project - Colab

colab.research.google.com/drive/1RbuwdcZ_sQg7DYwLsA6v1qSp5-g0xH38#scrollTo=f_XFpGCGJ5RY

project

File Edit View Insert Runtime Tools Help

Share

Commands + Code + Text Run all

RAM Disk

[1] ✓ 0s

```
#Predicting Home Loan Approval
```

[2] ✓ 1s

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

[3] ✓ 1s

```
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.linear_model import LogisticRegression
from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.naive_bayes import GaussianNB
from xgboost import XGBClassifier
from sklearn.metrics import accuracy_score, precision_score, confusion_matrix
```

[4] ✓ 0s

```
df = pd.read_csv('/content/loan_sanction_train.csv')
df.head()
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Propert
0	LP001002	Male	No	0	Graduate	No	5849	0.0	NaN	360.0	1.0	
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128.0	360.0	1.0	
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66.0	360.0	1.0	

Variables Terminal

5:00 PM Python 3

30°C Sunny

Search

IN 12/13/2025

project - Colab

colab.research.google.com/drive/1RbuwdcZ_sQg7DYwLsA6v1qSp5-g0xH38#scrollTo=f_XFpGCGJ5RY

project

File Edit View Insert Runtime Tools Help

Share

Commands + Code + Text Run all

RAM Disk

3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120.0	360.0	1.0
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141.0	360.0	1.0

Next steps: Generate code with df New interactive sheet

[5] 0s

```
df.sample(3)
```

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Proper
86	LP001280	Male	Yes	2	Not Graduate	No	3333	2000.0	99.0	360.0	NaN	Se
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128.0	360.0	1.0	
31	LP001095	Male	No	0	Graduate	No	3167	0.0	74.0	360.0	1.0	

[6] 0s

```
df.shape
```

```
(614, 13)
```

[7] 0s

```
df.isnull().sum()
```

	0

Variables Terminal

5:00 PM Python 3

30°C Sunny

Search

IN 12/13/2025

Loan_ID	0
Gender	13
Married	3
Dependents	15
Education	0
Self_Employed	32
ApplicantIncome	0
CoapplicantIncome	0
LoanAmount	22
Loan_Amount_Term	14
Credit_History	50
Property_Area	0
Loan_Status	0

dtype: int64

[8]
```
df.duplicated().sum()
```
0s
```
np.int64(0)
```

[9]
```
df.info()
```
0s

project - Colab

colab.research.google.com/drive/1RbuwdcZ_sQg7DYwLsA6v1qSp5-g0xH38#scrollTo=f_XFpGCGJ5RY

project

File Edit View Insert Runtime Tools Help

Share

Commands + Code + Text Run all

RAM Disk

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 614 entries, 0 to 613
Data columns (total 13 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   Loan_ID            614 non-null    object
 1   Gender             601 non-null    object
 2   Married            611 non-null    object
 3   Dependents         599 non-null    object
 4   Education          614 non-null    object
 5   Self_Employed      582 non-null    object
 6   ApplicantIncome    614 non-null    int64
 7   CoapplicantIncome  614 non-null    float64
 8   LoanAmount         592 non-null    float64
 9   Loan_Amount_Term   600 non-null    float64
 10  Credit_History     564 non-null    float64
 11  Property_Area      614 non-null    object
 12  Loan_Status        614 non-null    object
dtypes: float64(4), int64(1), object(8)
memory usage: 62.5+ KB
```

[10] ✓ 0s

```
df.describe()
```

	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History
count	614.000000	614.000000	592.000000	600.00000	564.000000
mean	5403.459283	1621.245798	146.412162	342.00000	0.842199
std	6109.041673	2926.248369	85.587325	65.12041	0.364878
min	150.000000	0.000000	9.000000	12.00000	0.000000

Variables Terminal

5:00 PM Python 3

30°C Sunny

Search

ENG IN

5:10 PM 12/13/2025

project - Colab

colab.research.google.com/drive/1RbuwdcZ_sQg7DYwLsA6v1qSp5-g0xH38#scrollTo=f_XFpGCGJ5RY

project

File Edit View Insert Runtime Tools Help

Share

Commands + Code + Text Run all

RAM Disk

```
 2   Married            611 non-null    object
 3   Dependents         599 non-null    object
 4   Education          614 non-null    object
 5   Self_Employed      582 non-null    object
 6   ApplicantIncome    614 non-null    int64
 7   CoapplicantIncome  614 non-null    float64
 8   LoanAmount         592 non-null    float64
 9   Loan_Amount_Term   600 non-null    float64
 10  Credit_History     564 non-null    float64
 11  Property_Area      614 non-null    object
 12  Loan_Status        614 non-null    object
dtypes: float64(4), int64(1), object(8)
memory usage: 62.5+ KB
```

[10] ✓ 0s

```
df.describe()
```

	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History
count	614.000000	614.000000	592.000000	600.00000	564.000000
mean	5403.459283	1621.245798	146.412162	342.00000	0.842199
std	6109.041673	2926.248369	85.587325	65.12041	0.364878
min	150.000000	0.000000	9.000000	12.00000	0.000000
25%	2877.500000	0.000000	100.000000	360.00000	1.000000
50%	3812.500000	1188.500000	128.000000	360.00000	1.000000
75%	5795.000000	2297.250000	168.000000	360.00000	1.000000
max	81000.000000	41667.000000	700.000000	480.00000	1.000000

Variables Terminal

5:00 PM Python 3

30°C Sunny

Search

ENG IN

5:11 PM 12/13/2025