CREDIBILITY OF CLIENTS FOR LOAN APPLICATION

Problem Statement:

In this application, we are basically classifying if the company/bank should give the clients further credits if they request for loan application with the help of past 6 months data regarding their limit balance , past 6 months payment details ,age , gender , education.

The workflow of this application:

Step 1: The Customer applies for request of loan application to the bank .

Step 2: The bank employee now uses this application to check whether this Customer is eligible for loan approval or not.

Step 3: The UI page consists of a form in which the Employee will have to enter the Customer_ID and choose which classification algorithm he wants to apply in order to determine the result.

Step 4: After the employee presses the submit button, the values submitted in the form is passed into python code. Using the sys package, the python code provided the value given by the employee via filling the form.

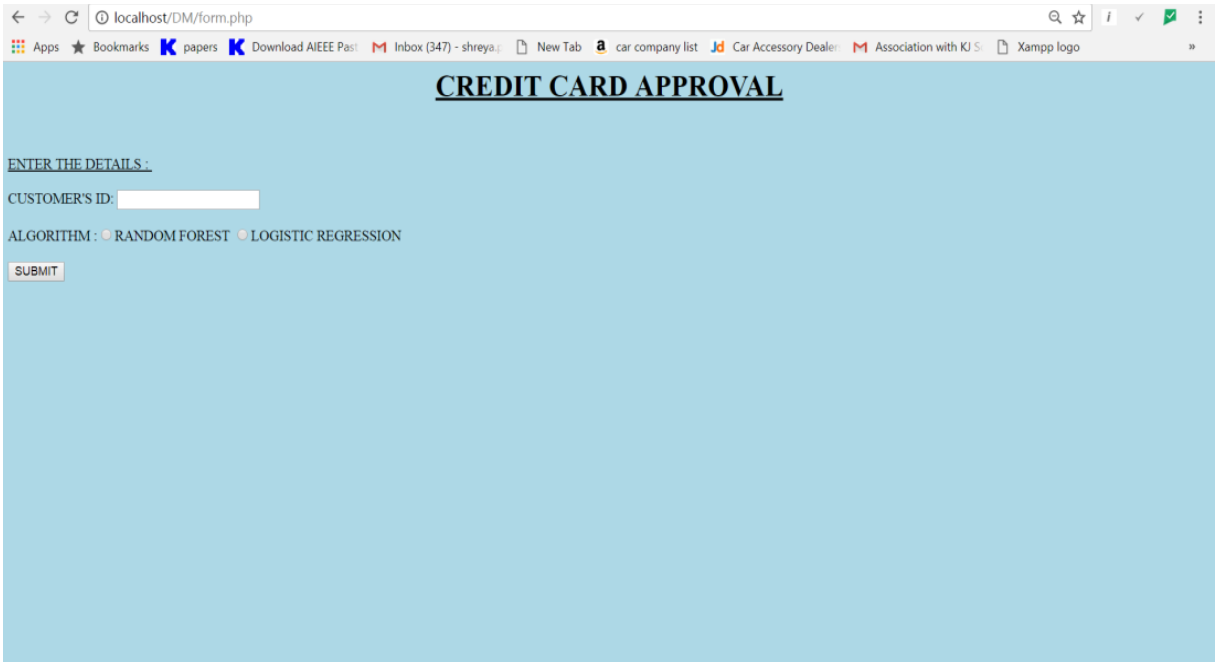
Step 5: The python code reads the excel sheet and it extracts the important features using rfe model of logistic regression which helps us to determine the features that contribute the most for the required classification.

Step 6: Then it trains the model based on the data from the input excel sheet using either of 1) Random Forest Classifier or 2)Logistic Regression classifications models and predicts the classification result according to it .

