LOAN APPROVAL PREDICTION SYSTEM

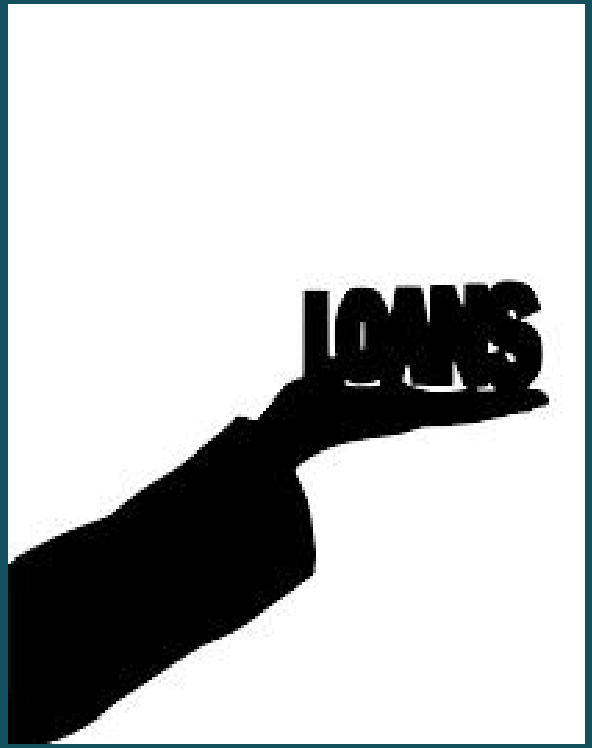
BY GROUP 13

GROUP MEMBERS

1. Maurice Njogu - maurice.njogu@student.moringaschool.com
2. James Mungai - james.mungai@student.moringaschool.com
3. Lynn Ndero - lynn.ndero@student.moringaschool.com
4. Mumbi Odhiambo - mumbi.odhiambo@student.moringaschool.com
5. Julie Chepngeno - julie.chepngeno@student.moringaschool.com
6. Sam Odongo - sam.odongo@student.moringaschool.com

INTRODUCING

- The loan Approval system which is envisioned as a state-of-the-art tool that will revolutionize the loan approval process of both lenders and borrowers.
- By incorporating Machine Learning algorithms our system aims to revolutionize and automate the decision-making process, leading to faster and more efficient loan approvals.
- The Loan approval system aims to provide a fair and unbiased loan approval process as it involves the elimination of human bias and subjective judgement.



CONTENT:

- ★ PROBLEM STATEMENT
- ★ GOALS
- ★ MODELS
- ★ RESULTS
- ★ CONCLUSION
- ★ RECOMMENDATIONS

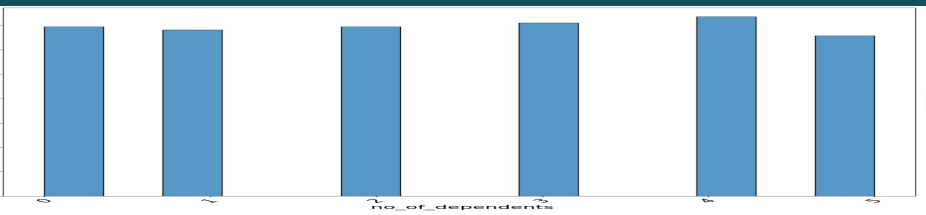
PROBLEM STATEMENT

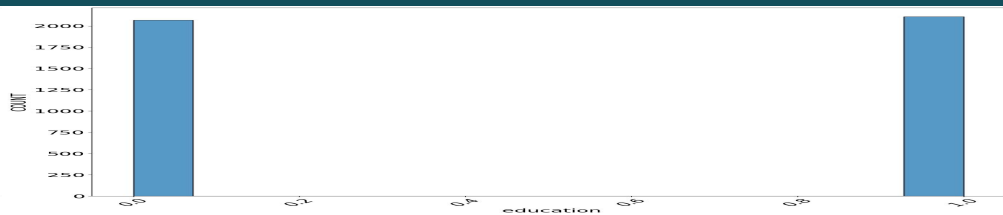
- The loan approval process is a critical aspect of the financial industry.
- Its accuracy directly impacts the success and stability of lending institutions
- Manual evaluation used by most companies currently tends to be to be time-consuming and ends up having very many inconsistencies more often than not.

