FRA MODULE 2

PROJECT REPORT

BY

M.P. KARTHIKEYAN

Credit Risk

Problem Statement

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.

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

Sample Trainset

	Networth_Next_Year	Equity_Paid_Up	Networth	Capital_Employed	Total_Debt	Gross_Block	Net_Working_Capital	Curr_Assets	Curr_Liab_and_Prov	Total_As
0	-0.53	-0.10	-0.54	-0.67	-0.70	-0.70	-0.45	-0.66	-0.70	
1	-0.12	-0.68	-0.14	-0.45	-0.70	-0.66	-0.50	-0.60	-0.54	
2	1.88	2.08	0.56	1.94	1.96	2.01	1.93	1.97	1.96	
3	-0.65	-0.48	-0.66	-0.73	-0.68	-0.70	-0.62	-0.73	-0.69	
4	-0.31	1.71	-0.29	-0.53	-0.68	-0.60	-0.34	-0.61	-0.71	

The trainset has 2402 rows and 64 columns.

Random Forest Classifier

The best parameters used for building random forest classifier

```
{'max_depth': 7,
 'min_samples_leaf': 5,
 'min_samples_split': 15,
 'n_estimators': 50}
```

Classification report on Train data

	precision	recall	f1-score	support
0.0	1.00	1.00	1.00	2142
1.0	0.99	0.98	0.98	260
accuracy			1.00	2402
macro avg	0.99	0.99	0.99	2402
weighted avg	1.00	1.00	1.00	2402

The model offers higher accuracy on train set.

The Recall value is also much higher.

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

Test Data set

	Networth_Next_Year	Equity_Paid_Up	Networth	Capital_Employed	Total_Debt	Gross_Block	Net_Working_Capital	Curr_Assets	Curr_Liab_and_Prov	Total_As
0	1.88	0.86	1.89	1.94	1.96	2.01	1.93	1.97	1.96	
1	1.18	-0.47	1.12	1.03	1.72	1.81	1.49	0.71	0.15	
2	-0.17	-0.29	-0.22	-0.46	-0.56	0.02	-0.76	-0.46	-0.01	
3	1.88	2.08	1.89	1.94	-0.70	-0.68	1.93	1.97	-0.47	
4	-0.58	-0.57	-0.59	-0.70	-0.69	-0.70	-0.52	-0.70	-0.71	

The trainset has 1184 rows and 64 columns.

Classification report on Test data

	precision	recall	f1-score	support
0.0	0.99	0.99	0.99	1056
1.0	0.94	0.92	0.93	128
accuracy			0.99	1184
macro avg	0.97	0.96	0.96	1184
weighted avg	0.99	0.99	0.99	1184

The recall values are slightly lesser than train set.

The model offers higher accuracy.

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

Linear Discriminant Analysis

Confusion Matrix on Train data

	precision	recall	f1-score	support
0.0	0.94	0.99	0.96	2142
1.0	0.86	0.46	0.60	260
accuracy			0.93	2402
macro avg	0.90	0.73	0.78	2402
weighted avg	0.93	0.93	0.92	2402

Although the accuracy of the model is better the Recall values are much lower.

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

Confusion Matrix on Test data

	precision	recall	f1-score	support
0.0	0.93	0.99	0.96	1056
1.0	0.81	0.43	0.56	128
accuracy			0.93	1184
macro avg	0.87	0.71	0.76	1184
weighted avg	0.92	0.93	0.92	1184

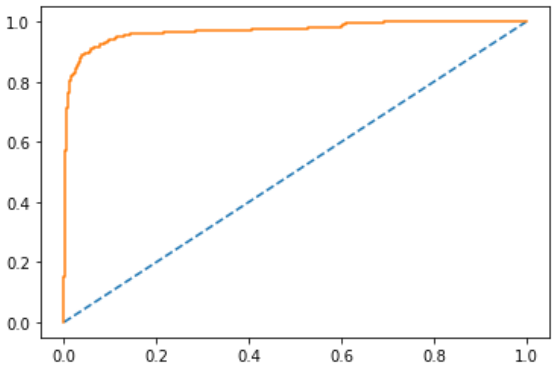
The model behaves same on test set as in train set.