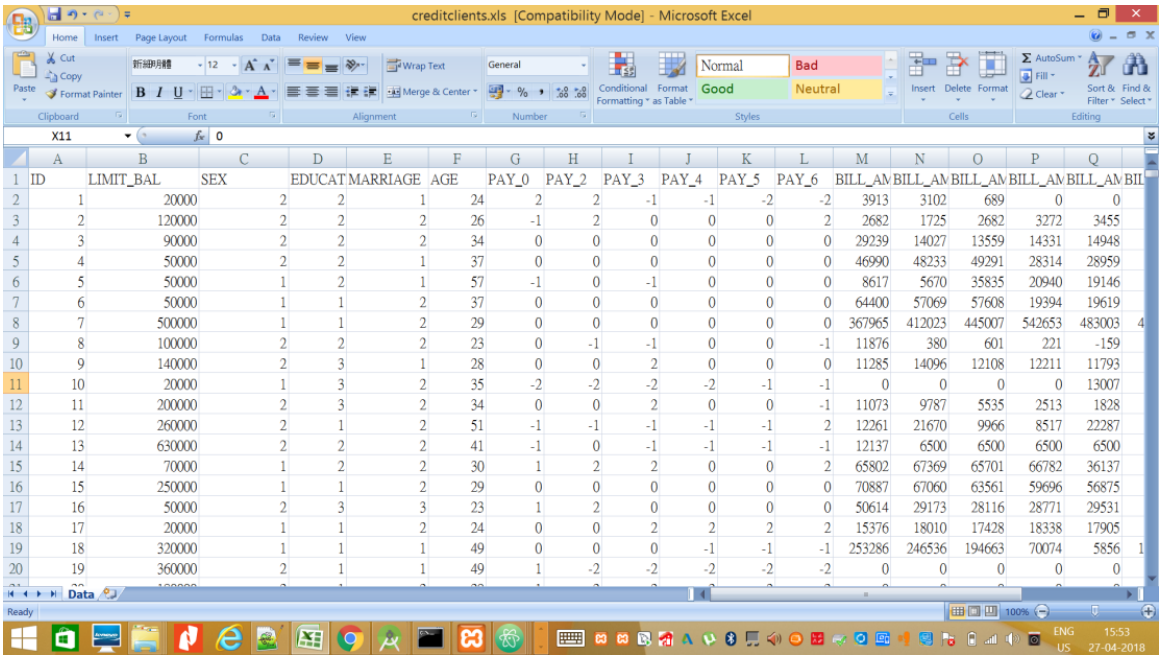
Step 7: The result predicted by the classification models in passed to the UI page where the result gets displayed whether the loan application is granted to the customer or not.

Screenshots Of The Workflow:

