DETAIL PROJECT REPORT (DPR)

ADULT CENSUS INCOME PREDICTION

Revision Number: 1.0

Last date of revision: 27/09/2023

Table of Contents

1. Objective 3

2. Benefits 3

3. Architecture 3

3.1 Training Flow.....3

3.2 Prediction Flow 4

3.3 Application Flow 4

4. Data Preprocessing 5

5. Model Training 5

6. Application 5

7. Q & A 6

1 Objective

Development of a predictive model which can classify the input values of a user into two of the following:

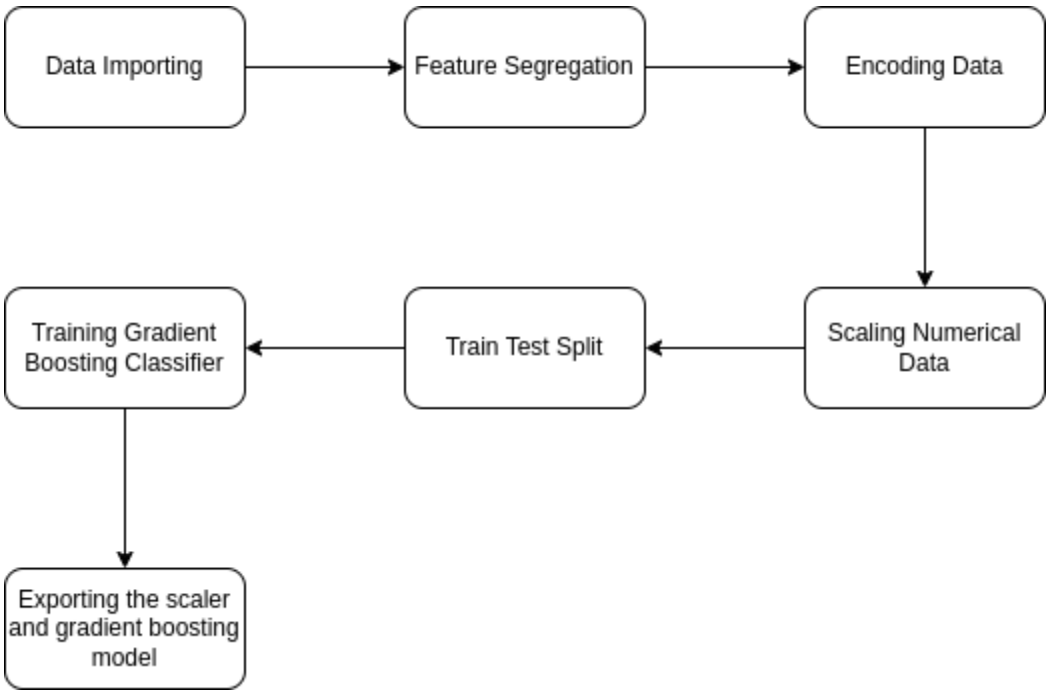
1. <50K income
2. >= 50K income

2 Benefits

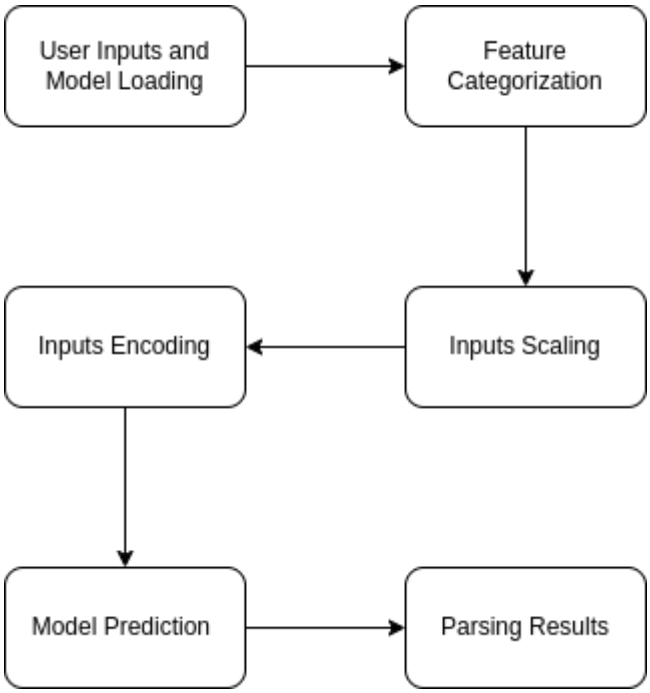
1. It will help users to make wise decisions on their finances
2. It will show how their financial standing is.

