VISHNU
UNIVERSAL LEARNING

BVRIT HYDERABAD

College of Engineering for Women

Loan Default Prediction

Team No: 2
Team Members:
Y.Jansi , 20wh1a1208
P.Jyothi Sivani, 20wh1a1209
Rida Arshad Khan, 20wh1a1210
R.Sangeetha, 20wh1a1211

VISHNU
UNIVERSAL LEARNING

AGENDA

- Problem statement
- Python Packages used
- Algorithm
- output
- Comparison table

VISHNU
UNIVERSAL LEARNING

Problem Statement

Loan Default Dataset contains 10,000 clients and 5 different features for each one of them. Employed is a binary feature where 1 represents people that are employed and 0 represents those who are unemployed. Bank Balance is the amount of money the client has available in his/her bank account. Annual Salary is the annual salary for each client and, Defaulted is the target variable where 0 represents those who didn't default and 1 represents those who defaulted. Process the data (if required). Use Logistic Regression to check whether a client will default or not and calculate your model accuracy.

VISHNU
UNIVERSAL LEARNING

Python Packages used

- numpy
- pandas
- seaborn
- matplotlib.pyplot
- imblearn
- sklearn

VISHNU
UNIVERSAL LEARNING

Algorithm

- Logistic Regression
- KNN
- Random Forest
- Gradient boosting

VISHNU
UNIVERSAL LEARNING

Logistic Regression

Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables,the outcome must be a categorical or discrete value. It can be either Yes or No, 0 or 1, true or False

VISHNU
UNIVERSAL LEARNING

K nearest Neighbour

K-Nearest Neighbors Algorithm. The k-nearest neighbors algorithm, also known as KNN or k-NN, is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point.

VISHNU
UNIVERSAL LEARNING

Random Forest

Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

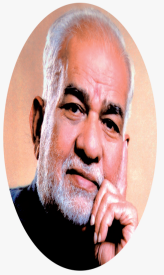
VISHNU
UNIVERSAL LEARNING

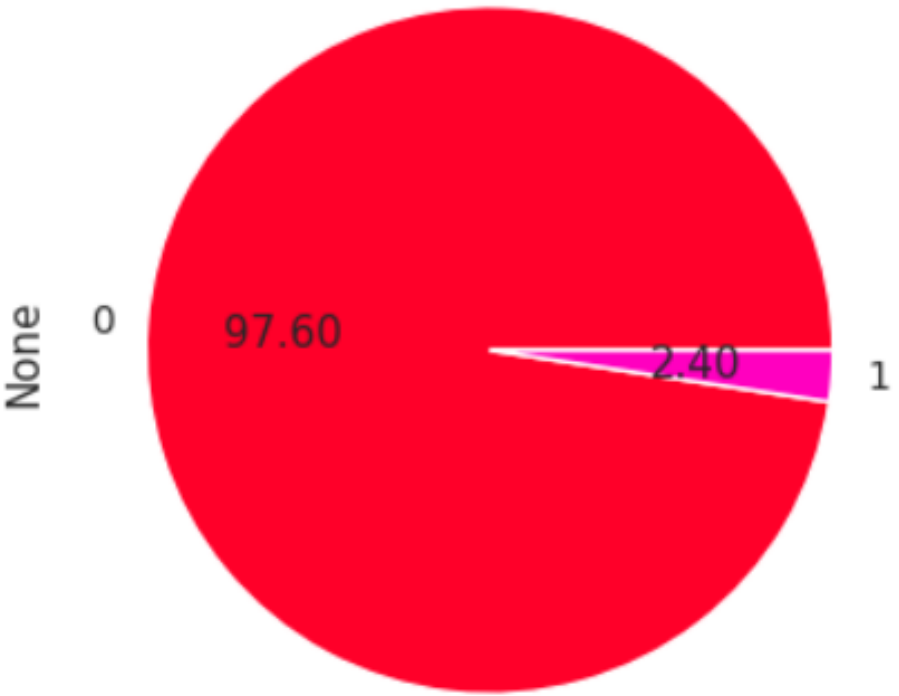
Gradient Boosting

Gradient boosting is a method standing out for its prediction speed and accuracy, particularly with large and complex datasets.

VISHNU
UNIVERSAL LEARNING

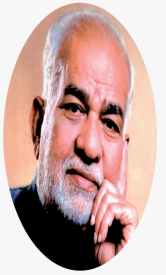
Output





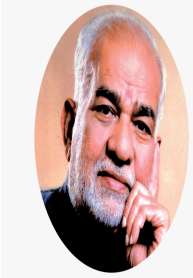
VISHNU
UNIVERSAL LEARNING

Comparison Table



	Models	Accuracies
0	Logistic Regression	0.96125
1	K-Neighbors Classifier	0.96250
2	Random Forest Classifier	0.97375
3	Gradient Boosting Classifier	0.96875

VISHNU
UNIVERSAL LEARNING



THANK YOU