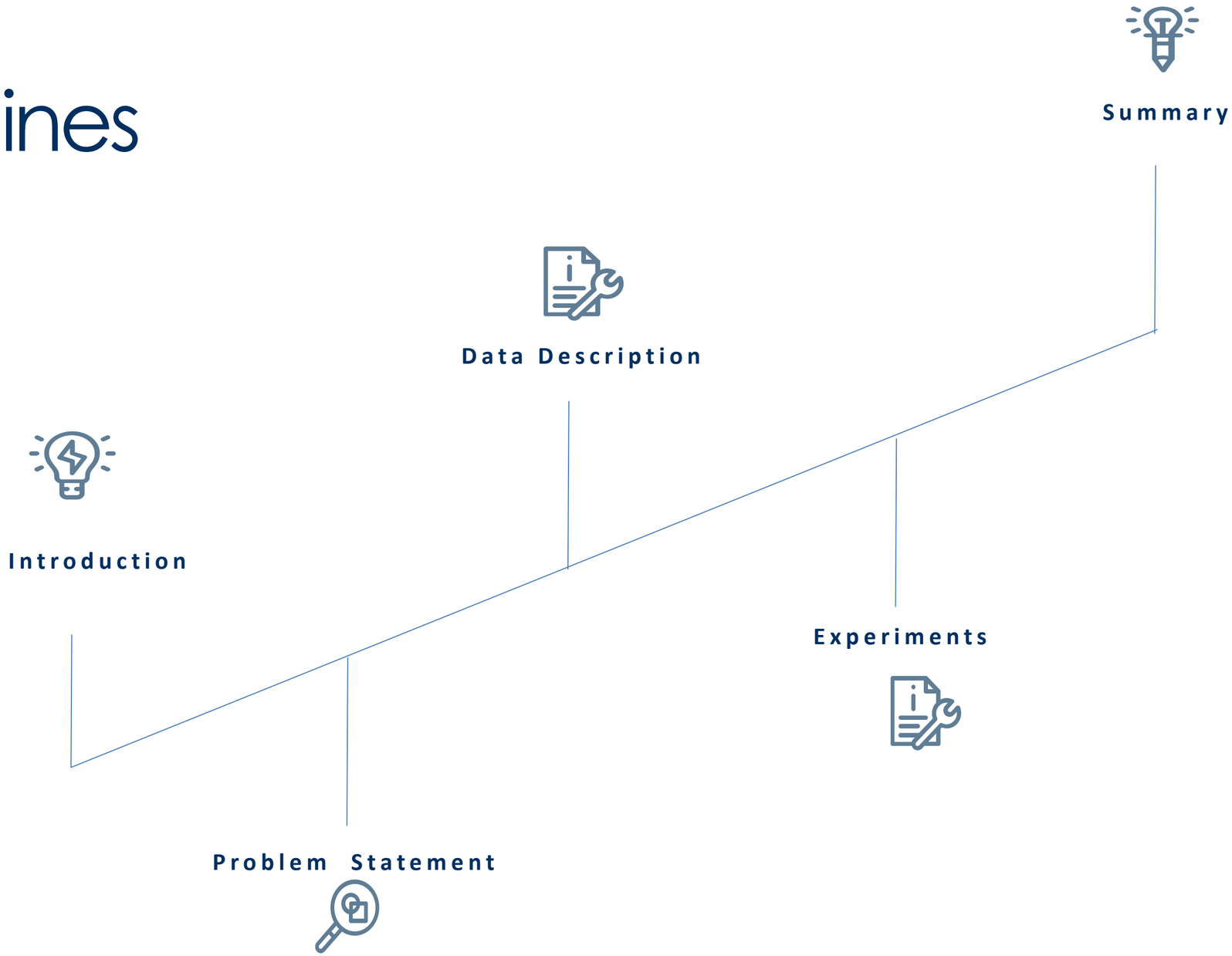
Classification for Housing Loans

Presented by

Naif Alsulaihem
Nasser Alshehri

Outlines


Summary
Data Description
Introduction
Experiments
Problem Statement

Introduction

Housing is one of the necessities of life that must be owned. Our company provides a financing service for people to obtain housing. And the agreement with the customer to pay this amount in loans within a specified period of time

