

FRA MILESTONE 2-BUSINESS REPORT

This Business Report shall provide detailed explanation of how we approached each problem given in the assignment. It shall also provide relative resolution and explanation with regards to the problems

PRABHJOT KAUR CHAHAL (PGP-DSBA -Online Mar_C 2021)

2021

Contents

MILESTONE:2:Credit Risk	3
1.8 Build a Random Forest Model on Train Dataset. Also showcase your model building approach	4
1.9 Validate the Random Forest Model on test Dataset and state the performance matrices. Also state interpretation from the model	4
1.10 Build a LDA Model on Train Dataset. Also showcase your model building approach	9
1.11 Validate the LDA Model on test Dataset and state the performance matrices. Also state interpretation from the model	9
1.12 Compare the performances of Logistics, Radom Forest and LDA models (include ROC Curve)	13
1.13 State Recommendations from the above models.	15
PROBLEM 2:MARKET RISK.....	16
1.1 Draw Stock Price Graph (Stock Price vs Time) for any 2 given stocks with inference.	17
1.2 Calculate Returns for all stocks with inference.....	18
2.3 Calculate Stock Means and Standard Deviation for all stocks with inference.....	20
2.4 Draw a plot of Stock Means vs Standard Deviation and state your inference	21
2.5 Conclusion and Recommendations.....	22

MILESTONE:2:Credit Risk

Businesses or companies can fall prey to default if they are not able to keep up their debt obligations. Defaults will lead to a lower credit rating for the company which in turn reduces its chances of getting credit in the future and may have to pay higher interests on existing debts as well as any new obligations. From an investor's point of view, he would want to invest in a company if it is capable of handling its financial obligations, can grow quickly, and is able to manage the growth scale.

A balance sheet is a financial statement of a company that provides a snapshot of what a company owns, owes, and the amount invested by the shareholders. Thus, it is an important tool that helps evaluate the performance of a business.

Data that is available includes information from the financial statement of the companies for the previous year (2015). Also, information about the Networth of the company in the following year (2016) is provided which can be used to drive the labeled field.

Explanation of data fields available in Data Dictionary, 'Credit Default Data Dictionary.xlsx'

Hints :

Dependent variable - We need to create a default variable which should take the value of 1 when net worth next year is negative & 0 when net worth is positive. Test Train Split - Split the data into Train and Test dataset in a ratio of 67:33 and use random_state=42. Model Building is to be done on Train Dataset and Model Validation is to be done on Test Dataset.

OBJECTIVE

We have a dataset which has financial statement of the companies for year 2015 and net worth of the company in 2016

We need to derive the predictions of the net worth of the companies to help the investors make the correct choice

APPROACH

We have done pre-processing of the dataset (Treating missing values, outliers).Now will work with different models and compare them with the help of python.

1.8 Build a Random Forest Model on Train Dataset. Also showcase your model building approach

1.9 Validate the Random Forest Model on test Dataset and state the performance matrices. Also state interpretation from the model

I have performed the train – test split as mentioned in the question and performed the random forest model on the train data.

I created two models using Random Forest- without SMOTE data and with SMOTE data.

I have not scaled the data for Random Forest as tree based models are not distance based models and can handle varying ranges of features.

We also used gridsearchcv for hyper parameter tuning.

Below are the performance metrics on **train data** for random forest model **without smote data**.

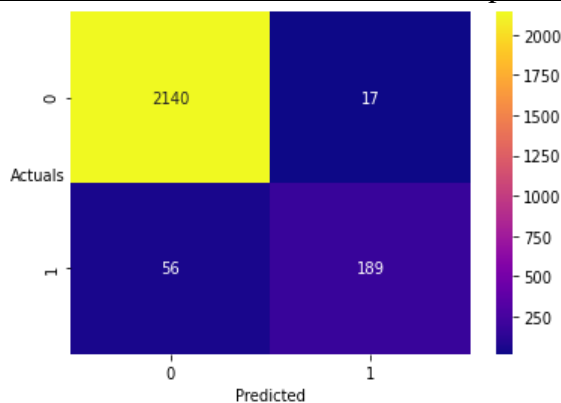
- Best Grid parameters:

```
{'max_depth': 20,  
'min_samples_leaf': 30,  
'min_samples_split': 90,  
'n_estimators': 100}
```

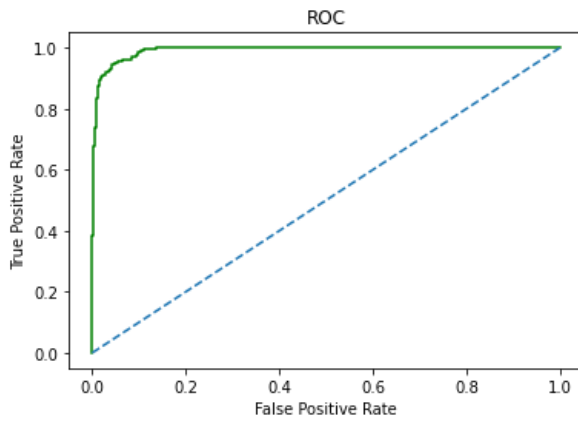
- **Accuracy of Train Data = 97.0**

Checking the accuracy of the model using confusion matrix for training set:

- Confusion matrix and classification report and ROC Curve for train data



	precision	recall	f1-score	support
0.0	0.97	0.99	0.98	2157
1.0	0.92	0.77	0.84	245
accuracy			0.97	2402
macro avg	0.95	0.88	0.91	2402
weighted avg	0.97	0.97	0.97	2402



Below are the performance metrics on **train data** for random forest model **with smote data**.