UI PAGE



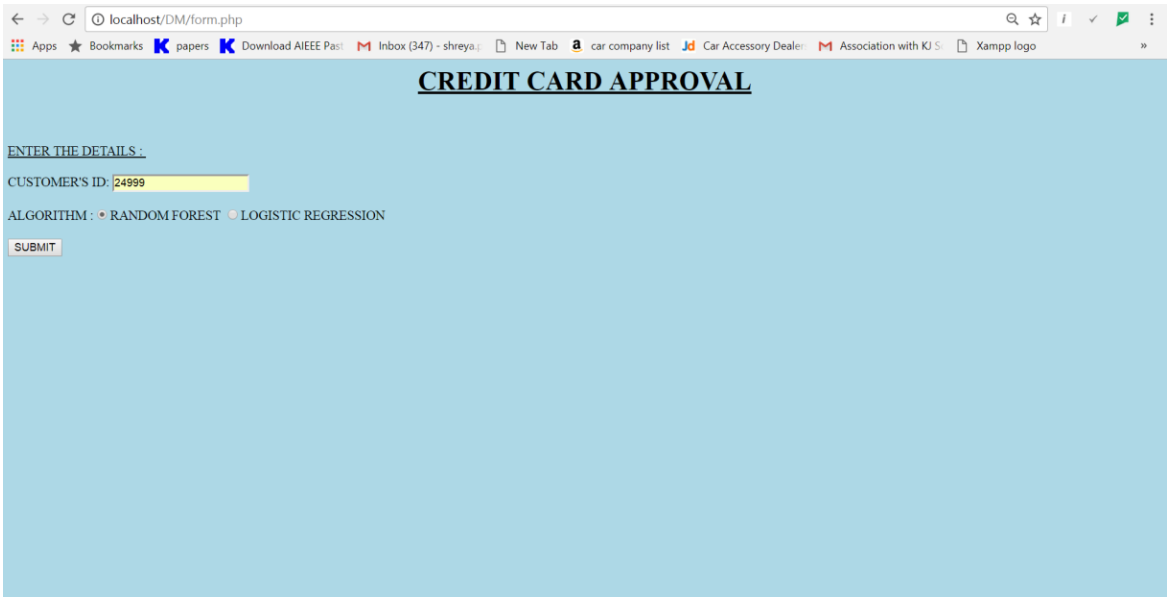
DATASET



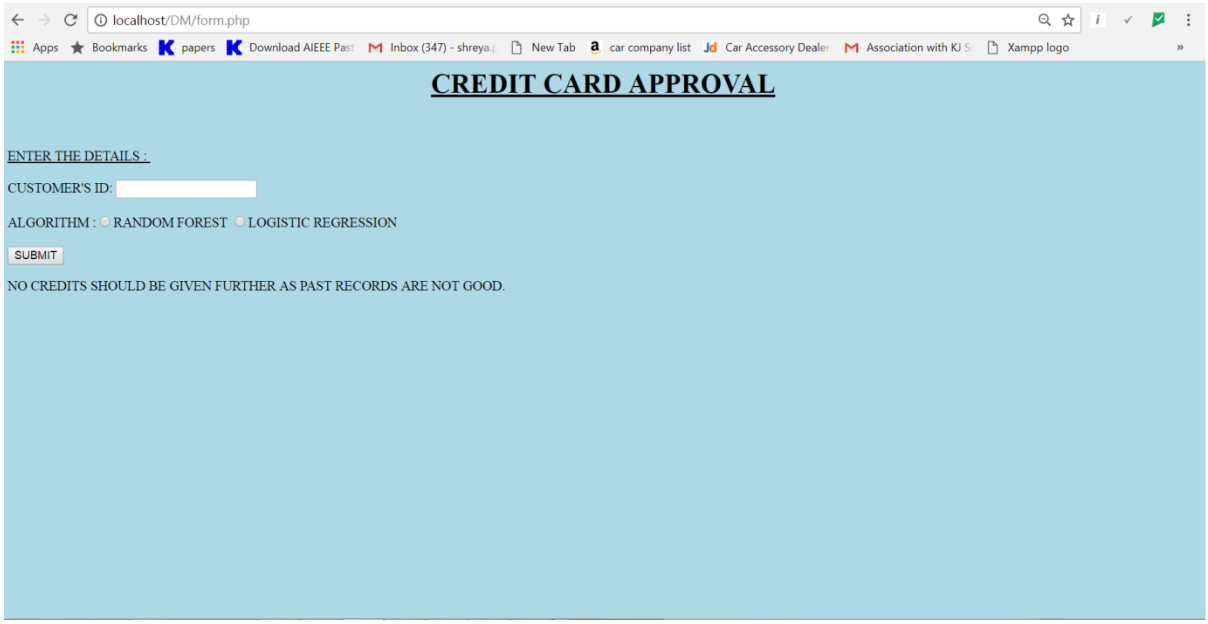
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	ID	LIMIT_BAL	SEX	EDUCAT	MARRIAGE	AGE	PAY_0	PAY_2	PAY_3	PAY_4	PAY_5	PAY_6	BILL_AM	BILL_AM	BILL_AM	BILL_AM	BILL_AM
2	1	20000	2	2	1	24	2	2	-1	-1	-2	-2	3913	3102	689	0	0
3	2	120000	2	2	2	26	-1	2	0	0	0	2	2682	1725	2682	3272	3455
4	3	90000	2	2	2	34	0	0	0	0	0	0	29239	14027	13559	14331	14948
5	4	50000	2	2	1	37	0	0	0	0	0	0	46990	48233	49291	28314	28959
6	5	50000	1	2	1	57	-1	0	-1	0	0	0	8617	5670	35835	20940	19146
7	6	50000	1	1	2	37	0	0	0	0	0	0	64400	57069	57608	19394	19619
8	7	500000	1	1	2	29	0	0	0	0	0	0	367965	412023	445007	542653	483003
9	8	100000	2	2	2	23	0	-1	-1	0	0	-1	11876	380	601	221	-159
10	9	140000	2	3	1	28	0	0	2	0	0	0	11285	14096	12108	12211	11793
11	10	20000	1	3	2	35	-2	-2	-2	-2	-1	-1	0	0	0	0	13007
12	11	200000	2	3	2	34	0	0	2	0	0	-1	11073	9787	5535	2513	1828
13	12	260000	2	1	2	51	-1	-1	-1	-1	-1	2	12261	21670	9966	8517	22287
14	13	630000	2	2	2	41	-1	0	-1	-1	-1	-1	12137	6500	6500	6500	6500
15	14	70000	1	2	2	30	1	2	2	0	0	2	65802	67369	65701	66782	36137
16	15	250000	1	1	2	29	0	0	0	0	0	0	70887	67060	63561	59696	56875
17	16	50000	2	3	3	23	1	2	0	0	0	0	50614	29173	28116	28771	29531
18	17	20000	1	1	2	24	0	0	2	2	2	2	15376	18010	17428	18338	17905
19	18	320000	1	1	1	49	0	0	0	-1	-1	-1	253286	246536	194663	70074	5856
20	19	360000	2	1	1	49	1	-2	-2	-2	-2	-2	0	0	0	0	0

PS: DATASET ATTACHED TO THE REPORT.