It offers lower Recall value**.**

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

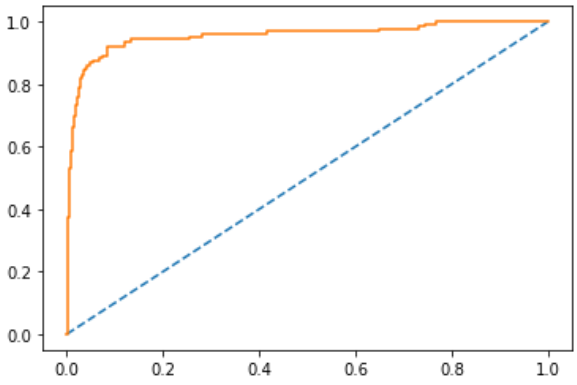
ROC Cuve for Logistic Regression

Train set



Fpr

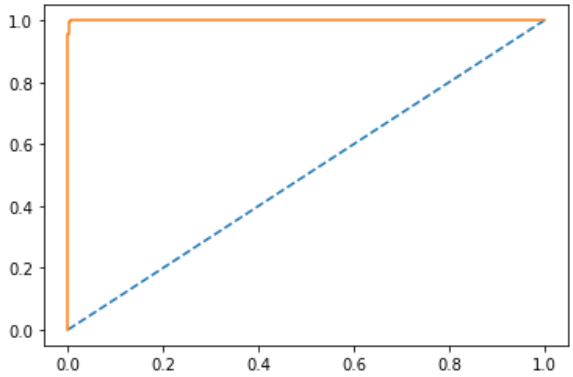
Test set



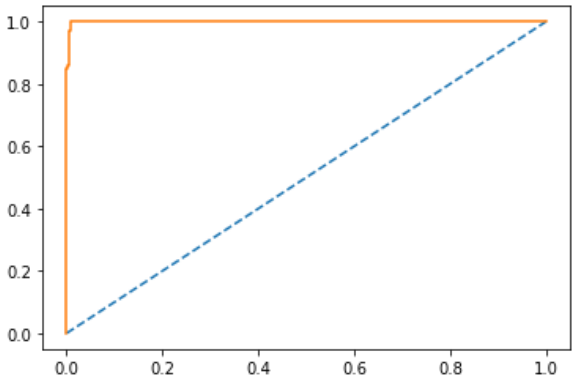
Fpr

ROC Cuve for Random Forest

Train set

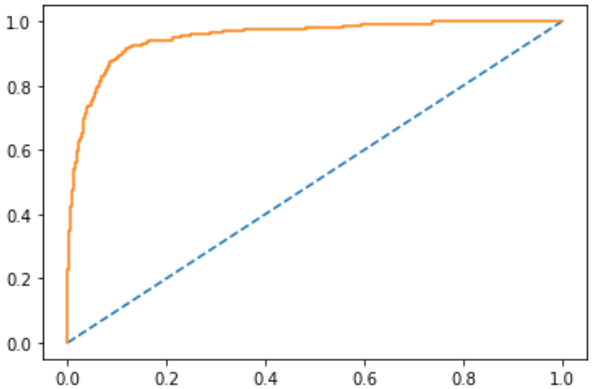


Test set

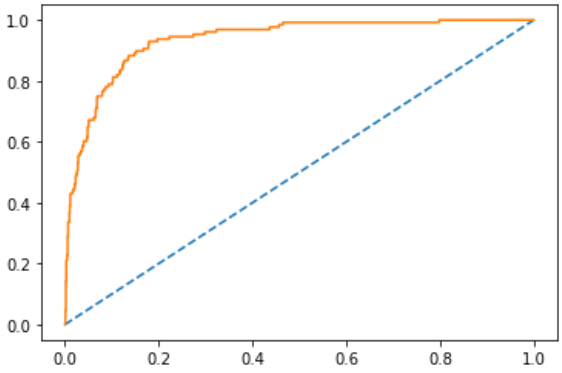


ROC Cuve for LDA

Train set



Test set



AUC Score on Train data

1.Logistic Regression -0.97

2. Random Forest -1.0

3.LDA – 0.952

AUC Score on Test data

1.Logistic Regression -0.97

2. Random Forest -0.99

3.LDA -0.935

Accuracy on Train data

1.Logistic Regression -0.96

2. Random Forest -1.0

3.LDA – 0.93

AUC Score on Test data

1.Logistic Regression -0.95

2. Random Forest -0.99

3.LDA -0.93

Based on the accuracy score & Recall value 'Radom Forest' is considered as the best model for predicting defaults.