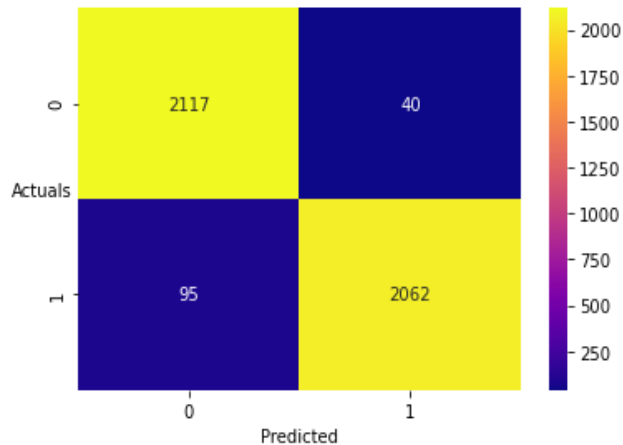
- Best Grid parameters:

```
{'max_depth': 30,
 'min_samples_leaf': 30,
 'min_samples_split': 90,
 'n_estimators': 100}
```

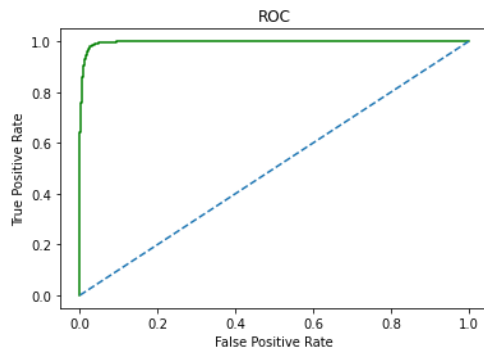
- Accuracy of Train Data using smote = 96.87

Checking the accuracy of the model using confusion matrix for training set:

- Confusion matrix and classification report and ROC Curve for train data



	precision	recall	f1-score	support
0.0	0.96	0.98	0.97	2157
1.0	0.98	0.96	0.97	2157
accuracy			0.97	4314
macro avg	0.97	0.97	0.97	4314
weighted avg	0.97	0.97	0.97	4314

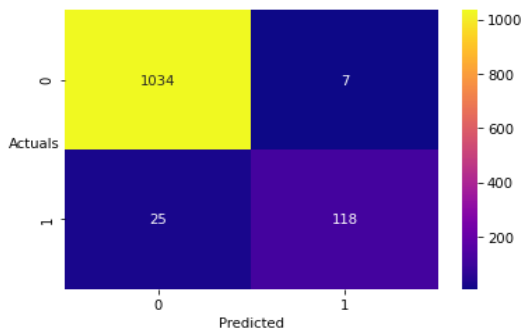


Both the random forest models were then evaluated on the test data.
Below are the results for **test data** for random forest model **without smote data**.

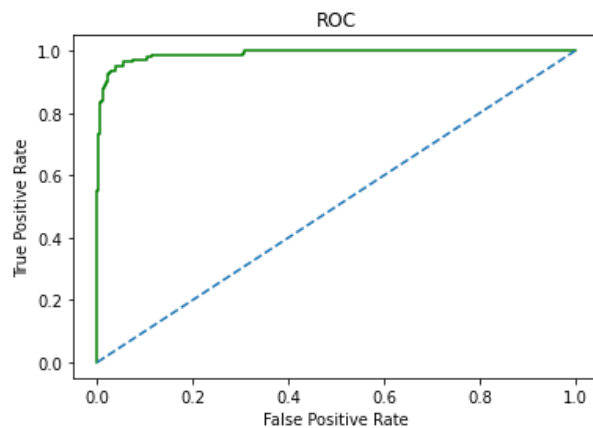
- Accuracy of Test Data = 97.3

Checking the accuracy of the model using confusion matrix for testing set:

- Confusion matrix and classification report and ROC Curve for test data



	precision	recall	f1-score	support
0.0	0.98	0.99	0.98	1041
1.0	0.94	0.83	0.88	143
accuracy			0.97	1184
macro avg	0.96	0.91	0.93	1184
weighted avg	0.97	0.97	0.97	1184

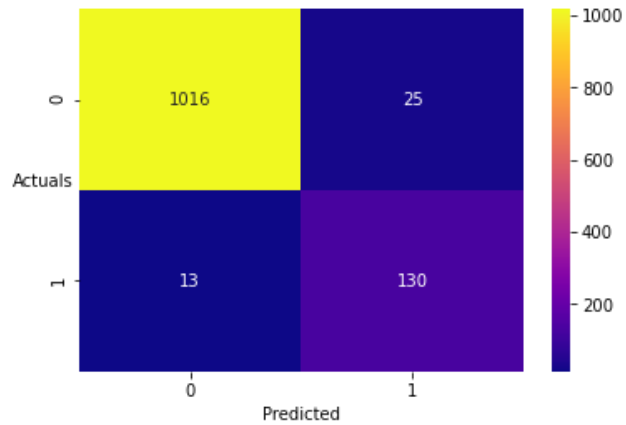


Below is the performance metrics on **test data** for random forest model **with smote**

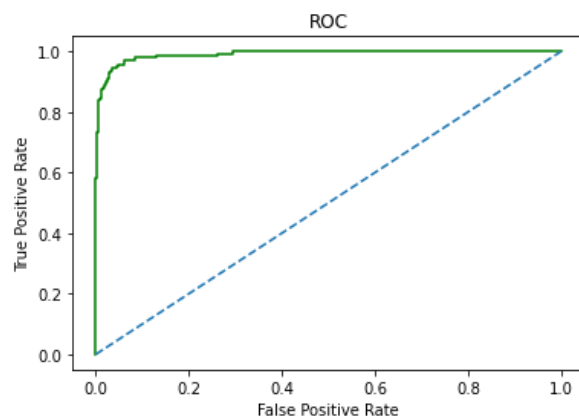
- Accuracy of Test Data using smote = 96.78999999999999

Checking the accuracy of the model using confusion matrix for testing set:

- Confusion matrix and classification report and ROC Curve for test data



	precision	recall	f1-score	support
0.0	0.99	0.98	0.98	1041
1.0	0.84	0.91	0.87	143
accuracy	0.97			1184
macro avg	0.91	0.94	0.93	1184
weighted avg	0.97	0.97	0.97	1184



Comparison of various evaluation metrics in tabular format for both the models i.e. with and without smote.