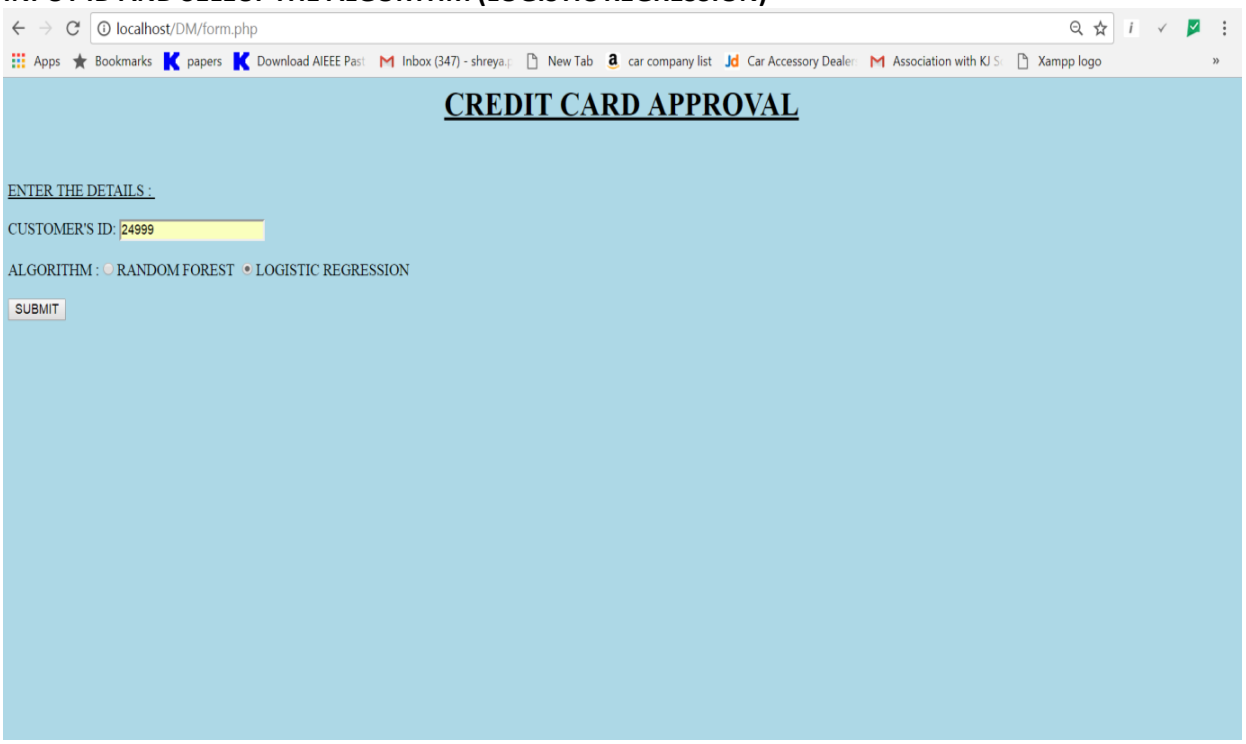
INPUT ID AND SELECT THE ALGORTHM (RANDOM FOREST ALGORITHM)