GOALS

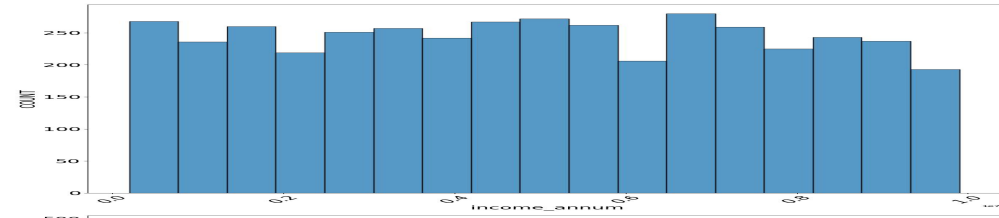
- Efficiency and speed during the loan approval process
- Risk mitigation for the lenders as it involves analysis of historical data and applicant characteristics thus reducing the risk of defaulters getting loans.

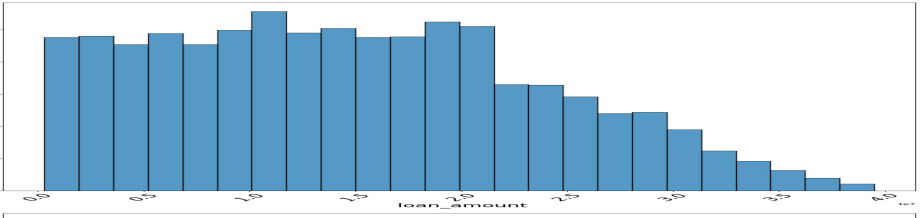
BAR GRAPHS SHOWING THE DISTRIBUTION OF THE VARIABLES IN THE DATASET

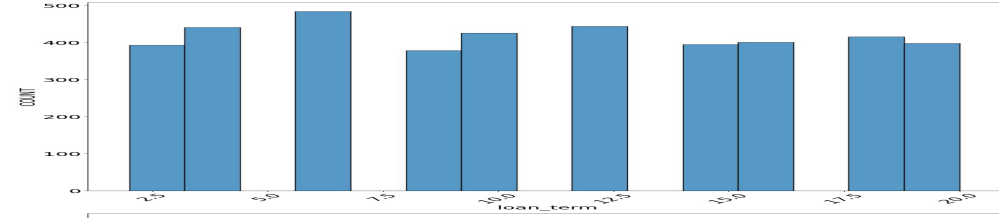

0
1
2
3
4
5
no_of_dependents

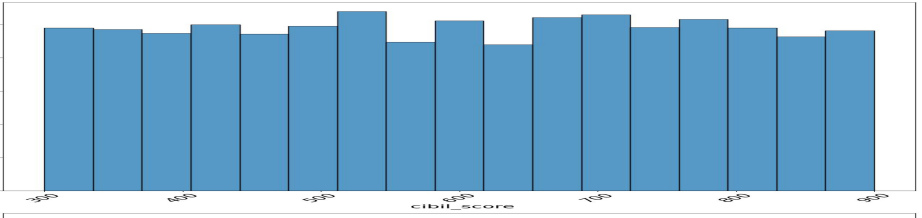

COUNT
2000
1750
1500
1250
1000
750
500
250
0
0.0
0.2
0.4
0.6
0.8
1.0
education

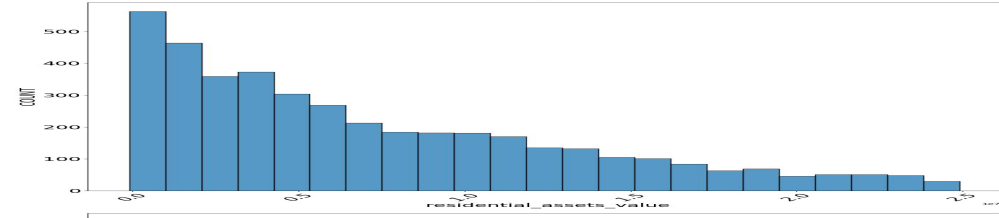

0.0
0.2
0.4
0.6
0.8
1.0
self_employed

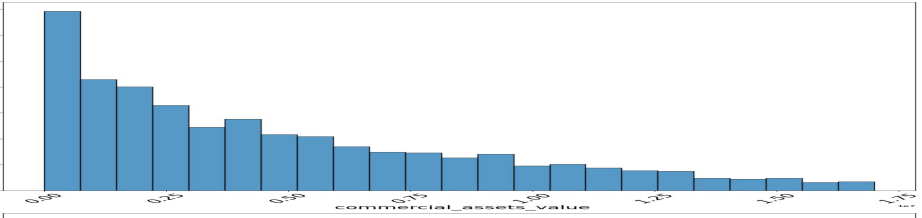

COUNT
250
200
150
100
50
0
0.0
0.2
0.4
0.6
0.8
1.0
income_annum

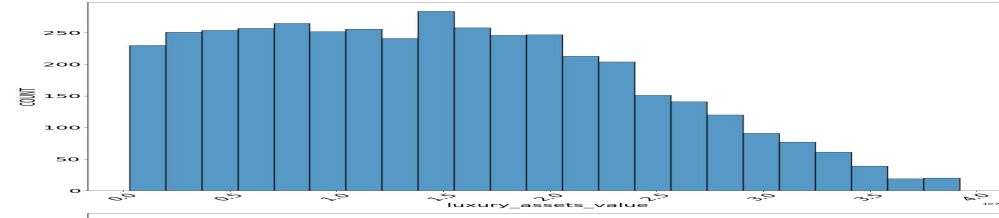

0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
loan_amount


COUNT
500
400
300
200
100
0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
loan_term