	Accuracy	AUC	Recall	Precision	F1 Score
Random Forest Train	0.97	0.99	0.77	0.92	0.84
Random Forest Test	0.97	0.99	0.83	0.94	0.88
Random Forest Train with SMOTE	0.97	1.00	0.96	0.98	0.97
Random Forest Test with SMOTE	0.97	0.99	0.91	0.84	0.87

According to the above comparison, even though Random Forest without SMOTE has similar accuracy as of with SMOTE, but Random Forest with SMOTE surpasses the other model significantly in terms of Recall and AUC score.

Also there is not a significant difference in the accuracies of both the models.

Therefore, we can say **Random Forest with SMOTE data** is the better of the two models with 97% accuracy, 84% precision, 91% recall and F1 score of 0.87 respectively.

1.10 Build a LDA Model on Train Dataset. Also showcase your model building approach

1.11 Validate the LDA Model on test Dataset and state the performance matrices. Also state interpretation from the model

I tried building the Linear Discriminant Analysis model on the same train data as above.

I have built the model without any parameters passed on the model; thereby the model will consider all of its default parameters while fitting the model.

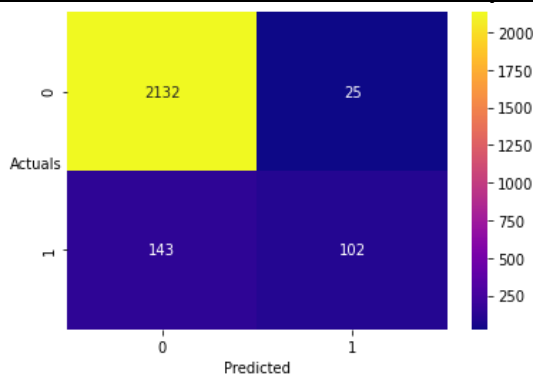
I have created two models using LDA, one without SMOTE data and one with SMOTE data. I have not scaled the data for LDA as it finds its coefficients using the variation between the classes, hence scaling doesn't matter.

Below are the performance metrics on **train data** for LDA model **without smote data**.

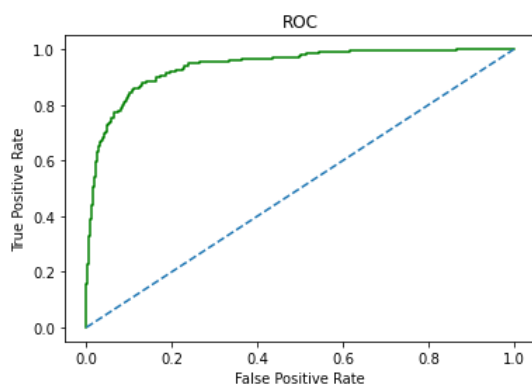
Accuracy of Train Data = 93.0

Checking the accuracy of the model using confusion matrix for training set:

- Confusion matrix and classification report and ROC Curve for train data



	precision	recall	f1-score	support
0.0	0.94	0.99	0.96	2157
1.0	0.80	0.42	0.55	245
accuracy			0.93	2402
macro avg	0.87	0.70	0.76	2402
weighted avg	0.92	0.93	0.92	2402

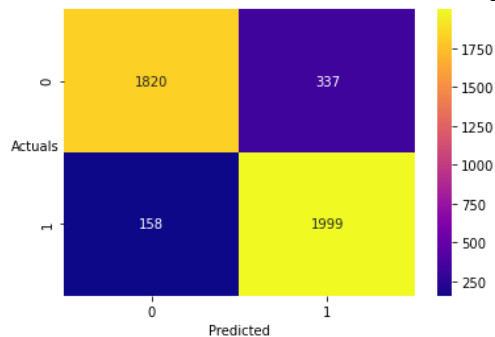


Below are the performance metrics on **train data** for LDA model **with smote data**.

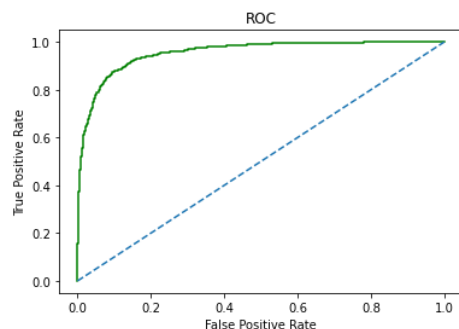
- Accuracy of Train Data using smote = 88.53

Checking the accuracy of the model using confusion matrix for training set:

- Confusion matrix and classification report and ROC Curve for train data



	precision	recall	f1-score	support
0.0	0.92	0.84	0.88	2157
1.0	0.86	0.93	0.89	2157
accuracy			0.89	4314
macro avg	0.89	0.89	0.89	4314
weighted avg	0.89	0.89	0.89	4314