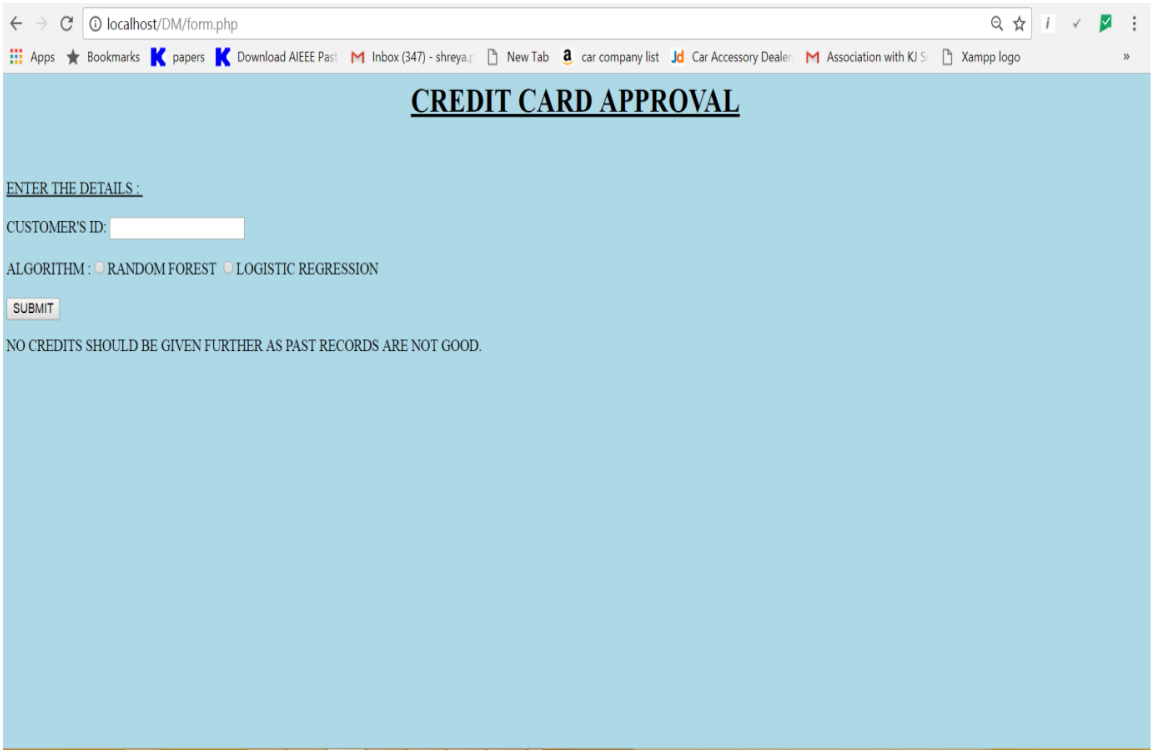
OUTPUT



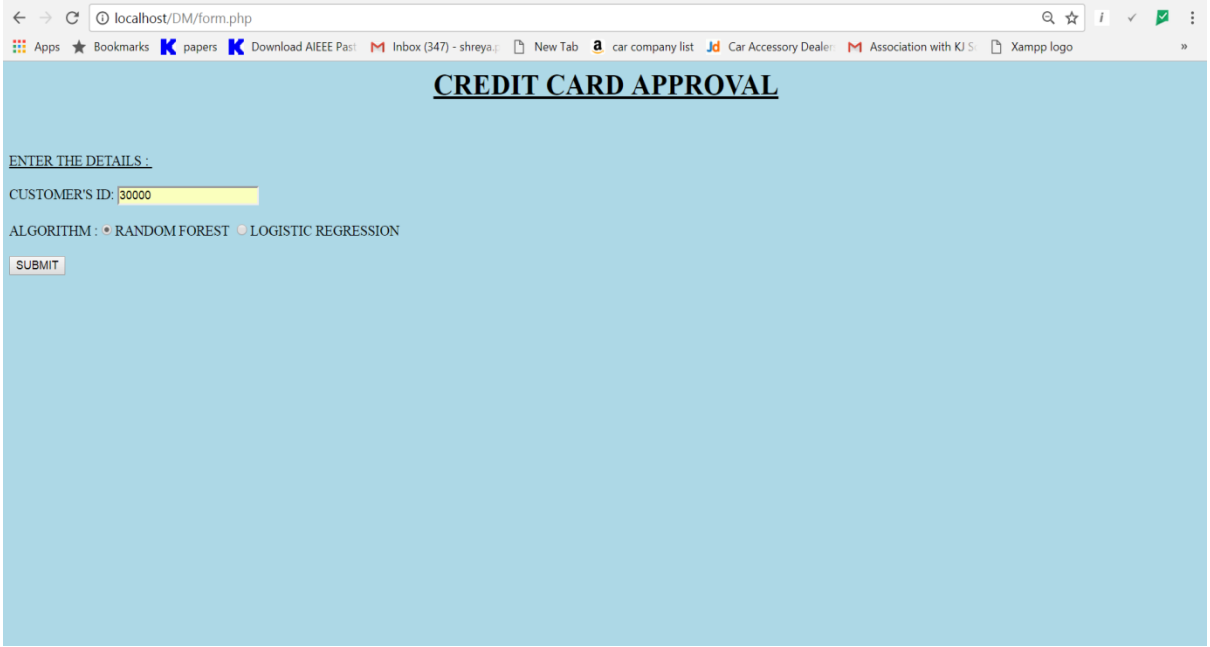
INPUT ID AND SELECT THE ALGORTHM (LOGISTIC REGRESSION)