300
400
500
600
700
800
900
cibil_score

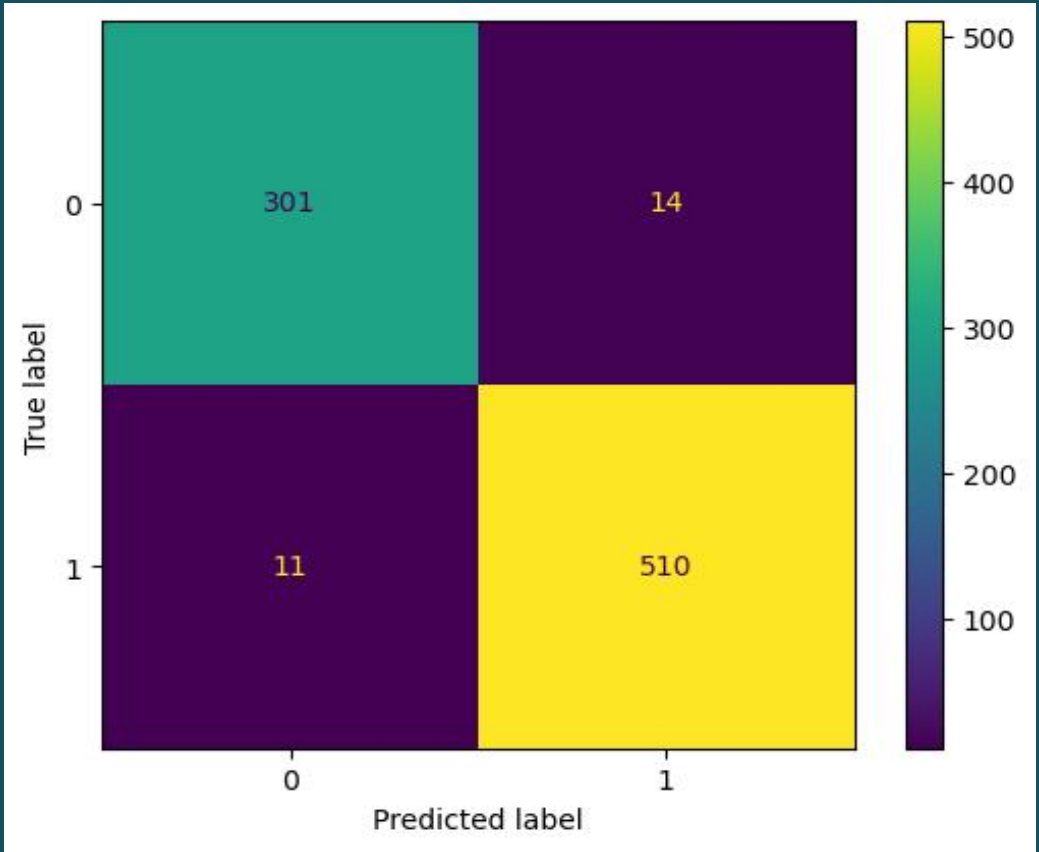

COUNT
500
400
300
200
100
0
0.0
0.5
1.0
1.5
2.0
2.5
residential_assets_value


0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