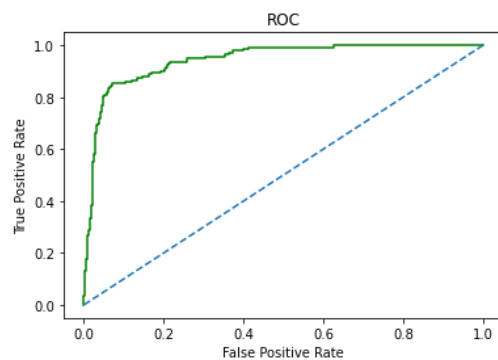
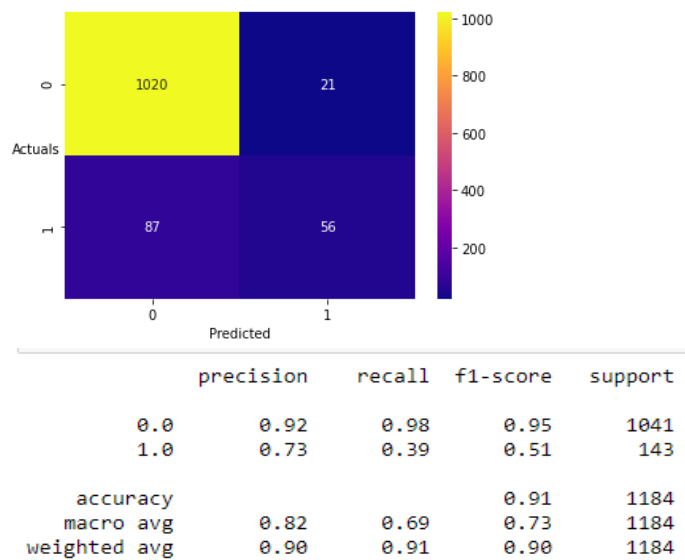
Both the LDA models were then evaluated and validated on the test data.

Below are the results for **test data** for LDA model **without smote data**.

- Accuracy of Test Data = 90.88000000000001

Checking the accuracy of the model using confusion matrix for testing set:

- Confusion matrix and classification report and ROC Curve for test data

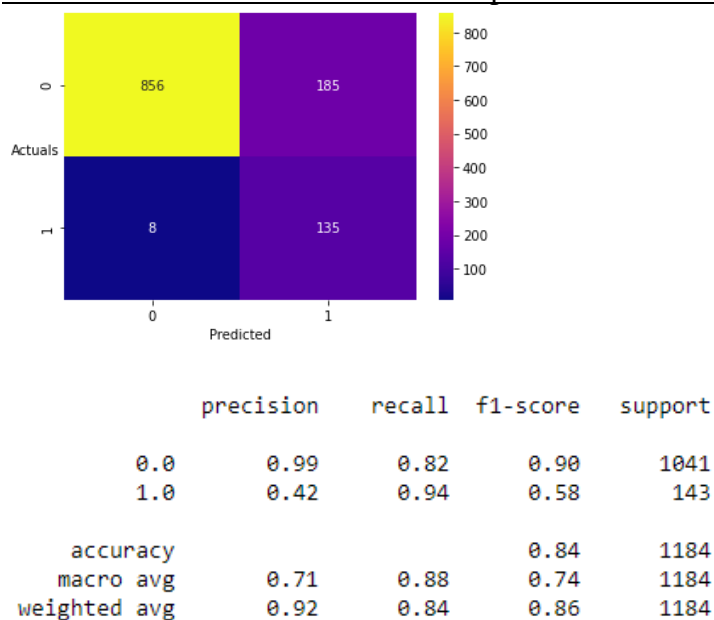


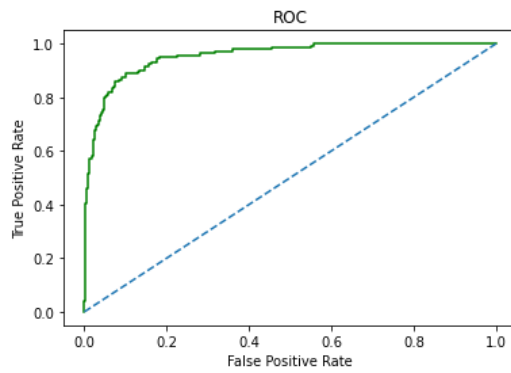
Below are the performance metrics on **test data** for LDA model **with smote data**.

- Accuracy of Test Data using smote = 83.7

Checking the accuracy of the model using confusion matrix for testing set:

- Confusion matrix and classification report and ROC Curve for test data





Below is the comparison of various evaluation matrices in tabular format for both the models i.e. with and without smote.

	Accuracy	AUC	Recall	Precision	F1 Score
LDA Train	0.93	0.94	0.42	0.80	0.55
LDA Test	0.91	0.94	0.39	0.73	0.51
LDA Train with SMOTE	0.89	0.96	0.93	0.86	0.89
LDA Test with SMOTE	0.84	0.95	0.94	0.42	0.58

Analyzing and comparing the two models, we can observe LDA model without SMOTE data has higher accuracy on test data. However it has really bad precision, recall and F1 scores compared to the other model.

On purely accuracy perspective model without smote data performs better. But in real world scenario model with smote will perform better given better recall and precision values.

1.12 Compare the performances of Logistics, Radom Forest and LDA models (include ROC Curve)

We made following models as part of this exercise:

- Logistic regression using statsModel without SMOTE
- Logistic regression using statsModel with SMOTE
- Random Forest using sklearn without SMOTE
- Random Forest using sklearn with SMOTE
- LDA using sklearn without SMOTE
- LDA using sklearn with SMOTE

