Solution Design:

Submission 3: Modeling and Strategy Development

Project Instructions:

Modeling

1. Decide on an algorithm or group of algorithms (for example, ensemble techniques).
2. Fit the model
3. Show that it works out of sample, and use appropriate cross-validation techniques.
4. Provide the following performance metrics:

    (a) ROC curves,

    (b) Confusion Matrix,

    (c) Precision, Recall, F1-Score, Accuracy, and AUC.

    5. Analysis of metrics and report

Fund factsheet

Create a fund factsheet for your new investment strategy.

Project Libraries Used:

Pandas==0.23.4

Numpy==1.15.1

Statsmodels==0.9.0

Matplotlib==2.2.2

Seaborn==0.9.0

Missingno==0.4.1

Scipy==1.1.0

Pathlib

Step 1:

Import a dataset "**clean_Google_features** " as a parq file from submission two (feature engineering)

Step2:

Develop a trading strategy of moving average of slow and fast windows of 12 and 26

Step3:

Used an advanced gradient boosting algorithm "**XGBoost**" for assembling technique. XGBoos has a high predictive power which enable us to make decision on the direction of the stock.

Step 4

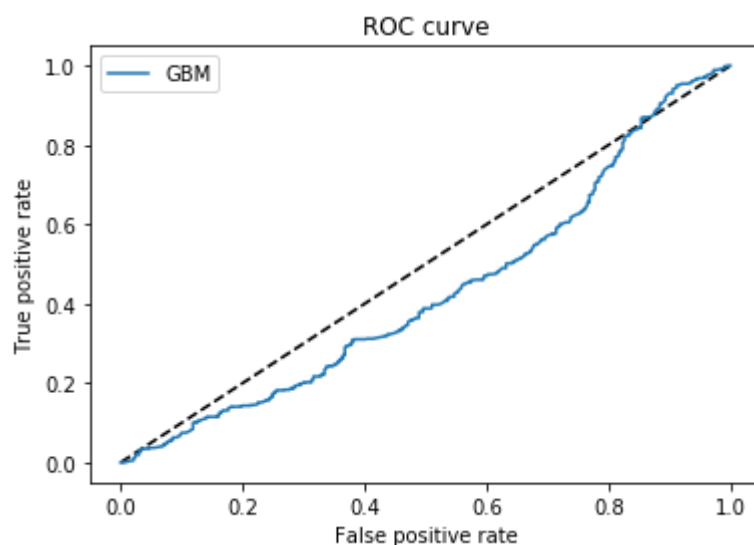Analyse the report and confusion matrix to measure the performance of the model.

Step 5

Create a fund factsheet for the equity "**$1,000**".

Report Analyst and Confusion Matrix

Performance Matrix:

	precision	recall	f1-score	support
no_trade	0.00	0.00	0.00	329
trade	0.69	1.00	0.82	747
avg /total	0.48	0.69	0.57	1076

What can we learn from this matrix?

- There are two possible predicted classes: "trade" and "no trade". If we were predicting the position of the stock in the market, for example, "trade" would mean that we have to trade, and "no trade" would mean we don't have to trade.
- The classifier made an average precision of 48% prediction and average recall of 69% prediction.
- Out of the total volting of 1076, 747 support us to trade while 329 support us not to trade.

Therefore, from our confusion matrix we are expected to trade Google stock in the market.

Fund factsheet

```
Stats for Equity from 2004-08-20 00:00:00 - 2018-11-15 00:00:00
```

Annual risk-free rate considered: 0.00%

Summary: Total Return   Sharpe   CAGR Max Drawdown

```
--------      0.30%          0.7      0.02%     -0.11%
```

Annualized Returns: mtd   3m   6m   ytd   1y   3y   5y   10y   incep.
```
------ ------ ------ ----- ----- ----- ----- ----- --------
```

```
               -0.00% -0.01% -0.00% 0.00% 0.00% 0.01% 0.01% 0.02% 0.02%
```

Periodic: daily   monthly   yearly

```
------          ------- --------- --------
```

	daily	monthly	yearly
sharpe	0.70	0.70	0.44
mean	0.02%	0.02%	0.02%
vol	0.03%	0.03%	0.04%
skew	0.57	0.52	-0.88
kurt	9.87	1.87	2.37
best	0.02%	0.04%	0.08%
worst	-0.01%	-0.02%	-0.08%

Drawdowns:

max	avg	# days
-0.11%	-0.01%	43.09

Monthly Returns

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	YTD
2004	0	0	0	0	0	0	0	-0.01	0.02	0.04	-0	0.01	0.06
2005	0	-0	-0	0.02	0.02	0.01	-0	-0	0.01	0.02	0.01	0	0.08
2006	0	-0.02	0.01	0.01	-0.01	0.01	-0.01	-0	0.01	0.02	0	-0.01	0.01
2007	0.01	-0.01	0	0	0.01	0	-0	0	0.01	0.02	-0	-0	0.04
2008	-0.02	-0.02	-0.01	0.03	0	-0.01	-0.01	-0	-0.01	-0.01	-0.02	0	-0.08
2009	0.01	-0	0	0.01	0.01	0	0	0	0.01	0.01	0.01	0.01	0.07
2010	-0.02	-0	0.01	-0.01	-0.01	-0.01	0.01	-0.01	0.02	0.02	-0.01	0.01	-0
2011	0	0	-0	-0.01	-0	-0	0.02	-0.01	-0	0.01	0	0.01	0.01
2012	-0.01	0.01	0	-0.01	-0	-0	0.01	0.01	0.01	-0.01	0	0	0.01
2013	0.01	0.01	-0	0	0.01	0	0	-0	0	0.02	0	0.01	0.05
2014	0.01	0	-0.01	-0	0.01	0	-0	0	0	-0	-0	-0	-0.01
2015	0	0	-0	-0	-0	-0	0.02	-0	-0	0.01	0	0	0.04
2016	-0	-0.01	0.01	-0.01	0.01	-0.01	0.01	-0	0	0	-0	0	0
2017	0	0	0	0.01	0.01	-0.01	0	0	0	0.01	0	0	0.03
2018	0.01	-0.01	-0.01	-0	0.01	0	0.01	0	-0	-0.01	-0	0	0

Equity Curve

Equity Price Series
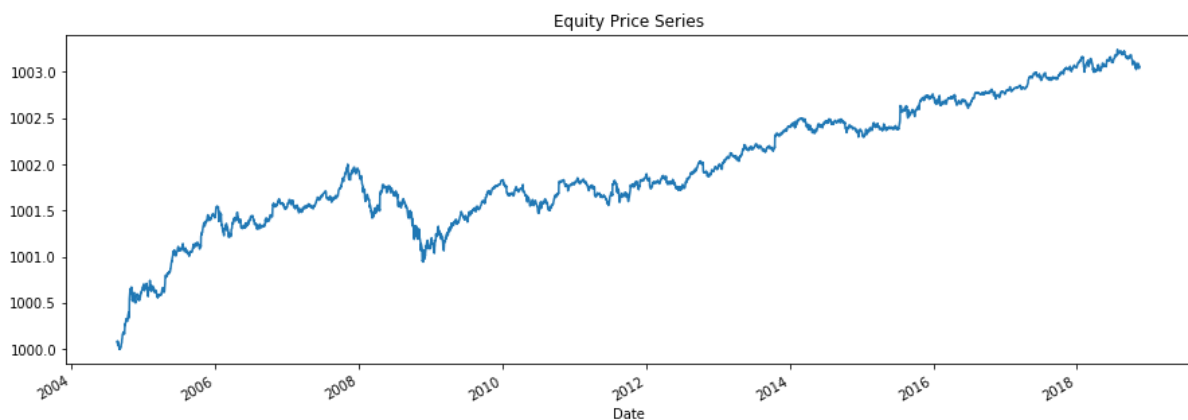
Drawdown
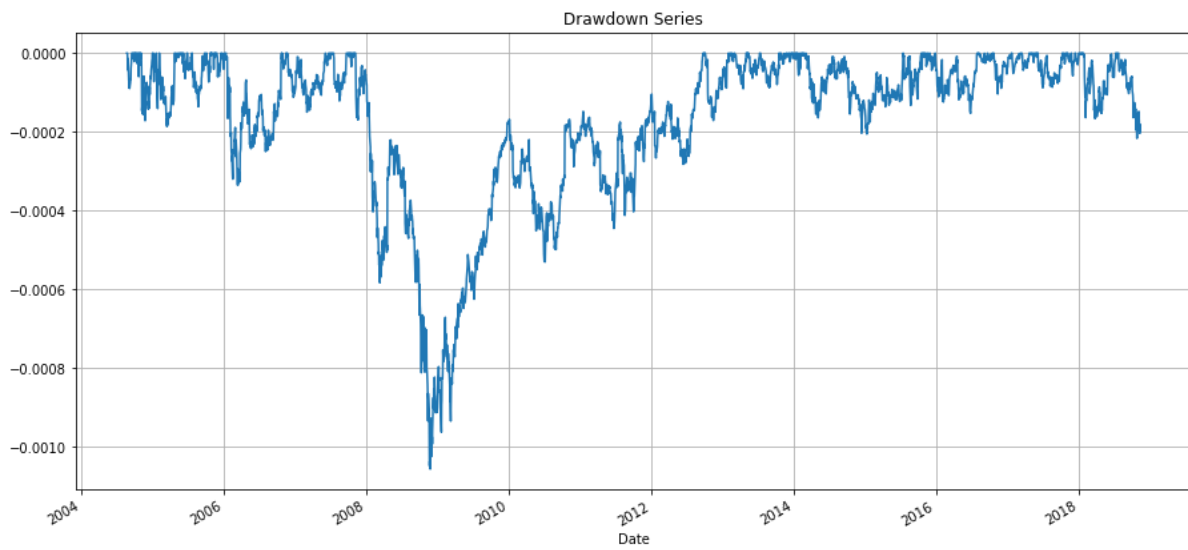
Artifacts:

The ReadMe file was written in Jupyter (GitHub Flavored Markdown) and saved in md file.

The code was run on Windows system.

The png files generated are included in this document above.