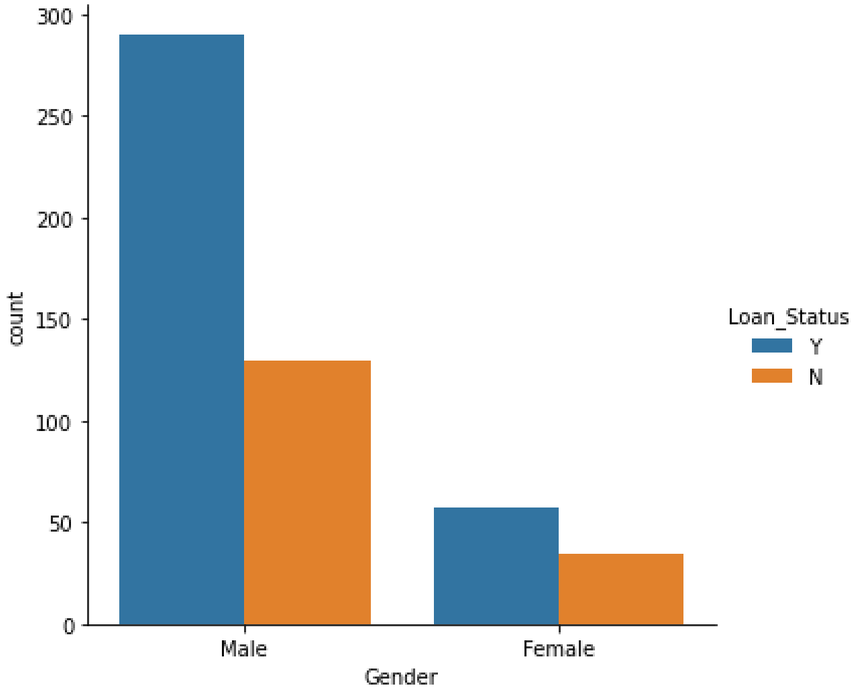
Problem Statement

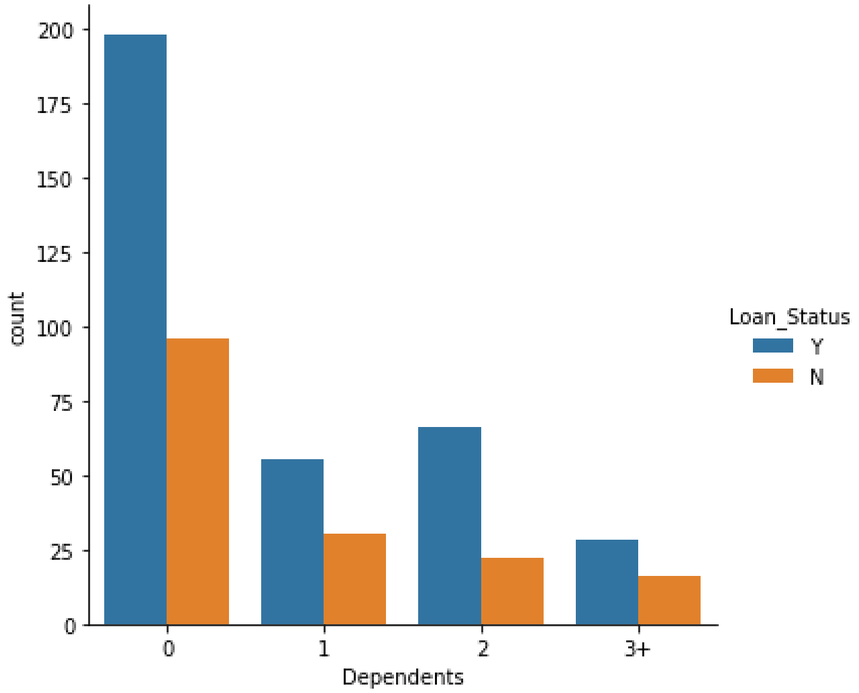
We face the problem of defaulting on some customers' payments, which makes us form a team to analyze customers and find out who will default, and this is very difficult, so we will do this process by machine learning algorethems and trying to predict through data of customer.