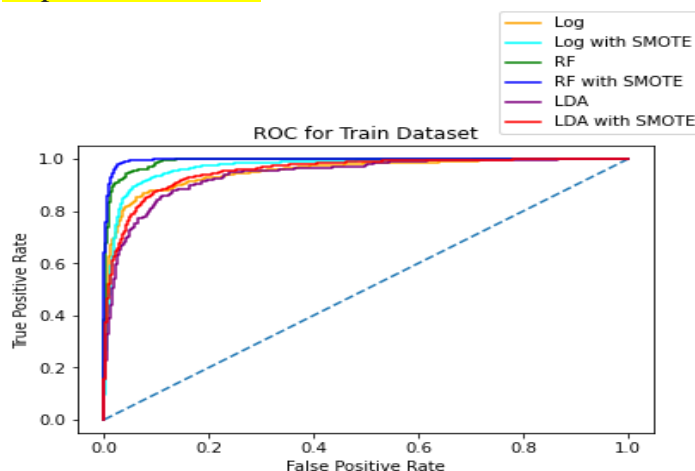
Comparison chart for all the models:

	Accuracy	AUC	Recall	Precision	F1 Score
Log Train	0.95	0.95	0.58	0.89	0.70
Log Test	0.95	0.97	0.69	0.85	0.76
Log Train with SMOTE	0.92	0.97	0.93	0.91	0.92
Log Test with SMOTE	0.89	0.97	0.96	0.52	0.68
Random Forest Train	0.97	0.99	0.77	0.92	0.84
Random Forest Test	0.97	0.99	0.83	0.94	0.88
Random Forest Train with SMOTE	0.97	1.00	0.96	0.98	0.97
Random Forest Test with SMOTE	0.97	0.99	0.91	0.84	0.87
LDA Train	0.93	0.94	0.42	0.80	0.55
LDA Test	0.91	0.94	0.39	0.73	0.51
LDA Train with SMOTE	0.89	0.96	0.93	0.86	0.89
LDA Test with SMOTE	0.84	0.95	0.94	0.42	0.58

For Random Forest and LDA scaling was not taken into account as these models are not impacted by varying magnitudes of the variables.

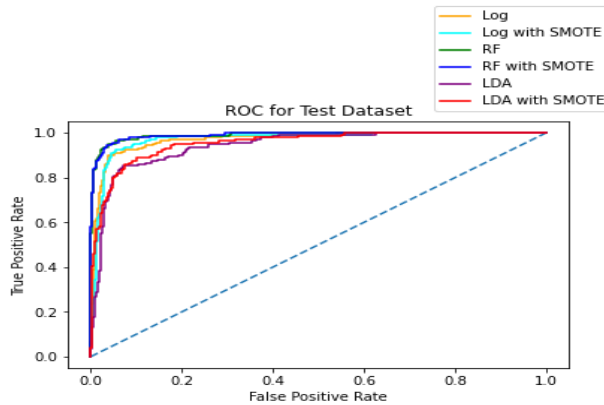
- As per above comparison chart we can observe, Random Forest model has excelled all other models in terms of accuracy.
- Both Random Forest with and without smote feature as the top 2 models.
- While Random Forest was the best performing model, LDA models were the worst performers of all. LDA with smote being the worst of the lot.

Comparison of different evaluation metrics for all the models on the training data set with the help of ROC Curve:



- This clearly depicts, the dark blue line indicating "Random Forest with SMOTE" being on top of all other models, consistently for all the evaluation parameters.
- While the purple line for LDA with smote is slightly higher than LDA with smote in terms of accuracy and auc score, it slips down significantly on all other parameters, making it the worst performing model of the lot.

Comparison of different evaluation metrics for all the models on the test data set using ROC Curve:



- When it comes to test data set, which is the deciding factor, here too, we see Random Forest with SMOTE indicated by dark blue line performing the best on almost all the fronts.
- We can see the dark blue line dip only very slightly on accuracy compared to Random Forest without smote. On all other fronts the blue line stays on the top and is the clear winner amongst all the models.

Random Forest model with smote dataset is the best model amongst all the models with 97% accuracy, 84% precision, 91% recall and F1 score of 0.87 respectively.

1.13 State Recommendations from the above models.

- We are choosing the Random Forest Model as our optimum model as we have higher performance metrics on all values – accuracy, precision and recall values for both the training and testing sets.
- We have choose this model as our best , because logistic regression model gives better results only when we use SMOTE technique for our unbalanced dataset issue. Logistic Regression model gives good recall value when compared to all our models, but our precision value takes a hit for our defaulters (people who are predicted to be defaulters).
- We set out to identify the potential customers in the bank who are predicted to default and our model should give us high recall value, but we also need to consider a model which gives high precision value also.
- Random model also gives high AUC value of 99% in the testing set, which will give better results in identifying our potential defaulters and it is important for us as a bank to identify our defaulters.

PROBLEM 2: MARKET RISK

The dataset contains 6 years of information (weekly stock information) on the stock prices of 10 different Indian Stocks. Calculate the mean and standard deviation on the stock returns and share insights.

DEFINING BUSINESS PROBLEM

The Main aim of conducting this case study is studying the stock prices in depth for advising the portfolio creation.

Approach:

We have imported the dataset and we could see the stock values of various companies across the years 2014 to 2021.

We have done the cleanup of the dataset of 314 observations and 11 variables by updating the variable names in python readable form.

