!!! WIP Play this in your Browser: https://www.beautiful.ai/player/-NPFy-xefouUTwx8N-7B

Predicting Loans with Machine Learning and Streamlit

Presented by Team#3: Emma, Lucas and Nav

6 March 2023

THE PROBLEM

Industry Overview

Most of the loan application process is laborious. The outcome is a simple 'Yes' or 'No' without further explanation.

Rejection is tough - don't let it make your business tougher. There is communication gap to be bridged.

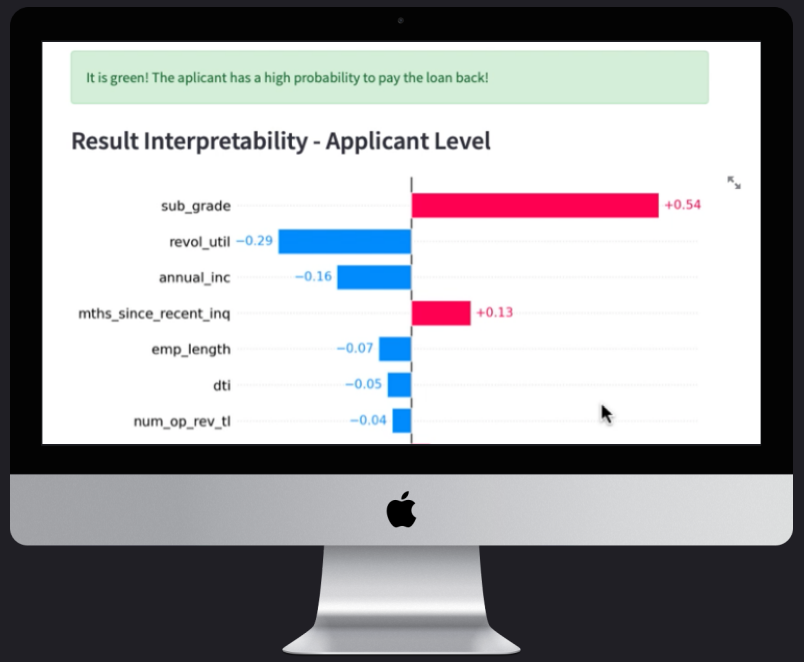
SOLUTION

Unlocking the Insights of Machine Learning

Despite its powerful capabilities, machine learning's charm was often hidden beneath technical jargon and the mystery of its "black box" operations.

