केन्द्रीय प्रौद्योगिकी संस्थान कोकराझार
CENTRAL INSTITUTE OF TECHNOLOGY KOKRAJHAR
DEEMED TO BE UNIVERSITY UNDER MOE, GOVT. OF INDIA

Mini Project

Loan Default Prediction

Guided by :
Dr. Anup Kumar Barman
Assistant Professor
Dept. of Computer Science and Engineering

Submitted by :
Swmdwn Choudhury(202102021011)
MD Dildar Mandal (202102021043)
Gulnaaz Parveen (202102021055)

Introduction

“Loan default prediction” using machine learning is crucial for financial institutions to assess credit risk and optimize lending decisions. This project aims to develop a predictive model that can forecast whether a borrower is likely to default on a loan based on historical data and key features. By leveraging advanced analytics, we seek to enhance risk management practices and improve loan approval processes.

PROBLEM STATEMENT

- Time-Consuming Manual Process
- Inefficiencies and Delays
- Lack of Real-Time Responses
- Slower and Less Accurate Decisions

SOLUTION

- Implement Automated Loan Default Assessment
- Utilize Predictive Modeling
- Real-Time Decision Making
- Continuous Monitoring and Model Updating

Methodolgy

Loan Data (Input Data)

Data Preprocessing

Machine Learning Model Training

Model Evaluation

Deployment

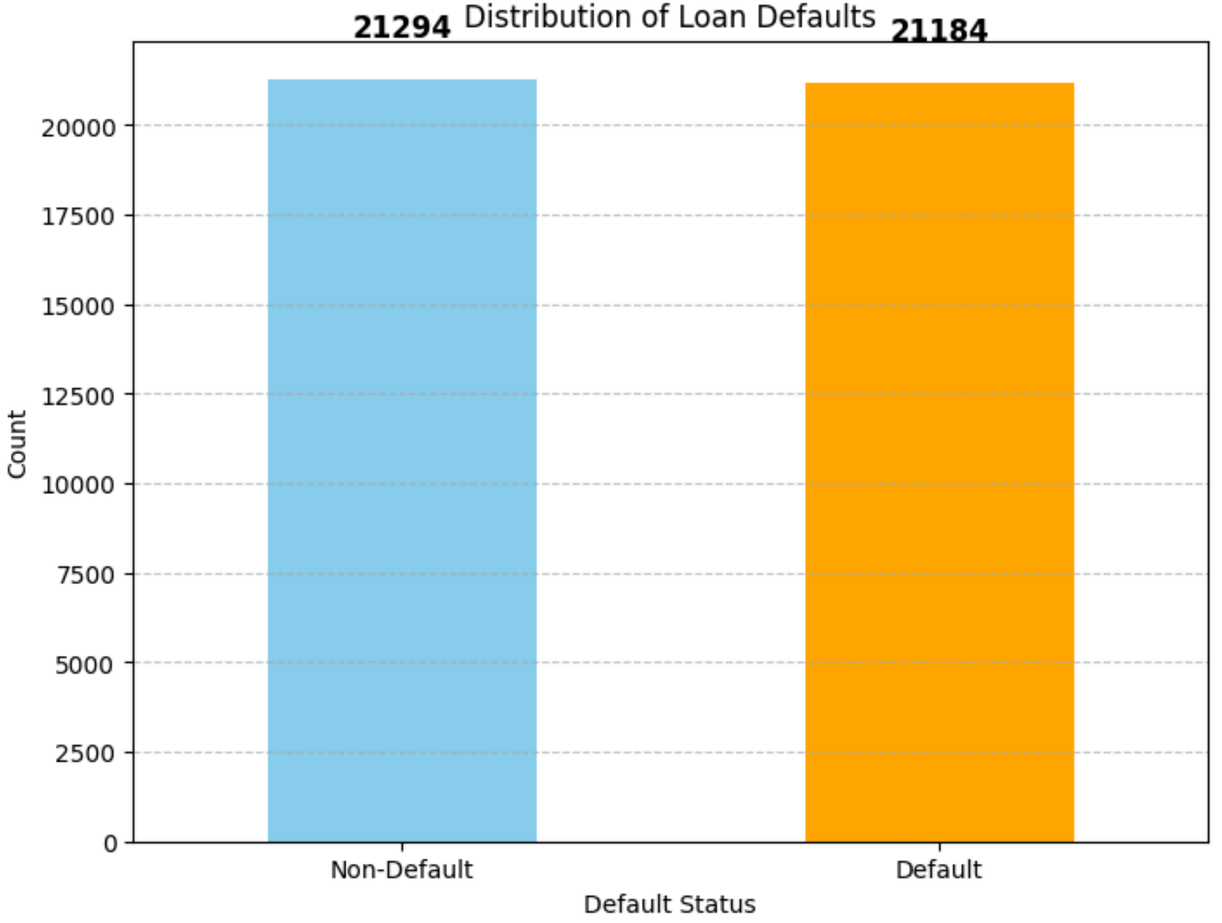
Loan Default Prediction

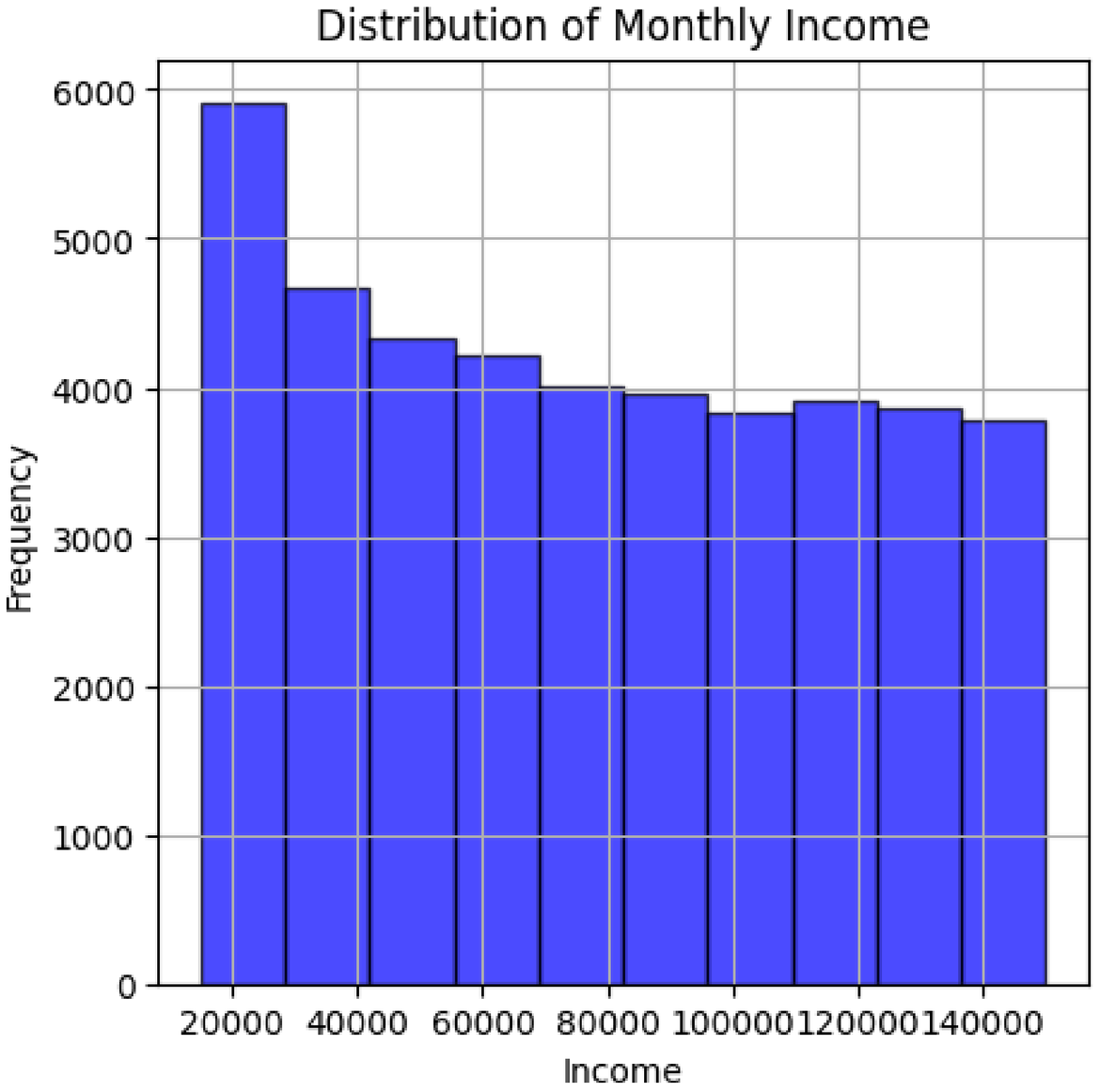
Output

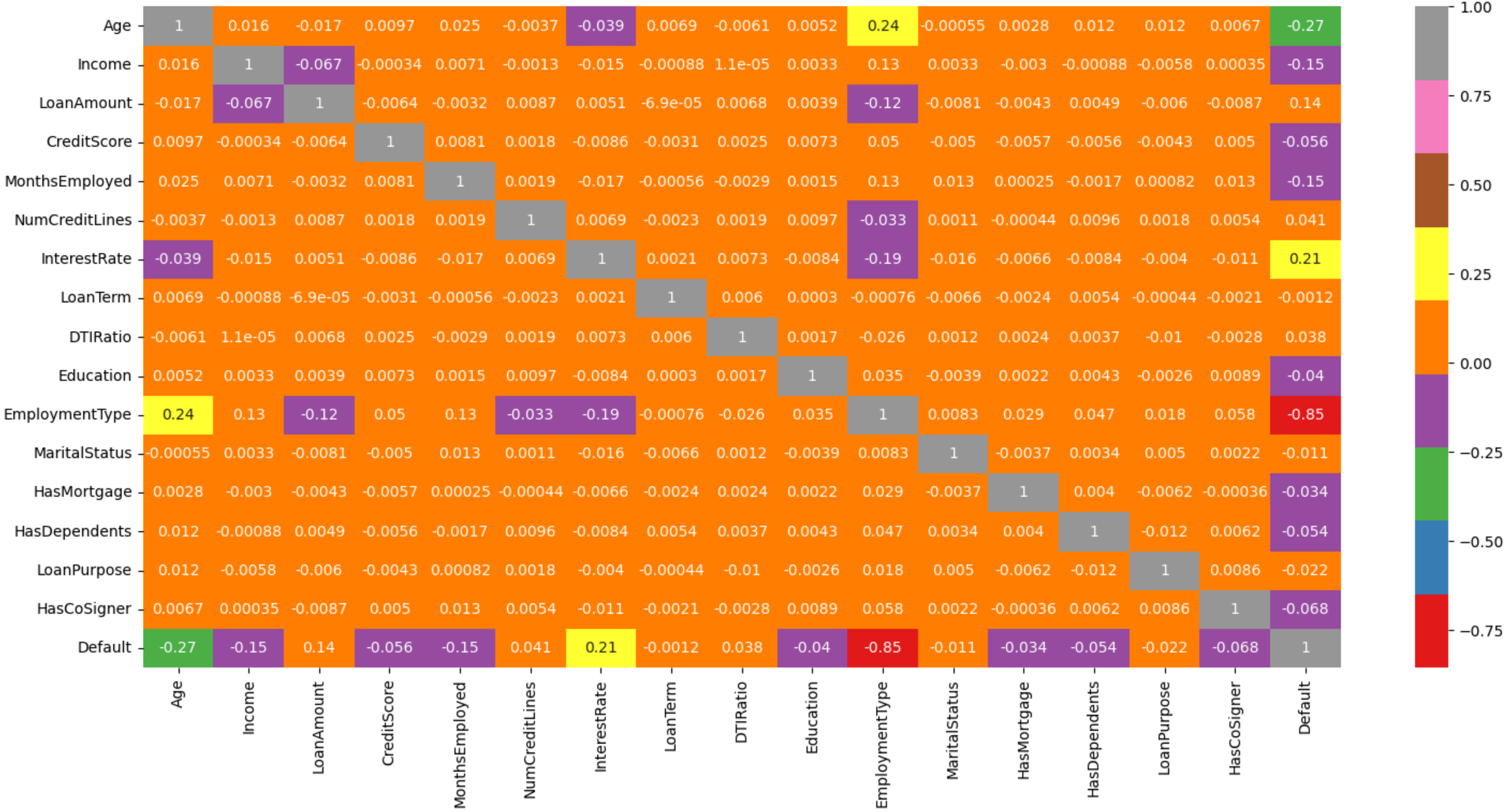
Data

- Age
- Income
- LoanAmount
- Credit Score
- Months Employed
- NumCreditLines
- Interest Rate
- Loan Term
- DTI Ratio
- Education
- Employment Type
- MaritalStatus
- Has Mortgage
- Has Dependents
- Loan Purpose
- Has Co Signer
- Loan Default

		Age	Income	LoanAmou	CreditScor	MonthsEm	NumCredit	InterestRa	LoanTerm	DTIRatio	Education	Employme	MaritalSta	HasMortga	HasDepen	LoanPurpc	HasCoSign	Default
