Mohan.k mohankancherla519@gmail.com

ML-OUTSTANDING-1, OUTSTANDING-2, and OUTSTANDING-3

Q1)How cleaning/EDA was performed

1. Firstly finding the structure of the dataset
2. Finding the total non null values
3. Replacing the null values using median
4. Making the outliners using sns(seaborn) function plot
5. Counter plot of married and non married for a gender

Q2)Your independent and dependent feature

Dependent feature

For the approval or not the final answer

Independent feature

1. 'Gender'
2. Married'
3. 'Dependents'
4. 'Education'
5. 'Self Employed’
6. ’Property Area'

Q3)Why and how selection/engineering/scaling was performed

Algorithms affected by feature rescaling

Algorithms in which two dimensions affect the outcome will be affected by rescaling

SVM with RBF kernel

When you maximize the distance, you've 2 or more dimensions

K-means clustering

Feature Scaling in Scikit-learn

```
from sklearn.preprocessing import MinMaxScaler
import numpy as np
```

Interpretability

Insight

Curse of dimensionality

Q4) Which activation function was chosen and why?

ReLU (Rectified Linear Unit) is chosen cz

Computationally efficient—allows the network to converge very quickly

Non-linear—although it looks like a linear function, ReLU has a derivative function and allows for backpropagation

The Dying ReLU problem—when inputs approach zero, or are negative, the gradient of the function becomes zero, the network cannot perform backpropagation and cannot learn.

Q5) Which optimizer was chosen and why?

In SGD, the optimizer estimates the direction of steepest descent based on a mini-batch and takes a step in this direction. Because the step size is fixed, SGD can quickly get stuck on plateaus or in local minima.cz and its decent normal generalization

Q6) Which neural network and why? Describe how your neural structuring?

I have taken all the 4 networks as follows and their effeciencies

1. Logistic Regression-82
2. Random forest-81
3. SVM-82
4. Decision tree-81(CHOOSEN TO BE THE BEST NETWORK)

For LR

Loan_ID	Gender	Approval	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Gender_Female	Gender_M
LP001015	Male	Approved	5720	0	110.0	360.0	1.0	0	
LP001022	Male	Declined	3076	1500	126.0	360.0	1.0	0	
LP001031	Male	Declined	5000	1800	208.0	360.0	1.0	0	
LP001035	Male	Declined	2340	2546	100.0	360.0	1.0	0	
LP001051	Male	Approved	3276	0	78.0	360.0	1.0	0	
...	...	...	...	...	...	...	...	...	
LP002971	Male	Declined	4009	1777	113.0	360.0	1.0	0	
LP002975	Male	Declined	4158	709	115.0	360.0	1.0	0	
LP002980	Male	Declined	3250	1993	126.0	360.0	1.0	0	
LP002986	Male	Declined	5000	2393	158.0	360.0	1.0	0	
LP002989	Male	Approved	9200	0	98.0	180.0	1.0	0	

367 rows × 20 columns

Mohan.k mohankancherla519@gmail.com

For random forest

RF

```
6]: test_rf_predict = rf_model.predict(test_df)
predict=pd.DataFrame(data=test_rf_predict ,columns=["Approval"])
results=pd.concat([test_df,predict,gender,loanid],axis=1)
results.Approval.replace([1,0],['Approved','Declined'],inplace=True)
results.set_index(['Loan_ID','Gender','Approval'],inplace=False)
```

6]:

			ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Gender_Female	Gender_Male	Marri
Loan_ID	**Gender**	**Approval**								
LP001015	**Male**	**Declined**	5720	0	110.0	360.0	1.0	0	1	
LP001022	**Male**	**Declined**	3076	1500	126.0	360.0	1.0	0	1	
LP001031	**Male**	**Declined**	5000	1800	208.0	360.0	1.0	0	1	
LP001035	**Male**	**Declined**	2340	2546	100.0	360.0	1.0	0	1	
LP001051	**Male**	**Declined**	3276	0	78.0	360.0	1.0	0	1	
...	**...**	**...**	...	...	...	...	...	...	...	
LP002971	**Male**	**Declined**	4009	1777	113.0	360.0	1.0	0	1	
LP002975	**Male**	**Declined**	4158	709	115.0	360.0	1.0	0	1	
LP002980	**Male**	**Approved**	3250	1993	126.0	360.0	1.0	0	1	
LP002986	**Male**	**Declined**	5000	2393	158.0	360.0	1.0	0	1	
LP002989	**Male**	**Approved**	9200	0	98.0	180.0	1.0	0	1	

367 rows × 20 columns

FOR SVM

```
In [58]: test_sv_predict = svm.predict(test_df)
         test_sv_predict
```

```
Out[58]: array([1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
               1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1], dtype=int64)
```

FOR DT(DECISION TREE)

```
]: test_dt_predict = dt_clf.predict(test_df)
   predict=pd.DataFrame(data=test_dt_predict ,columns=["Approval"])
   results=pd.concat([test_df,predict,gender,loanid],axis=1)
   results.Approval.replace([1,0],['Approved','Declined'],inplace=True)
   results.set_index(['Loan_ID','Gender','Approval'],inplace=False)
]:
```

Loan_ID	Gender	Approval	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Gender_Female	Gender
LP001015	Male	Declined	5720	0	110.0	360.0	1.0	0	
LP001022	Male	Declined	3076	1500	126.0	360.0	1.0	0	
LP001031	Male	Declined	5000	1800	208.0	360.0	1.0	0	
LP001035	Male	Declined	2340	2546	100.0	360.0	1.0	0	
LP001051	Male	Declined	3276	0	78.0	360.0	1.0	0	
...	...	...	...	...	...	...	...	...	
LP002971	Male	Declined	4009	1777	113.0	360.0	1.0	0	
LP002975	Male	Declined	4158	709	115.0	360.0	1.0	0	
LP002980	Male	Approved	3250	1993	126.0	360.0	1.0	0	
LP002986	Male	Declined	5000	2393	158.0	360.0	1.0	0	
LP002989	Male	Declined	9200	0	98.0	180.0	1.0	0	

367 rows × 20 columns

```
]:
```

THANK U

MOHAN KANCHERLA