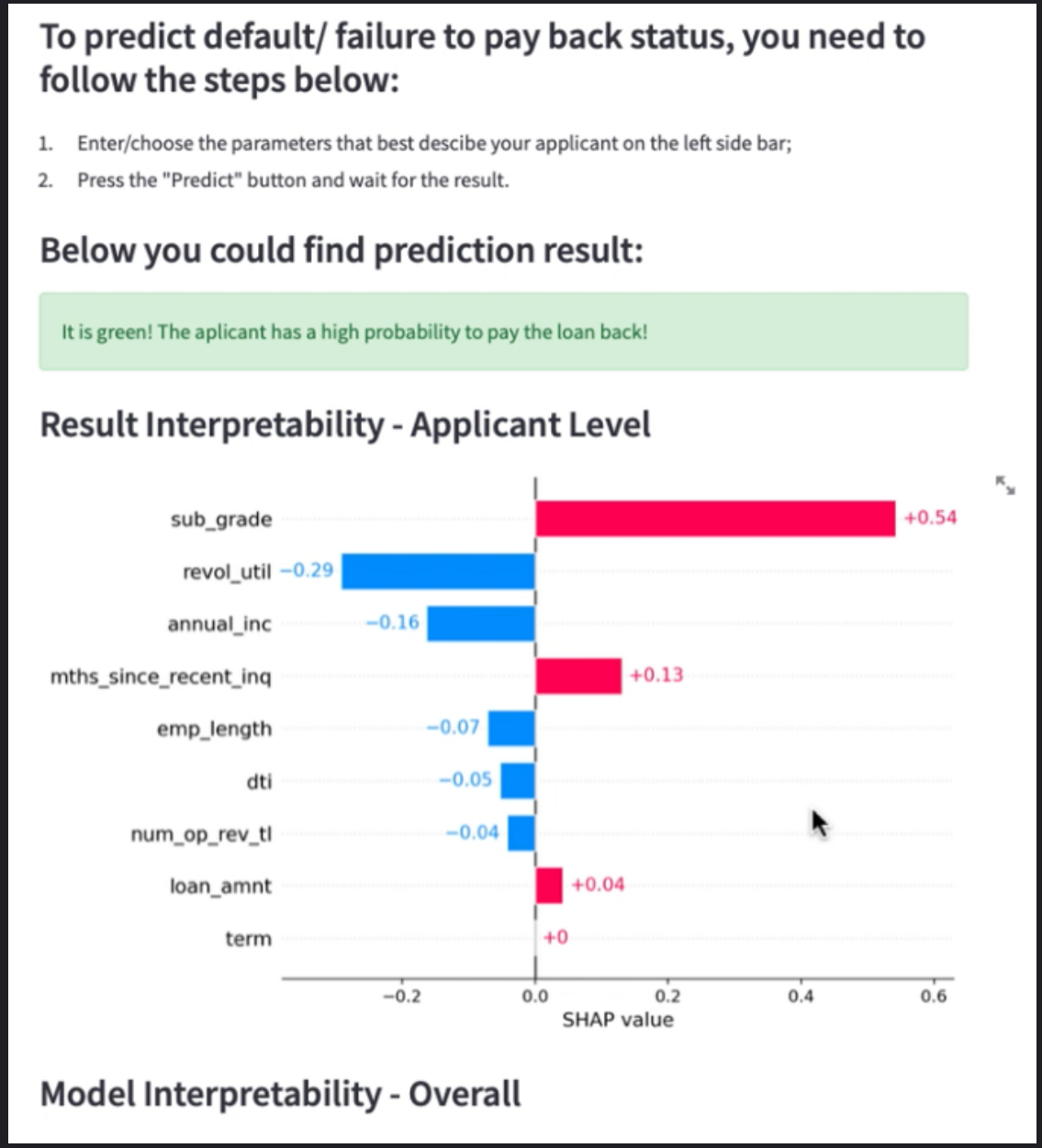
The web application created with Streamlit used visual exploration to demonstrate the power of machine learning, showcasing the dynamic input feature and its SHAP value that drove the decision.

This improved the application's transparency and helped to close the communication gap.



What are SHAP Values?

- Calculation that represents the relative influence of each input features within the algorithm.
- It can be thought of as the 'weighting' that each feature had upon the final result of creditworthiness that the model predicted.



Why Should I Care About SHAP Values?

Unveiling the Creditworthiness Prediction:

SHAP values provide an intuitive way to understand the impact of individual features on a model's predictions

SHAP values can be used to identify important features and detect potential bias in a model

SHAP values can be used to improve model performance by identifying redundant or irrelevant features

Discover How Your Financial Profile Affects Your Risk with SHAP Values

MISSION

Helping customer to learn from their data

Lending Club dataset is used for model training

Available from Kaggle

Total observations

2.26m

Total features

150

Annual income:

65k

Loan amount (median):