21184	ZZY5HORA	22	123437	9163	582	108	1	8.39	36	0.69	PhD	Unemploy	Single	No	Yes	Education	No	0
21185	ZZZDJ23OK	30	118720	93079	619	51	1	5.41	60	0.65	High Schoc	Unemploy	Married	No	Yes	Other	No	0
21186	002EJYQA(	39	61393	167431	841	68	4	8.93	60	0.26	Master's	Full-time	Married	No	No	Education	No	1
21187	0042NIHB	24	95283	159587	539	50	4	20.61	36	0.63	Master's	Full-time	Divorced	No	Yes	Business	Yes	1
21188	004WNVB	52	111151	191688	843	54	4	23.93	12	0.27	High Schoc	Full-time	Married	No	Yes	Other	No	1
21189	00DWW2X	25	90046	155267	367	8	4	24.39	12	0.48	PhD	Full-time	Single	No	No	Education	No	1
21190	00W5DU2	23	124025	31936	664	44	2	24.61	48	0.52	Master's	Full-time	Divorced	No	No	Education	No	1
21191	013DGAEB	34	19665	126745	541	55	1	17.36	12	0.77	PhD	Full-time	Divorced	Yes	No	Business	Yes	1
21192	01HCWM6	36	35479	193321	489	14	1	17.92	48	0.15	High Schoc	Full-time	Single	No	Yes	Auto	No	1
