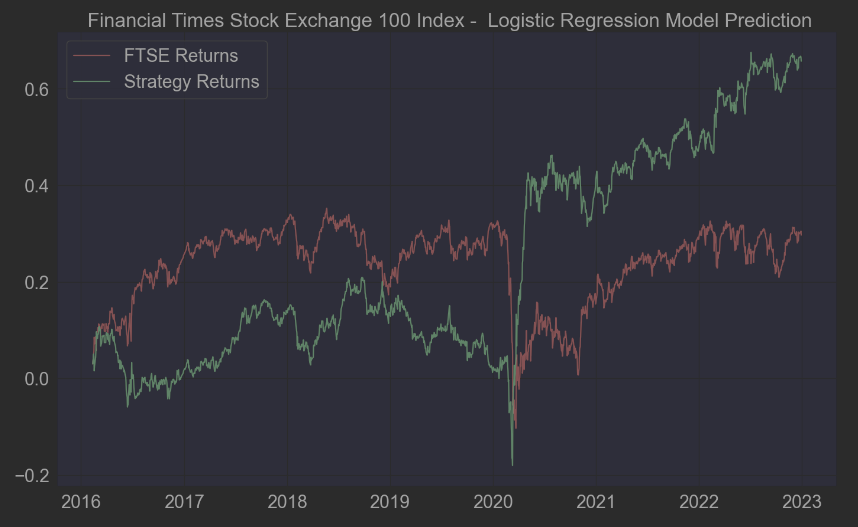
We cross-check the accuracy of the model using 10-fold cross-validation.



The accuracy is still 52% which means the model is working fine.

Basics of the Trading Strategy using the Model

We will predict the signal to buy (1) or sell (-1) and calculate the cumulative FTSE 100 returns for the test dataset. Next, we will calculate the cumulative strategy return based on the signal predicted by the model in the test dataset. We will also plot the cumulative returns.



It can be observed that the above Logistic Regression model predicts the classes with an accuracy of approximately 52% and generates good returns.