$13k

Avg interest rate

13%

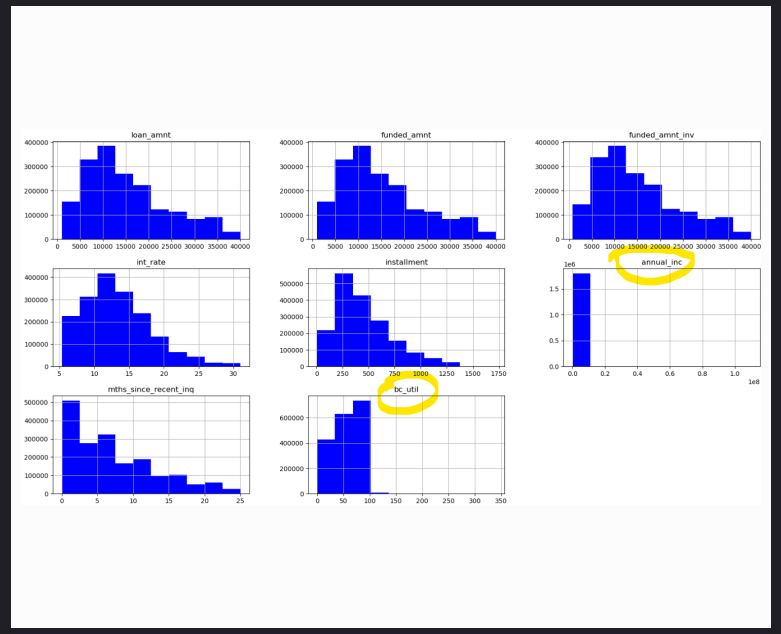
Selected features

17

Big Data challenges

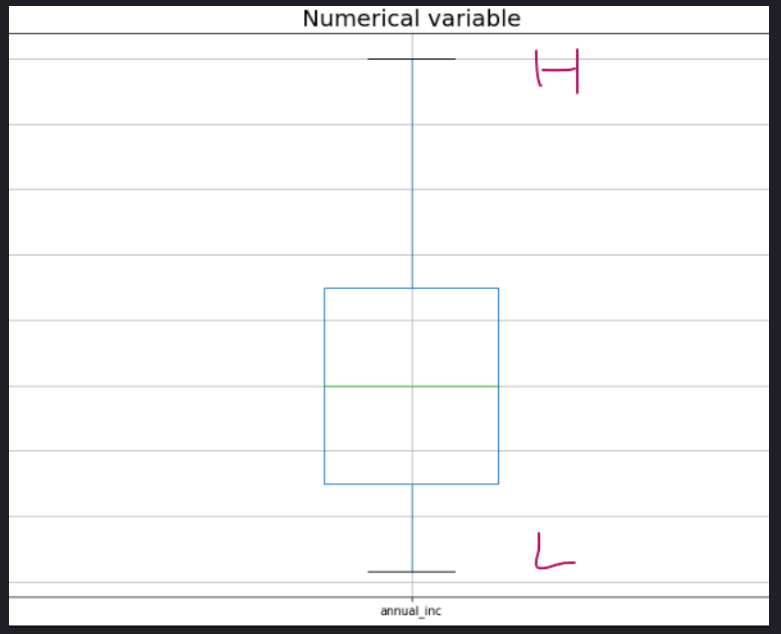
Univariate Analysis

Two major components: Outliers and Skewness



Outlier

can significantly distort the feature distribution and ML algorithms such as: Linear Regression, Logistic Regression, Support Vector Machine

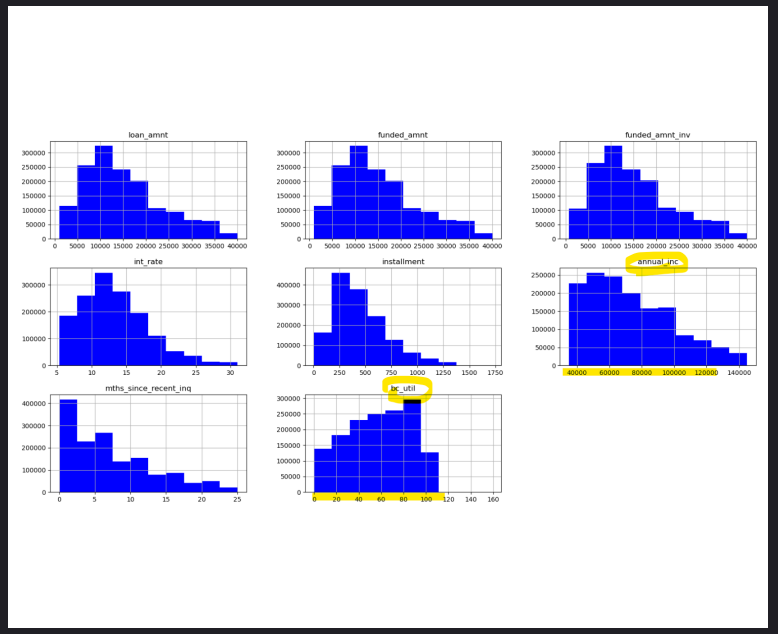


Remove outlier

low_q = df.quantile(0.08)

high_q = df.quantile(0.92)