21193	01LTICCI6'	32	18684	67703	743	61	4	3.07	48	0.11	Master's	Full-time	Divorced	Yes	Yes	Home	Yes	1
21194	01YOH9E9	50	129370	232769	810	16	3	11.37	48	0.11	High Schoc	Full-time	Married	No	Yes	Auto	Yes	1
21195	02D4E6RR	40	122111	178171	705	97	1	24.63	36	0.33	High Schoc	Full-time	Divorced	Yes	No	Business	No	1
21196	02GDBSIJD	47	38647	63255	449	5	3	20.13	24	0.29	PhD	Full-time	Divorced	Yes	Yes	Other	No	1
21197	02IKVK76V	23	22161	127466	753	103	3	17.45	12	0.54	PhD	Full-time	Married	Yes	Yes	Other	Yes	1
21198	02MA61ET	55	97676	157845	558	107	3	13.47	36	0.28	PhD	Full-time	Married	No	Yes	Education	Yes	1
21199	03IO75CAI	36	18586	97436	512	102	1	2.01	36	0.71	High Schoc	Full-time	Married	Yes	No	Business	No	1
21200	03L2IR9QT	58	36191	131550	736	7	2	23.25	60	0.64	PhD	Full-time	Single	Yes	No	Other	Yes	1
21201	03NS8IFYC	27	43996	166090	417	25	2	18.15	60	0.24	Master's	Full-time	Married	No	Yes	Auto	No	1