1.13 State Recommendations from the above models

The companies with higher debts have significantly higher risk of detault,Providing loans to such companies must be reviewed carefully.

The companies with higher Net working capital & Current Assets have significantly lower risk of defaults making it eligible for Loans.

Market Risk

2.Problem Statement

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.

Sample Dataset

	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

Dataset shape

The dataset ha 314 rows

The dataset has 11 columns

Dataset Info

```
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Date                314 non-null    object
 1   Infosys             314 non-null    int64
 2   Indian Hotel        314 non-null    int64
 3   Mahindra & Mahindra 314 non-null    int64
 4   Axis Bank           314 non-null    int64
 5   SAIL                314 non-null    int64
 6   Shree Cement        314 non-null    int64
 7   Sun Pharma          314 non-null    int64
 8   Jindal Steel        314 non-null    int64
 9   Idea Vodafone       314 non-null    int64
 10  Jet Airways         314 non-null    int64
```

The dataset has 10 integer & 1 object type variables.

The ‘Date’ variable is then converted from ‘object’ type to ‘datetime’ variable

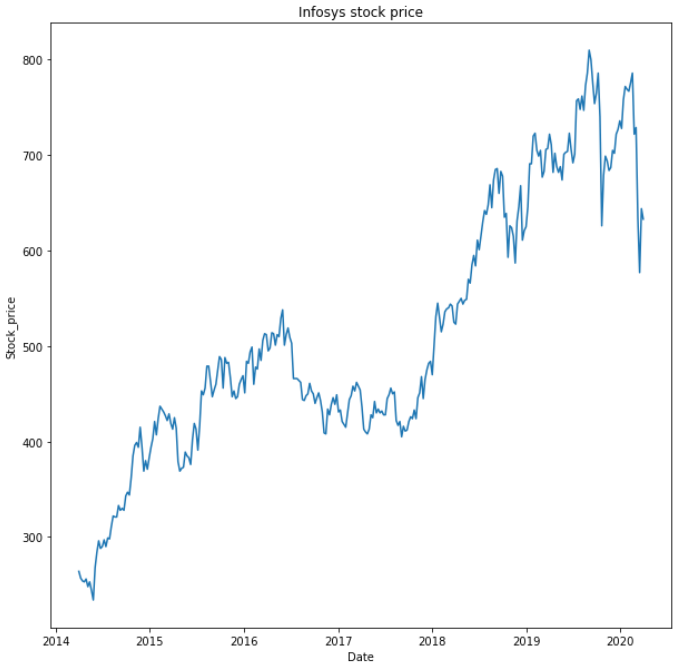
```
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Date                 314 non-null    datetime64[ns]
 1   Infosys              314 non-null    int64
 2   Indian Hotel         314 non-null    int64
 3   Mahindra & Mahindra  314 non-null    int64
 4   Axis Bank            314 non-null    int64
 5   SAIL                 314 non-null    int64
 6   Shree Cement         314 non-null    int64
 7   Sun Pharma           314 non-null    int64
 8   Jindal Steel         314 non-null    int64
 9   Idea Vodafone        314 non-null    int64
 10  Jet Airways          314 non-null    int64
dtypes: datetime64[ns](1)  int64(10)
```