	Date	Infosys	Indian Hotel	Mahindra & Mahindra	Axis Bank	SAIL	Shree Cement	Sun Pharma	Jindal Steel	Idea Vodafone	Jet Airways
0	31-03-2014	264	69	455	263	68	5543	555	298	83	278
1	07-04-2014	257	68	458	276	70	5728	610	279	84	303
2	14-04-2014	254	68	454	270	68	5649	607	279	83	280
3	21-04-2014	253	68	488	283	68	5692	604	274	83	282
4	28-04-2014	256	65	482	282	63	5582	611	238	79	243

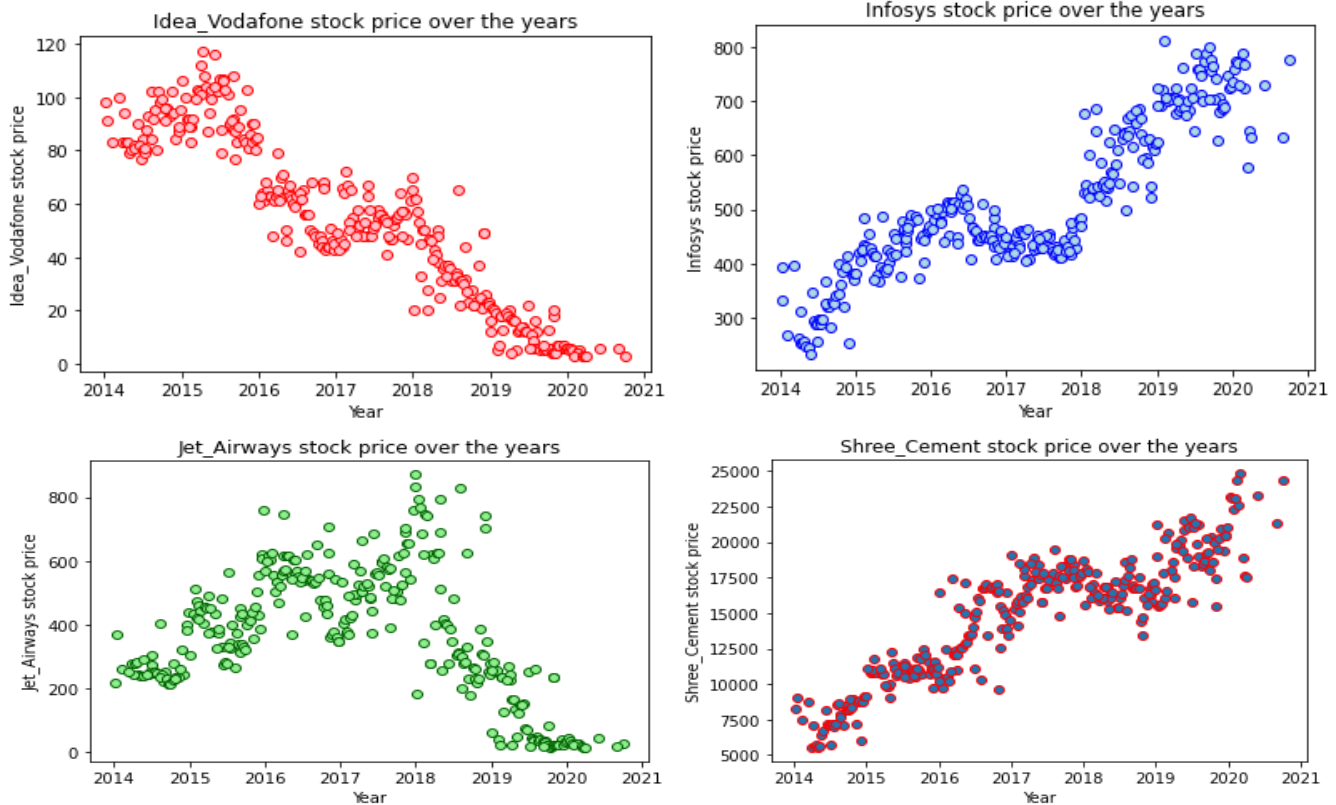
Descriptive Statistics:

	Infosys	Indian_Hotel	Mahindra_&_Mahindra	Axis_Bank	SAIL	Shree_Cement	Sun_Pharma	Jindal_Steel	Idea_Vodafone	Jet_Airways
count	314.000000	314.000000	314.000000	314.000000	314.000000	314.000000	314.000000	314.000000	314.000000	314.000000
mean	511.340764	114.560510	636.678344	540.742038	59.095541	14806.410828	633.468153	147.627389	53.713376	372.659236
std	135.952051	22.509732	102.879975	115.835569	15.810493	4288.275085	171.855893	65.879195	31.248985	202.262668
min	234.000000	64.000000	284.000000	263.000000	21.000000	5543.000000	338.000000	53.000000	3.000000	14.000000
25%	424.000000	96.000000	572.000000	470.500000	47.000000	10952.250000	478.500000	88.250000	25.250000	243.250000
50%	466.500000	115.000000	625.000000	528.000000	57.000000	16018.500000	614.000000	142.500000	53.000000	376.000000
75%	630.750000	134.000000	678.000000	605.250000	71.750000	17773.250000	785.000000	182.750000	82.000000	534.000000
max	810.000000	157.000000	956.000000	808.000000	104.000000	24806.000000	1089.000000	338.000000	117.000000	871.000000

1.1 Draw Stock Price Graph (Stock Price vs Time) for any 2 given stocks with inference.

We have Stock prices of Infosys, Indian_Hotel, Mahindra and Mahindra, Axis Bank, SAIL, Shree Cement, Sun_Pharma, Jindal Steel, Idea_Vodafone and Jet Airways for the period Mar 2014 to Mar 2020.

We drew up the Stock price graph for Idea Vodafone, Infosys, Jet Airways and Shree_Cement.



For **Idea_Vodafone**, stock price is decreasing over the years and we could see that the prices are inversely proportional to the year. We find the company has higher risk in terms of stock market investments, as it could improve the share value even after merger acquisition. It is a company with descent stock price over the year.

For **Infosys**, we could see that the stock price seems to be on up rise from the years 2014 to 2021. We could see a slight dip in stock prices in the year 2018 but the Company got hold of its problems and from then on, there is steep increase in the stock price. It is a Company with good stock prices over the year.

For **Jet_Airways**, shares has been fluctuating and has seen a sharp decrease starting 2018.

For **Shree_Cement**, had an increasing trend and has been stable in 2017 and has an increasing trend again.