commercial_assets_value

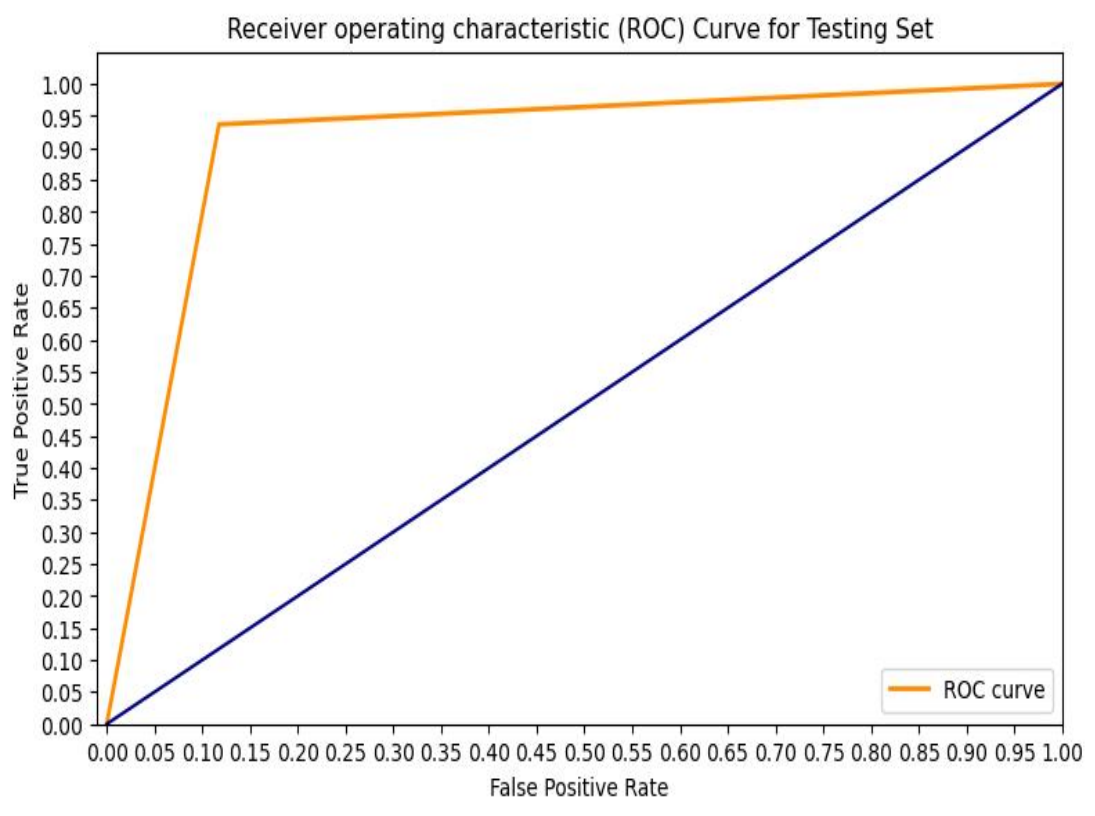

COUNT
250
200
150
100
50
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
luxury_assets_value