Summary of Dataset

	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

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

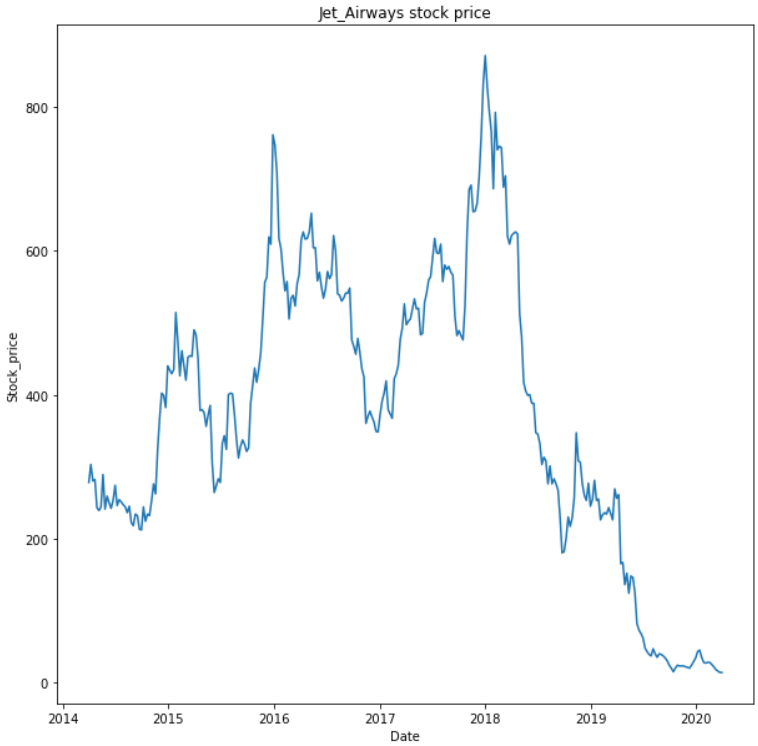
Stock 1 :Infosys



The stock price of Infosys have witnessed overall positive trend on an average over the years.

The stock reached it’s all time high during 2019.

Stock 2 :Jet_Airways



The stock price of Jet Airways is facing significant decline since 2018 .

The stock reached its All time low price at 2020.

2.2 Calculate Returns for all stocks with inference

The stock returns are calculated on weekly basis.

	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

Summary of stock returns

	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

The stock returns of all companies are fluctuating .

The highest weekly return made was 0.693 from Idea_vodafone .

The Lowest weekly return made was 0.127 from Axis Bank .

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

Stock Means

The Average returns made over the years from each stock are as follows

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

The highest average return was made from Shree Cement stock.

The average returns are positive for Infosys,Indian Hotel,Axis Bank & Shree cement.

The average returns are negative for Mahindra & Mahindra,SAIL,Sun Pharma,Jindal steel,Idea Vodafone & Jet Airways.

Standard Deviation for all stocks

The Standard Deviation of the stocks over the years are as follows

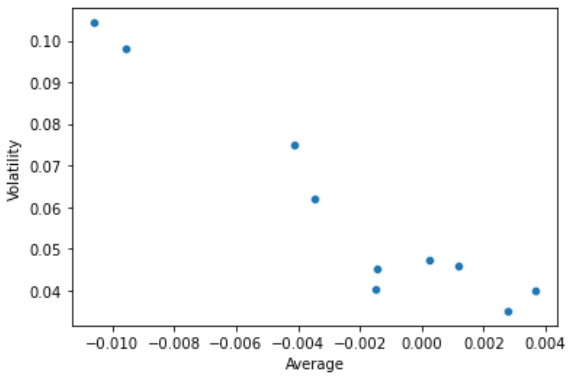
```
Infosys                 0.035070
Indian_Hotel            0.047131
Mahindra_&_Mahindra     0.040169
Axis_Bank               0.045828
SAIL                    0.062188
Shree_Cement            0.039917
Sun_Pharma              0.045033
Jindal_Steel            0.075108
Idea_Vodafone           0.104315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

The most volatile stock was Idea Vodafone followed by Jet Airways.

Infosys has significantly lesser volatility.

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

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



Idea Vodafone offered the lowest average return with highest Volatility over the years.

Jet Airways offered second lowest return with higher volatility.

Shree Cement stock made higher returns with less volatility.

Infosys made second highest return with least volatility.

2.5 Conclusion and Recommendations

Based on the weekly stock price Analysis made on the companies we can conclude the following things

Shree Cement & **Infosys** are the most preferable stocks, these stocks belonging to infrastructure & Tech sectors also has significant potential for growth in coming years making it an **attractive investment.**

Idea Vodafone is facing immense competition from it's competitors in telecom sector.The company has lost significant marketshare and the losses are piling up.The returns are seeing steady decline.

Jet Airways is facing losses since 2018 and this has been aggravated by the pandemic.The significant drop in passenger flights during covid have deepen the loss.

These two stocks discussed above are to be **avoided from investing** .

It is always better to maintain an investment portfolio with stocks from different sectors.