1.2 Calculate Returns for all stocks with inference

We calculate the returns by taking logarithmic value difference by subtracting the stock value from any day to its previous day value.

The first row will return a null value, as we don't have any previous day stock price value.

Below are the returns calculated for all stocks:

:

	Infosys	Indian_Hotel	Mahindra_&_Mahindra	Axis_Bank	SAIL	Shree_Cement	Sun_Pharma	Jindal_Steel	Idea_Vodafone	Jet_Airways
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	-0.026873	-0.014599	0.006572	0.048247	0.028988	0.032831	0.094491	-0.065882	0.011976	0.086112
2	-0.011742	0.000000	-0.008772	-0.021979	-0.028988	-0.013888	-0.004930	0.000000	-0.011976	-0.078943
3	-0.003945	0.000000	0.072218	0.047025	0.000000	0.007583	-0.004955	-0.018084	0.000000	0.007117
4	0.011788	-0.045120	-0.012371	-0.003540	-0.076373	-0.019515	0.011523	-0.140857	-0.049393	-0.148846

To analyze the returns further it's better to visualize returns of each stock.

	Infosys	Indian_Hotel	Mahindra_&_Mahindra	Axis_Bank	SAIL	Shree_Cement	Sun_Pharma	Jindal_Steel	Idea_Vodafone	Jet_Airways
count	313.000000	313.000000	313.000000	313.000000	313.000000	313.000000	313.000000	313.000000	313.000000	313.000000
mean	0.002794	0.000266	-0.001506	0.001167	-0.003463	0.003681	-0.001455	-0.004123	-0.010608	-0.009548
std	0.035070	0.047131	0.040169	0.045828	0.062188	0.039917	0.045033	0.075108	0.104315	0.097972
min	-0.167300	-0.236389	-0.285343	-0.284757	-0.251314	-0.129215	-0.179855	-0.283768	-0.693147	-0.458575
25%	-0.014514	-0.023530	-0.020884	-0.022473	-0.040822	-0.019546	-0.020699	-0.049700	-0.045120	-0.052644
50%	0.004376	0.000000	0.001526	0.001614	0.000000	0.003173	0.001530	0.000000	0.000000	-0.005780
75%	0.024553	0.027909	0.019894	0.028522	0.032790	0.029873	0.023257	0.037179	0.024391	0.036368
max	0.135666	0.199333	0.089407	0.127461	0.309005	0.152329	0.166604	0.243978	0.693147	0.300249

- Infosys stock returns shows, Mean is less than median and maximum value is much more than mean. Comparing the mean return and standard deviation which is risk measure we can say that at risk of 0.035, return of 0.0027 is assured. We can think of investing in Infosys stock.
- Indian Hotel stock returns shows, Mean are 0.000266 little more than median of 0.0 and maximum value 0.199 is much more than mean. Comparing the mean return and standard deviation which is risk measure we can say that at risk of 0.047, return of 0.000266 is assured. Thus this shows that risk is quite higher than the returns but some positive returns are assured there for we can think of investing in this stock.
- Mahindra and Mahindra stock returns shows, Mean is -0.001506 which is significantly less than median of 0.001526. Maximum value 0.89 is much more than mean. Comparing the mean return and standard deviation which is risk measure we can say that at risk of 0.040, return of -0.001506 is assured. Thus this shows that risk is very high and there are no associated returns rather, loss of return is linked. Therefore we better not invest in the Mahindra and Mahindra stock.
- Axis Bank stock returns shows left skewed data. Mean is 0.001167 which is less than median of 0.001614. Maximum value 0.127 is much more than mean. Comparing the mean return and standard deviation which is risk measure we can say that at risk of 0.0458, return of 0.001167 is assured. Thus this shows that risk is very high and returns

are very low. But somehow some returns are assured therefore we can make some investments in this stock.

- SAIL stock shows, Mean are -0.003463 which is significantly less than median of 0.0. Maximum value 0.309 which is much more than mean. Comparing the mean return and standard deviation which is risk measure we can say that at risk of 0.062188, negative return of 0.003463 is assured. Thus this shows that risk is very high and losses are assured. Therefore we better not invest in the SAIL.
- Shree Cement stock returns shows, Mean is 0.0036 which is more or less equal to median of 0.0031. Maximum value 0.152 which is much more than mean. Comparing the mean return and standard deviation which is risk measure we can say that at risk of 0.0399, return of 0.00368 is assured. Thus this shows that risk is significantly high than return. But some returns are assured therefore we can make some investments in this stock.
- Sun Pharma stock returns shows, Mean is -0.001455 which is less than median of 0.001530. Maximum value of 0.166 which is much more than mean. Comparing the mean return and standard deviation which is risk measure we can say that at risk of 0.045033, return of -0.001455 is assured. Thus this shows that risk is significantly high and losses are assured. Therefore we better not invest in the Shree Cement.
- Jindal steel stock returns shows, Mean is -0.004123 which is less than median of 0.0. Maximum value of 0.24 which is much more than mean. Comparing the mean return and standard deviation which is risk measure we can say that at risk of 0.075, return of -0.0041 is assured. Thus this shows that risk is significantly high and losses are assured. Therefore we better not invest in the Jindal Steel.
- Idea Vodafone stock returns shows, Mean are -0.0106 which is less than median of 0.0. Maximum value of 0.69 which is much more than mean. Comparing the mean return and standard deviation which is risk measure we can say that at risk of 0.104, return of -0.010608 is assured. Thus this shows that risk is significantly high and losses are assured. Therefore we better not invest in the Idea Vodafone stock.
- Jet Airways stock returns shows; Mean is -0.009548 which is less than median of -0.005780. Maximum value of 0.302 which is much more than mean. Comparing the mean return and standard deviation which is risk measure we can say that at risk of 0.0979, return of -0.00954 is assured. Thus this shows that risk is significantly high and losses are assured. Therefore we better not invest in the Jet Airways.