Distribution of Loan Defaults
21294
21184
Count
Default Status
Non-Default
Default
0
2500
5000
7500
10000
12500
15000
17500
20000


Distribution of Monthly Income
Frequency
Income
0
1000
2000
3000
4000
5000
6000
20000
40000
60000
80000
100000
120000
140000

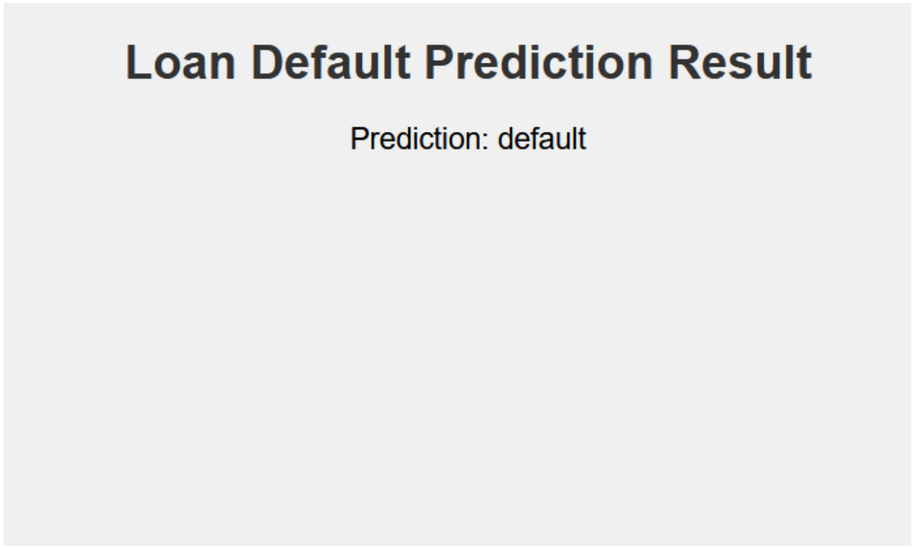

Age 1 0.016 -0.017 0.0097 0.025 -0.0037 -0.039 0.0069 -0.0061 0.0052 0.24 -0.00055 0.0028 0.012 0.012 0.0067 -0.27
Income 0.016 1 -0.067 -0.00034 0.0071 -0.0013 -0.015 -0.00088 1.1e-05 0.0033 0.13 0.0033 -0.003 -0.00088 -0.0058 0.00035 -0.15
LoanAmount -0.017 -0.067 1 -0.0064 -0.0032 0.0087 0.0051 -6.9e-05 0.0068 0.0039 -0.12 -0.0081 -0.0043 0.0049 -0.006 -0.0087 0.14
CreditScore 0.0097 -0.00034 -0.0064 1 0.0081 0.0018 -0.0086 -0.0031 0.0025 0.0073 0.05 -0.005 -0.0057 -0.0056 -0.0043 0.005 -0.056
MonthsEmployed 0.025 0.0071 -0.0032 0.0081 1 0.0019 -0.017 -0.00056 -0.0029 0.0015 0.13 0.013 0.00025 -0.0017 0.00082 0.013 -0.15
NumCreditLines -0.0037 -0.0013 0.0087 0.0018 0.0019 1 0.0069 -0.0023 0.0019 0.0097 -0.033 0.0011 -0.00044 0.0096 0.0018 0.0054 0.041
InterestRate -0.039 -0.015 0.0051 -0.0086 -0.017 0.0069 1 0.0021 0.0073 -0.0084 -0.19 -0.016 -0.0066 -0.0084 -0.004 -0.011 0.21
LoanTerm 0.0069 -0.00088 -6.9e-05 -0.0031 -0.00056 -0.0023 0.0021 1 0.006 0.0003 -0.00076 -0.0066 -0.0024 0.0054 -0.00044 -0.0021 -0.0012