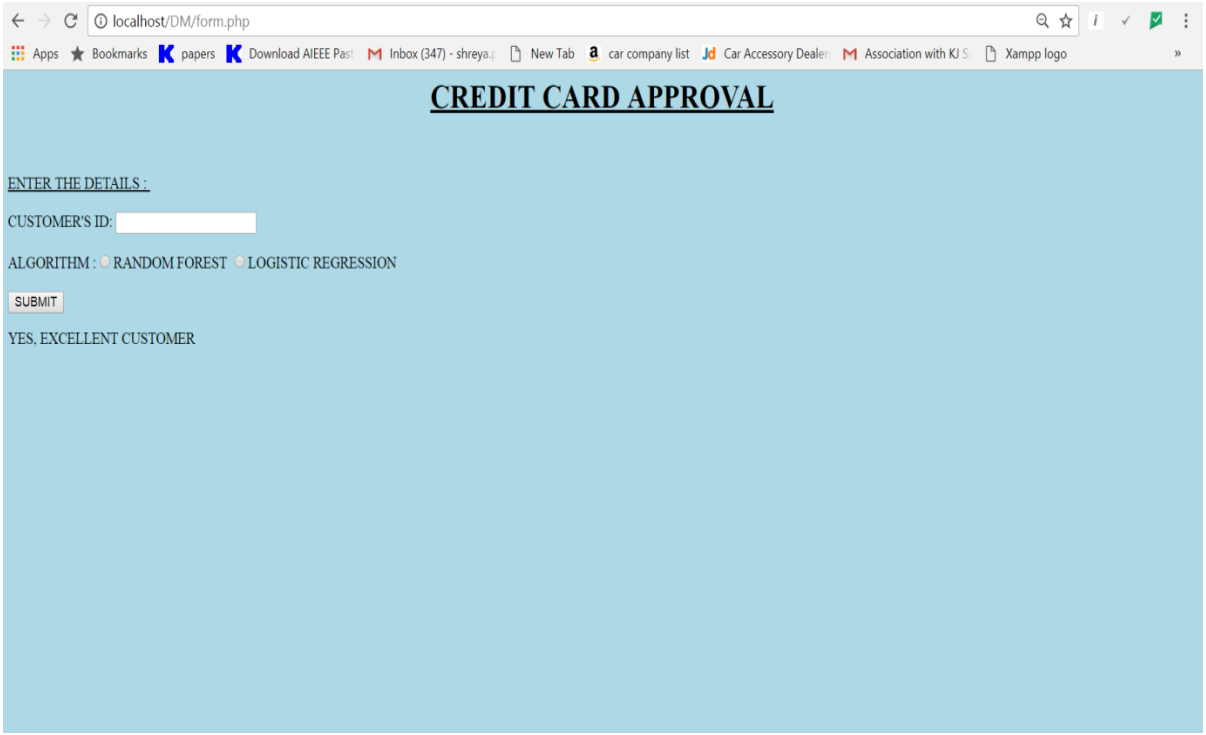
OUTPUT



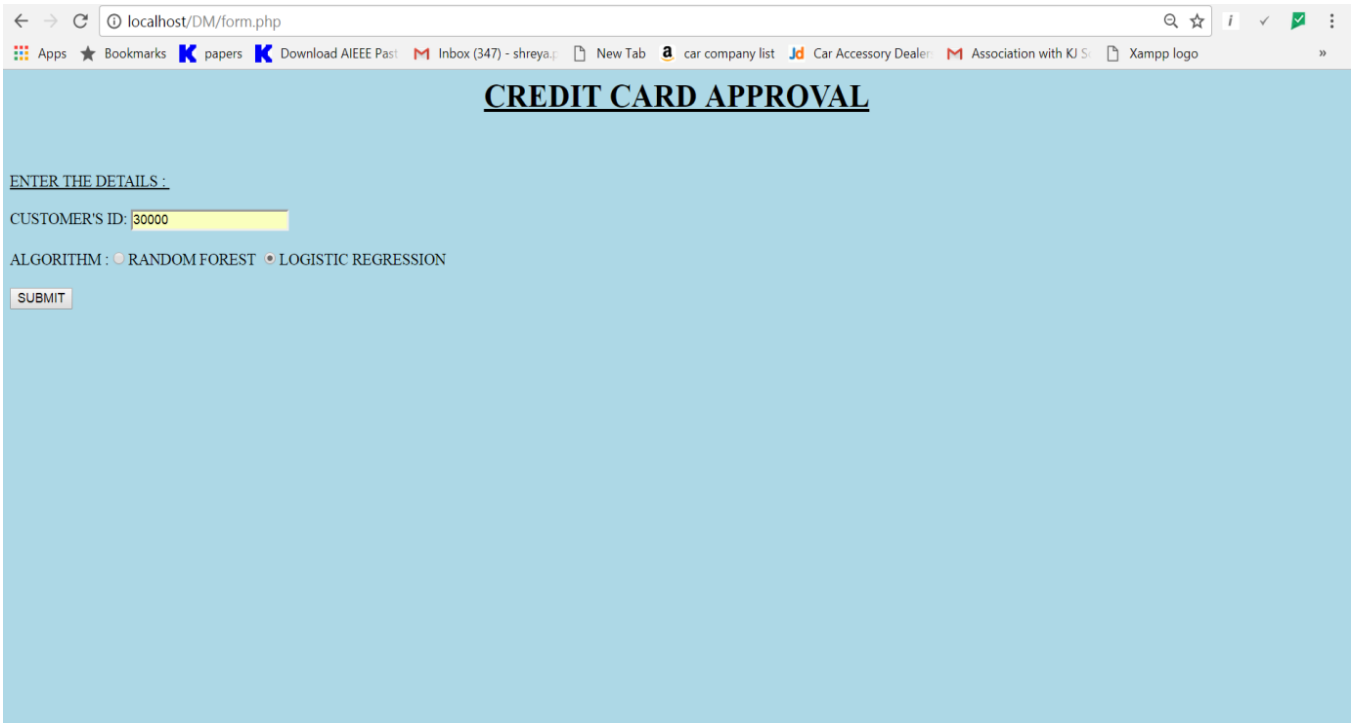
INPUT ID AND SELECT THE ALGORTHM (RANDOM FOREST ALGORITHM)