Data Description

- The data is customer data such as gender, Married and education..etc.
- The data was extracted from Kaggle and it contains 13 columns .
- The target is Loan Status.

Data Description





Experiments

Baseline	train	0.76
	val	0.76
Decision Tree	Train	0.814
	Val	0.756
Random Forest	Train	0.808
	Val	0.813
Knn	Train	0.812
	Val	0.80
Logistic Regression	train	0.816
	val	0.814
Naïve Bayes	train	0.794
	val	0.812

SVM	train	0.688
	val	0.699

Experiments - Logistic Regression

Train

0.816

Test

0.814

F 1

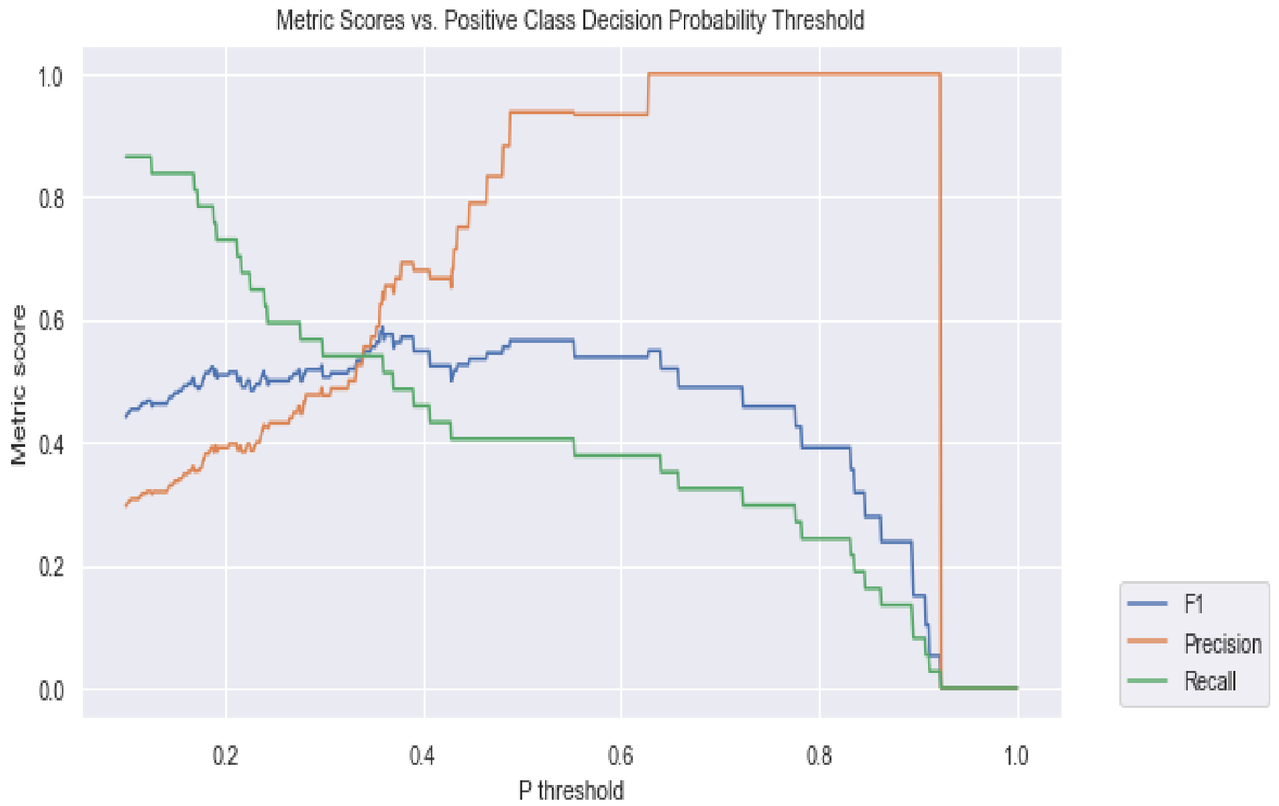
0.56

Precision

0.93

Recall

0.40



Summary

- We have used many algorithms to predict the loan status
- The highest rated algorithm is **Logistic Regression**

Thank You