DTIRatio -0.0061 1.1e-05 0.0068 0.0025 -0.0029 0.0019 0.0073 0.006 1 0.0017 -0.026 0.0012 0.0024 0.0037 -0.01 -0.0028 0.038
Education 0.0052 0.0033 0.0039 0.0073 0.0015 0.0097 -0.0084 0.0003 0.0017 1 0.035 -0.0039 0.0022 0.0043 -0.0026 0.0089 -0.04
EmploymentType 0.24 0.13 -0.12 0.05 0.13 -0.033 -0.19 -0.00076 -0.026 0.035 1 0.0083 0.029 0.047 0.018 0.058 -0.85
MaritalStatus -0.00055 0.0033 -0.0081 -0.005 0.013 0.0011 -0.016 -0.0066 0.0012 -0.0039 0.0083 1 -0.0037 0.0034 0.005 0.0022 -0.011
HasMortgage 0.0028 -0.003 -0.0043 -0.0057 0.00025 -0.00044 -0.0066 -0.0024 0.0024 0.0022 0.029 -0.0037 1 0.004 -0.0062 -0.00036 -0.034
HasDependents 0.012 -0.00088 0.0049 -0.0056 -0.0017 0.0096 -0.0084 0.0054 0.0037 0.0043 0.047 0.0034 0.004 1 -0.012 0.0062 -0.054
LoanPurpose 0.012 -0.0058 -0.006 -0.0043 0.00082 0.0018 -0.004 -0.00044 -0.01 -0.0026 0.018 0.005 -0.0062 -0.012 1 0.0086 -0.022
HasCoSigner 0.0067 0.00035 -0.0087 0.005 0.013 0.0054 -0.011 -0.0021 -0.0028 0.0089 0.058 0.0022 -0.00036 0.0062 0.0086 1 -0.068
Default -0.27 -0.15 0.14 -0.056 -0.15 0.041 0.21 -0.0012 0.038 -0.04 -0.85 -0.011 -0.034 -0.054 -0.022 -0.068 1
Age
Income
LoanAmount
CreditScore
MonthsEmployed
NumCreditLines
InterestRate
LoanTerm
DTIRatio
Education
EmploymentType
MaritalStatus
HasMortgage
HasDependents
LoanPurpose
HasCoSigner
Default
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75