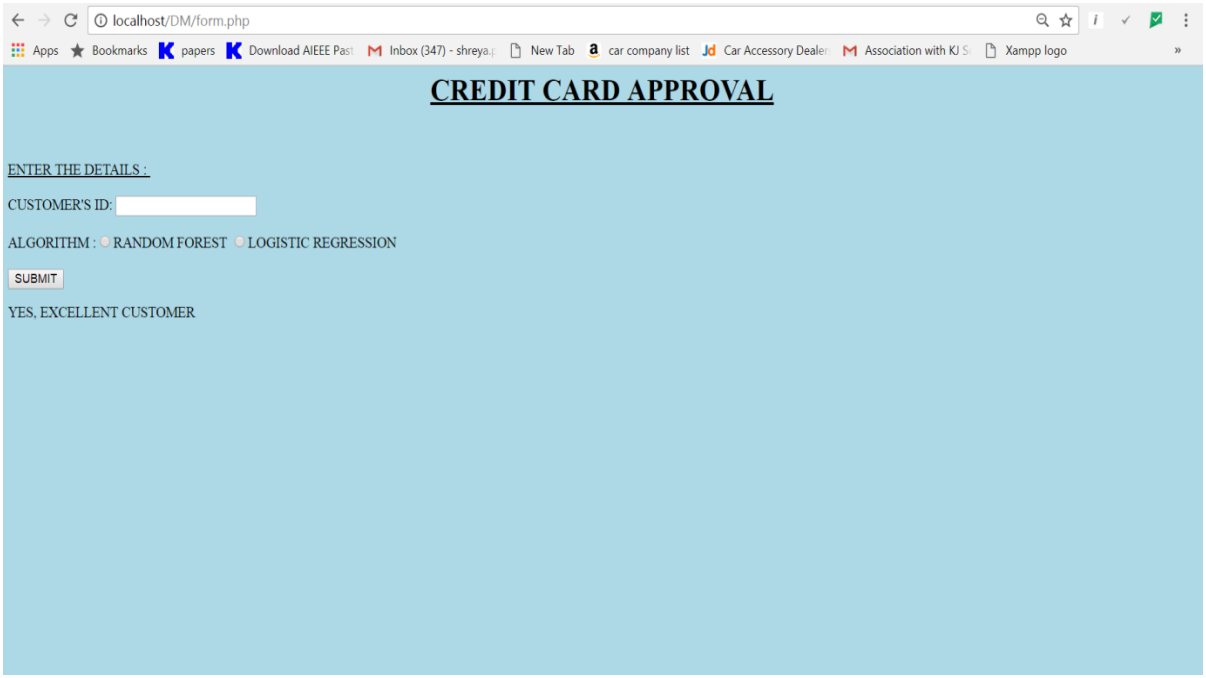
OUTPUT



INPUT ID AND SELECT THE ALGORTHM (LOGISTIC REGRESSION)



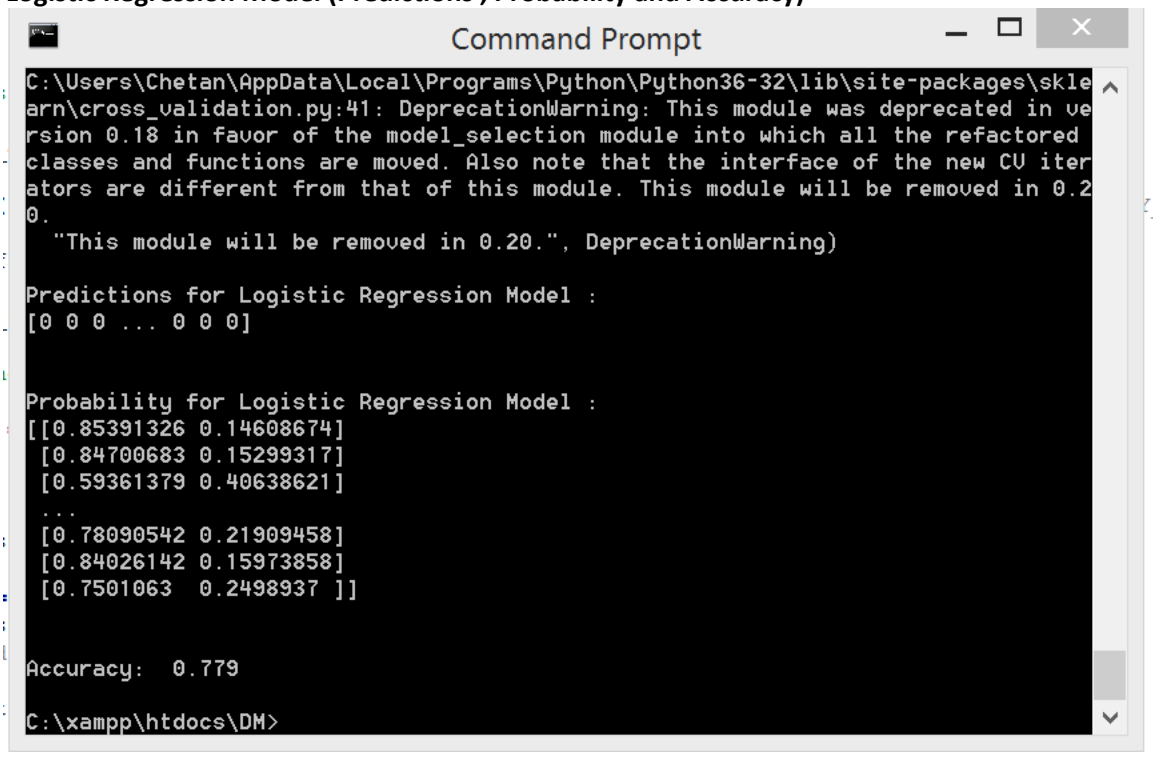
OUTPUT



Question:

1. Provide Verify accuracy of the mini algorithms using suitable measure.

Logistic Regression Model (Predictions , Probability and Accuracy)



Random Forest Classifier (Predictions , Probability and Accuracy)