3 Architecture

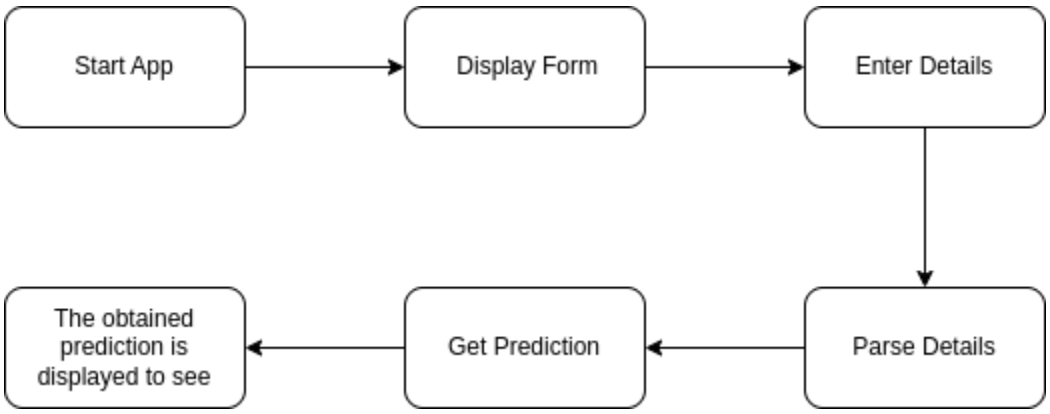
3.1 Training Flow



3.2 Prediction Flow



3.3 Application Flow



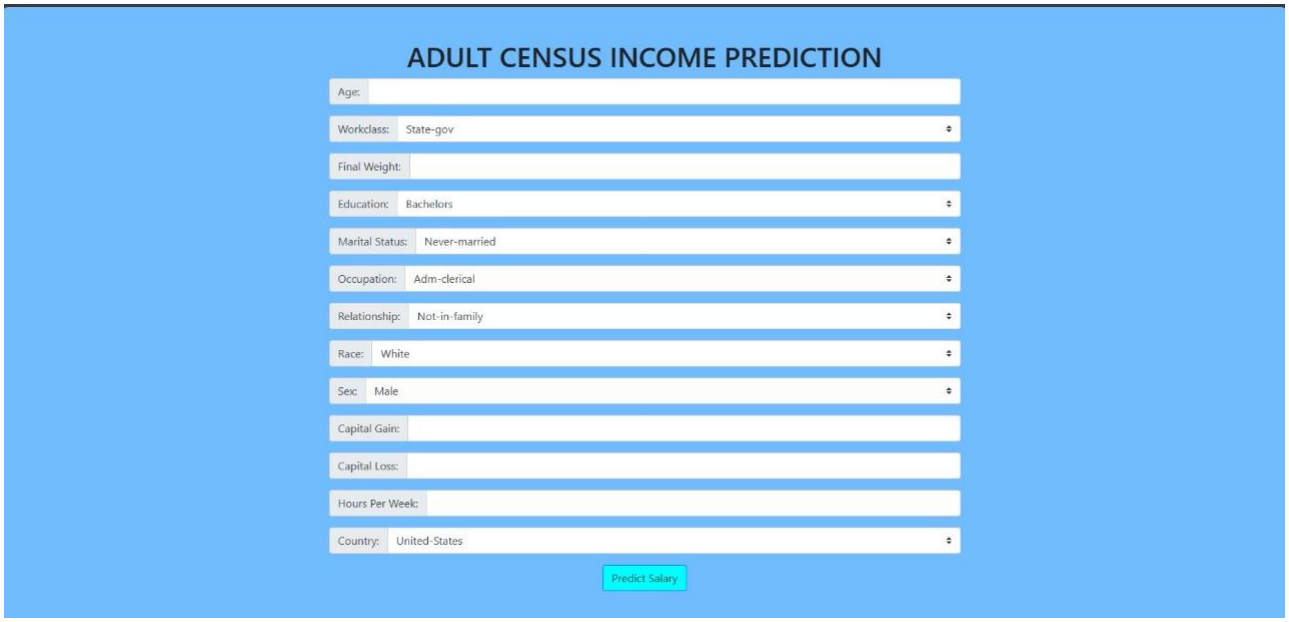
4 Data Preprocessing

The data is received from the form and is pre-processed and scaled using the standard scaler and passed to the model.

5 Model Training

The data is pre-processed by clearing the nans and scaled using Standard Scaler, Then The scaled dataset is given to Gradient boosting model and the model is trained. The resulting model is getting an accuracy of 89%

6 Application



The model is incorporated inside an Flask application It handles all the data delivery from the form to the model and preprocessing of the user input. Then gets the prediction from the model and returns to output page. The final application is deployed in render.

render

You can visit the website at:

https://adultcensusincomeprediction.onrender.com/

7 Q & A

Q1: What is the primary objective of the Adult Census Income Prediction project?

A1: The primary objective of the Adult Census Income Prediction project is to develop a predictive model that classifies user input values into one of two income categories: "<50K income" or ">= 50K income."

Q2: What are the two income categories that the predictive model classifies users into?

A2: The predictive model classifies users into two income categories: "<50K income" and ">= 50K income."

Q3: How is the data pre-processed in this project, and what scaling technique is used?

A3: The data is pre-processed by clearing the missing values (nans) and scaled using the Standard Scaler technique. This scaled data is then passed to the model for training and prediction.

Q4: What is the accuracy achieved by the Gradient Boosting model in this project?

A4: The Gradient Boosting model in this project achieves an accuracy of 89%.

Q5: Where is the final application of this predictive model deployed, and how can users access it?

A5: The final application of this predictive model is deployed on Render, and users can access it via the following website link: https://adultcensusincomeprediction.onrender.com