2.3 Calculate Stock Means and Standard Deviation for all stocks with inference

Below are the stock mean and standard deviation stock market values across the rows.

We are calculating the mean value and standard deviation using the python functionality.

Lesser the standard deviation, lesser is the risk in investing in that particular company and higher the mean value means that the stock price is on the higher side.

- Companies with higher mean – Shree Cement, Infosys
- Company with lowest mean – Idea_Vodafone

```
Shree_Cement      0.003681
Infosys           0.002794
Axis_Bank         0.001167
Indian_Hotel      0.000266
Sun_Pharma        -0.001455
Mahindra_&_Mahindra -0.001506
SAIL              -0.003463
Jindal_Steel      -0.004123
Jet_Airways       -0.009548
Idea_Vodafone     -0.010608
dtype: float64
```

- Companies with lesser standard deviation - Shree Cement, Infosys
- Idea_Vodafone has the highest risk factor.

```
Idea_Vodafone     0.104315
Jet_Airways       0.097972
Jindal_Steel      0.075108
SAIL              0.062188
Indian_Hotel      0.047131
Axis_Bank         0.045828
Sun_Pharma        0.045033
Mahindra_&_Mahindra 0.040169
Shree_Cement      0.039917
Infosys           0.035070
dtype: float64
```

2.4 Draw a plot of Stock Means vs Standard Deviation and state your inference

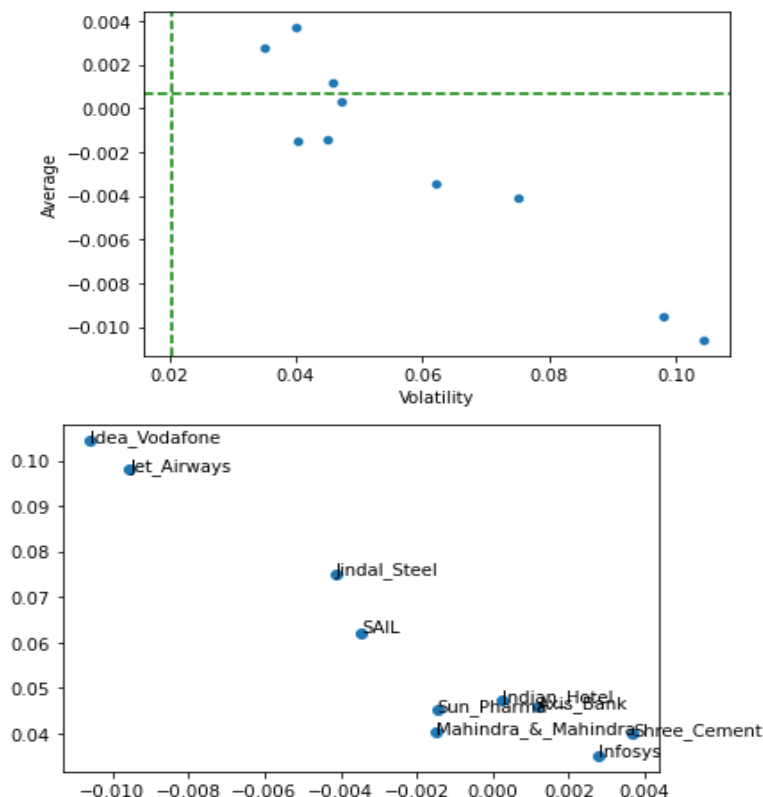
We created a data frame to show its mean value as average and the standard deviation value as its volatility across all companies.

	Average	Volatility
Infosys	0.002794	0.035070
Indian_Hotel	0.000266	0.047131
Mahindra_&_Mahindra	-0.001506	0.040169
Axis_Bank	0.001167	0.045828
SAIL	-0.003463	0.062188
Shree_Cement	0.003681	0.039917
Sun_Pharma	-0.001455	0.045033
Jindal_Steel	-0.004123	0.075108
Idea_Vodafone	-0.010608	0.104315
Jet_Airways	-0.009548	0.097972

We also created the plot for these columns – Average stock price Vs Standard Deviation (Volatility).

For X-axis , we need values with mean values higher than 0 , so we set the line at mean value = 0 and we need to check values higher than that , so we know which companies have higher stock mean value.

For Y-axis, we have set values with lower risk of 0.02 as the y-axis limit for finding out which companies has risk value near to the limit value.



Stocks higher up but on the far left indicate high volatility and low returns (Idea_Vodafone and Jet_Airways), while the stocks on the bottom right indicate low volatility and high returns (Infosys and Shree Cement).

This is a useful graph to find a balance between risk and reward when it comes to investing in different companies.

2.5 Conclusion and Recommendations

CONCLUSION:

Stocks with a lower mean & higher standard deviation do not play a role in a portfolio that has competing stock with more returns & less risk.

Thus for the data we have here, we are only left few stocks:

- One with highest return and lowest risk &
- One with lowest risk and highest return

Therefore:

- 1) Shree Cement, Infosys, Axis Bank and Indian hotel are Good stocks with lower risk and highest returns and are Good investment options
- 2) Sun Pharma, Mahindra and Mahindra, and SAIL have lower risk but yield lower returns
- 3) Idea_Vodafone, Jet Airways and Jindal steel have higher risk and lower returns and are poor investment options

RECOMMENDATION:

We would recommend using the stock means vs. standard deviation plot to assess the risk to reward ratio. More volatile stock might give short term gains but might not be a good investment in long term. Whereas a low volatile stock might not be a good investment in short term, but might give a good return in long term. Hence based on the type of investment that one is looking for, an inference should be made from the above mentioned plot.

THE END