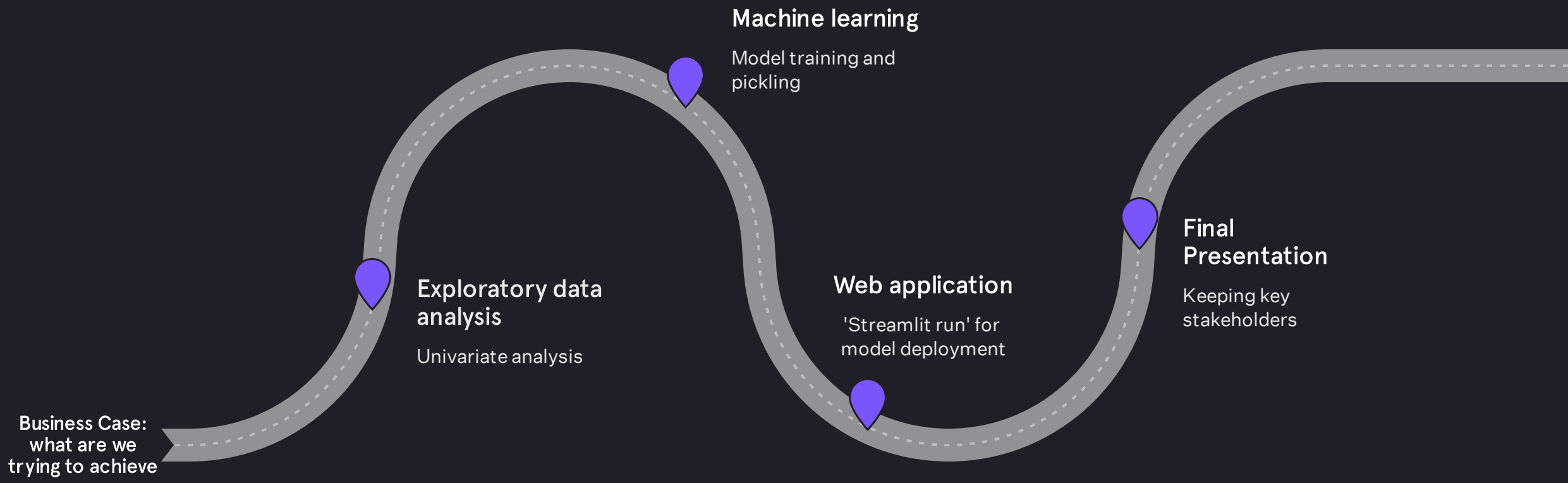
new_df = (df< q_hi) & (df > q_low)]



Distribution after dropping outliers

If you have a really good sense of what range the data should fall in, like people's ages, you can safely drop values that are outside of that range.

Roadmap


Machine learning
Model training and pickling
Exploratory data analysis
Univariate analysis
Web application
'Streamlit run' for model deployment
Final Presentation
Keeping key stakeholders
Business Case: what are we trying to achieve

Demos


Deloitte.
log.md
index.js
var express = require('express');
var router = express.Router();
var User = require('../models/user');
// GET /register
router.get('/register', function(req, res, next) {
return res.render('register', { title: 'Sign Up' });
});
// POST /register
router.post('/register', function(req, res, next) {
if (req.body.email &&
req.body.name &&
req.body.favoriteBook &&
req.body.password &&
req.body.confirmPassword) {
// confirm that user typed same password twice
if (req.body.password !== req.body.confirmPassword) {
var err = new Error('Passwords do not match.');
err.status = 400;
return next(err);
}
// create object with form input
var userData = {
email: req.body.email,
name: req.body.name,
favoriteBook: req.body.favoriteBook,
password: req.body.password
};
// use schema's `create` method to insert document into Mongo
User.create(userData, function (error, user) {
if (error) {
JavaScript-Authentication-Mongo-Express/routes/index.js 1:1
LF UTF-8 JavaScript 0 files