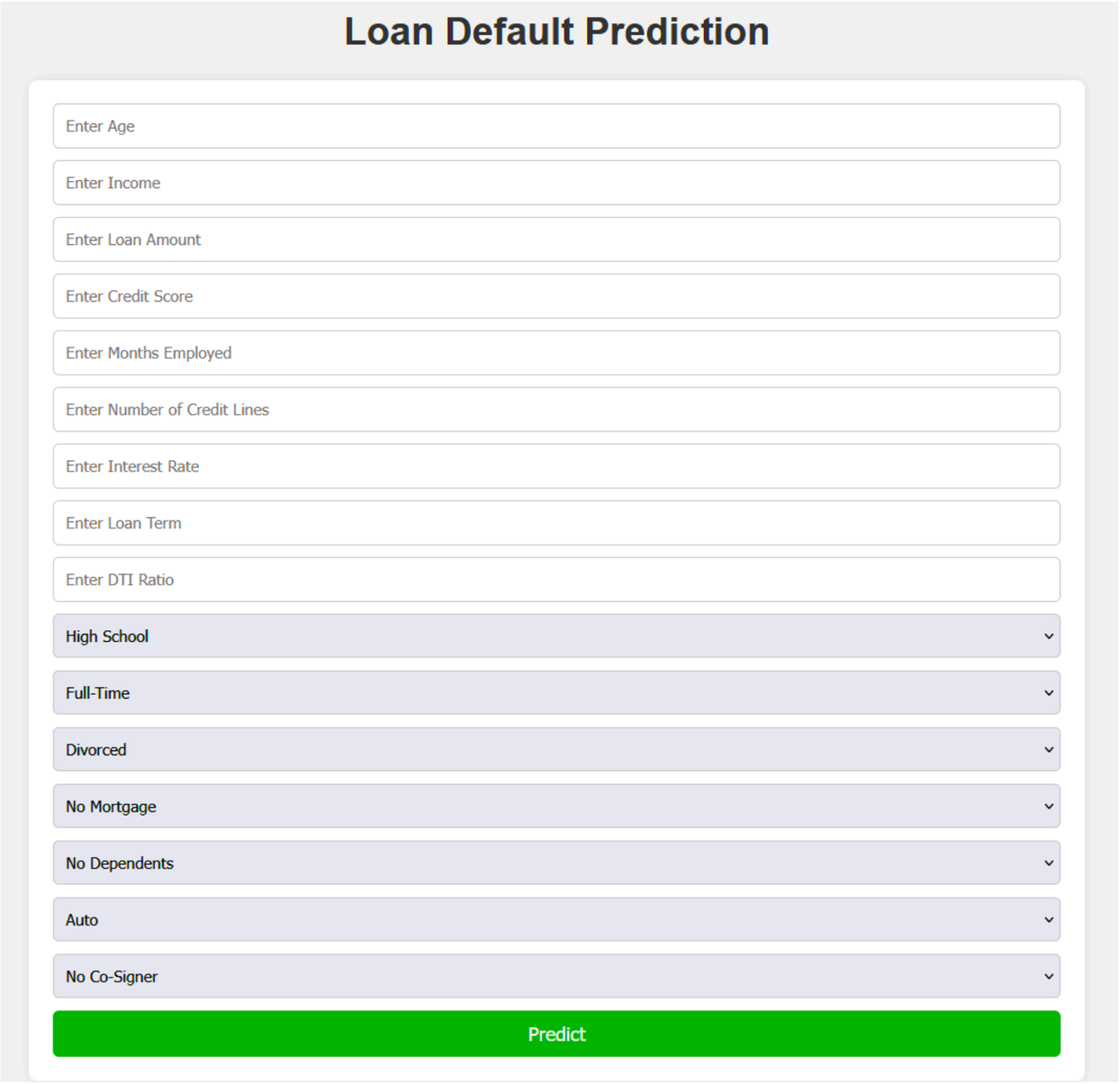
Machine Learning Models

models	accuracy
RandomForestClassifier	0.99
XGBClassifier	0.99
KNeighborsClassifier	0.92
GaussianNB	0.80
LogisticRegression	0.69

Result

exported model: Random forest





Loan Default Prediction

30

118720

93079

619

51

1

5.41

60

0.65

High School

Unemployed

Married

No Mortgage

Has Dependents

Other

No Co-Signer

Predict

Loan Default Prediction

39

61393

167431

841

68

4

8.93

60

0.26

Master's

Full-Time

Married

No Mortgage

No Dependents

Education

No Co-Signer

Predict

Loan Default Prediction Result

Prediction: no default

Loan Default Prediction Result

Prediction: default

Conclusion

- By identifying high-risk loans early, financial institutions can take proactive measures to mitigate losses.
- Some of the key factors influencing loan default are age, income, loan amount,credit score, and employment status.
- Future steps include exploring advanced algorithms and real time data for enhanced model performance.

Reference

1. Madaan, Mehul & Kumar, Aniket & Keshri, Chirag & Jain, Rachna & Nagrath, Preeti. (2021). Loan default prediction using decision trees and random forest: A comparative study. IOP Conference Series: Materials Science and Engineering. 1022. 012042. 10.1088/1757-899X/1022/1/012042.

2.ManjeetKumar, Vishesh Goel, Tarun Jain, Sahil Singhal,DR. Lalit Mohan Goel. (2018). Neural Network Approach To Loan Default Prediction, International Research Journal of Engineering and Technology (IRJET) , p-ISSN: 2395-0072

3.Xu Zhu, Qingyong Chu, Xinchang Song, Ping Hu, Lu Peng,Explainable prediction of loan default based on machine learning models,Data Science and Management, Volume 6, Issue 3,2023,Pages 123-133,ISSN 2666-7649,https://doi.org/10.1016/j.dsm.2023.04.003

Thank You !