MODELS USED

- KNN model: uses nearby points to generate predictions of the scaled testing data.
- Logistic Regression: was used to predict the probability of whether a loan application will be approved or rejected based on input features.
- Gaussian Naive Bayes Classification: we looked at multiple observations and conditional probabilities to help increase predictions of loan approval and aid in making well-informed decisions.
- Random Forest: was used to build multiple decision trees using different subsets of the training data and different subsets of features.



- Decision tree: was used to predict whether a loan applicant will be approved or not and results displayed as shown.



- SUPPORT VECTOR MACHINE:

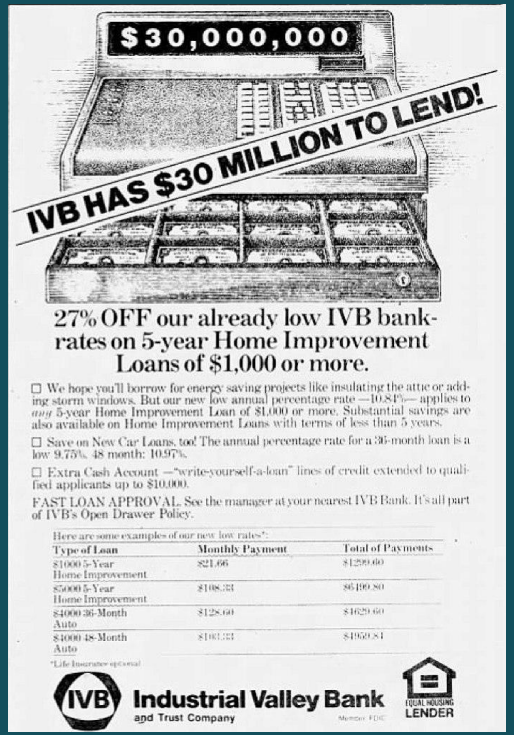
was used to help distinguish between the positive and negative classes effectively.

FINDINGS

After the evaluation of various models, we achieved the following metrics:

- ➔ KNN Model Accuracy(0.92), precision (0.96), recall(0.90), f1 score(0.90)
- ➔ Logistics Regression Model Accuracy(0.92)
- ➔ Decision Tree Model Accuracy (0.97), precision(0.97), recall(0.96), f1 score(0.99)
- ➔ Random Forest Model Accuracy(0.98), precision(0.98), recall(0.98), f1 score(0.98)
- ➔ Gaussian Naive Bayes Classification Accuracy(0.98)
- ➔ Support Vector Machine Accuracy (0.95)

CONCLUSIONS



- The Random Forest Algorithm is a strong choice for our Loan Approval Prediction System because it has the ability to handle non-linear relationships and interaction between features which most likely contributed to its high accuracy level of 97%.
- The high accuracy level is a ky indicator that this model is effectively capturing patterns in the data and differentiating between approved and rejected loan appliaions.

RECOMMENDATIONS

Based on our findings, we were able to come up with the following recommendations:

A. Integration of the Random Forest Algorithm with the pre existing loan approval systems
B. Providing the loan officers ad decision-makers with training and education about how the Random Forest Model works, how to interpret its predictions and using the model's insights to make informed decisions.
C. Developing clear guidelines on using the model's predictions in loan approval decisions
D. Fostering ongoing collaborations amongst data scientists.domain officers and loan officials to ensure exchange of continuously evolving information on the model between them
E. Ensuring the model is updated regularly.
F. Risk management measures which might involve coming up with contingency plans for scenarios where the model's contingency plans might fail.

LOAN

APPROVED.