Architecture Design

Fraud Transaction Detection

Document Control

Version	Date	Author	Comments
1	08/2022	Shubham Chavhan	

Index

Abstract

Machine Learning is a category of algorithms that allows software applications to become more accurate in predicting outcomes without being explicitly programmed. The basic premise of machine learning is to build models and employ algorithms that can receive input data and use statistical analysis to predict an output while updating outputs as new data becomes available. These models can be applied in different areas and trained to match the expectations of management so that accurate steps can be taken to achieve the organization's target. In this project, we will estimate the amount of insurance premium on the basis of personal health information. Taking various aspects of a dataset collected from people, and the methodology followed for building a predictive model.

1. Introduction

1.1 What is Architecture Design?

The goal of Architecture Design (AD) is to give the internal design of the actual program code for the Insurance Premium Prediction . AD describes the class diagrams with the methods and relation between classes and program specification. It describes the modules so that the programmer can directly code the program from the document.

1.2 Scope

Architecture Design (AD) is a component-level design process that follows a step-by-step refinement process. This process can be used for designing data structures, required software, architecture, source code, and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work. And the complete workflow.

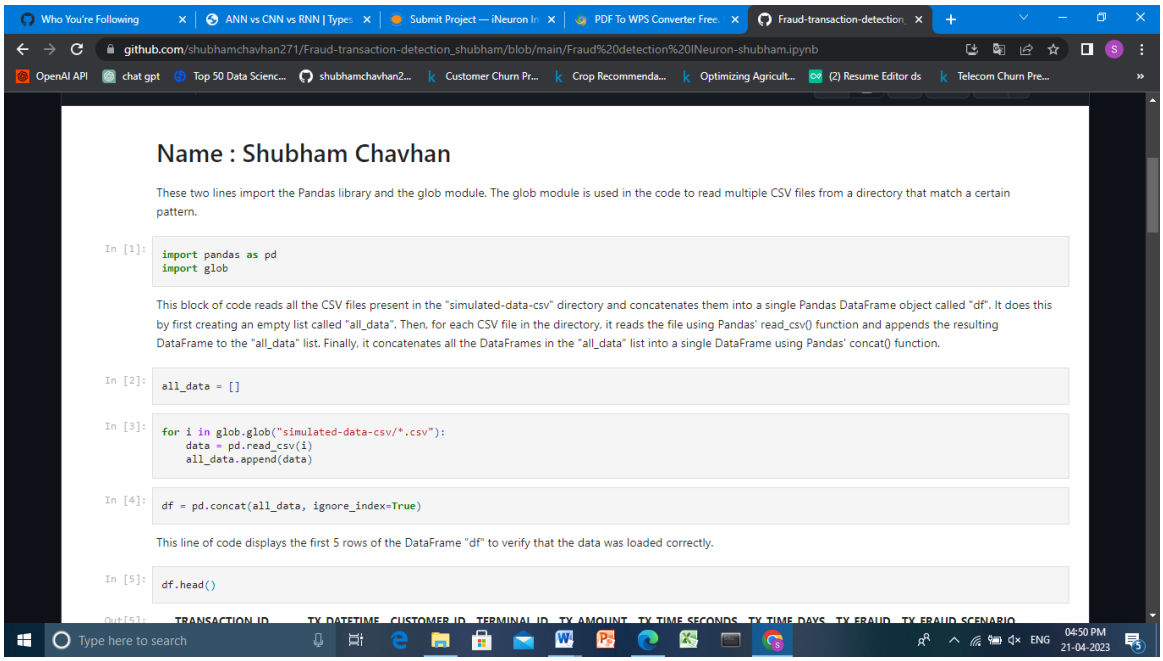
1.3 Constraints

We only predict the expected estimating cost of expenses customers based on some personal health information.

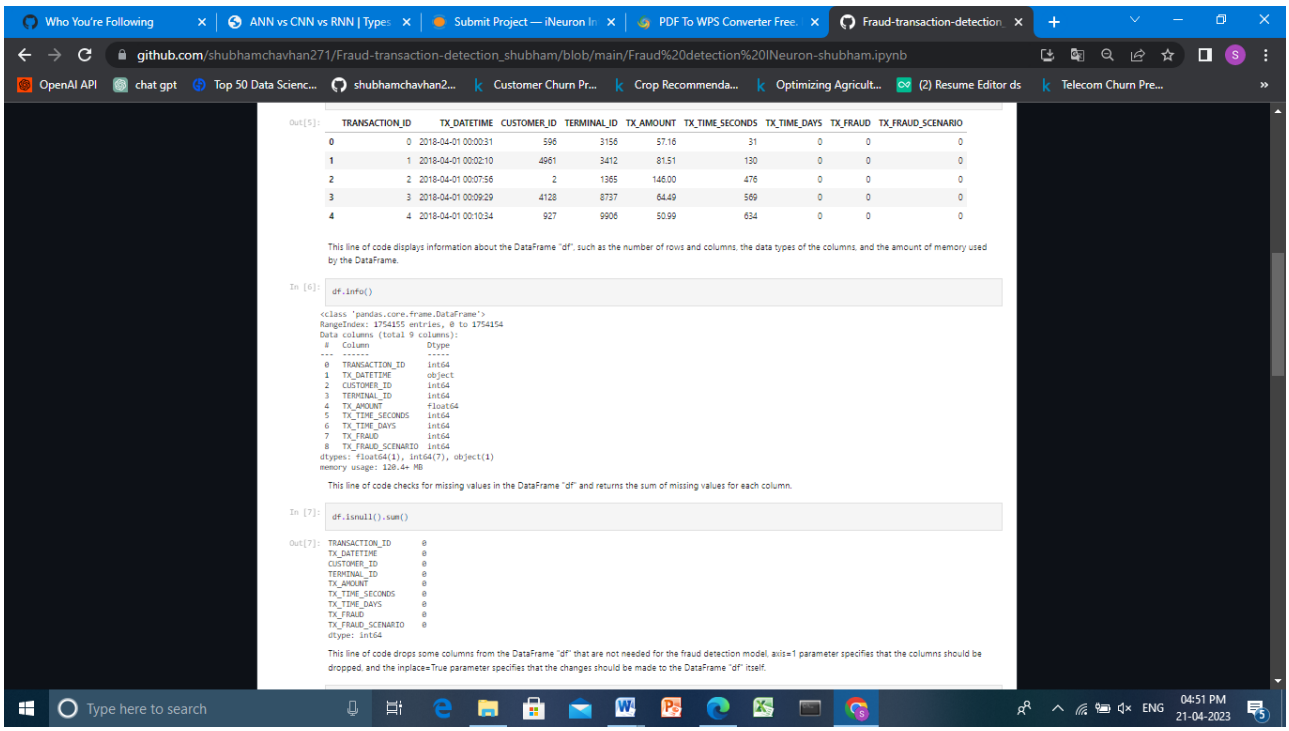
2. Technical Specification

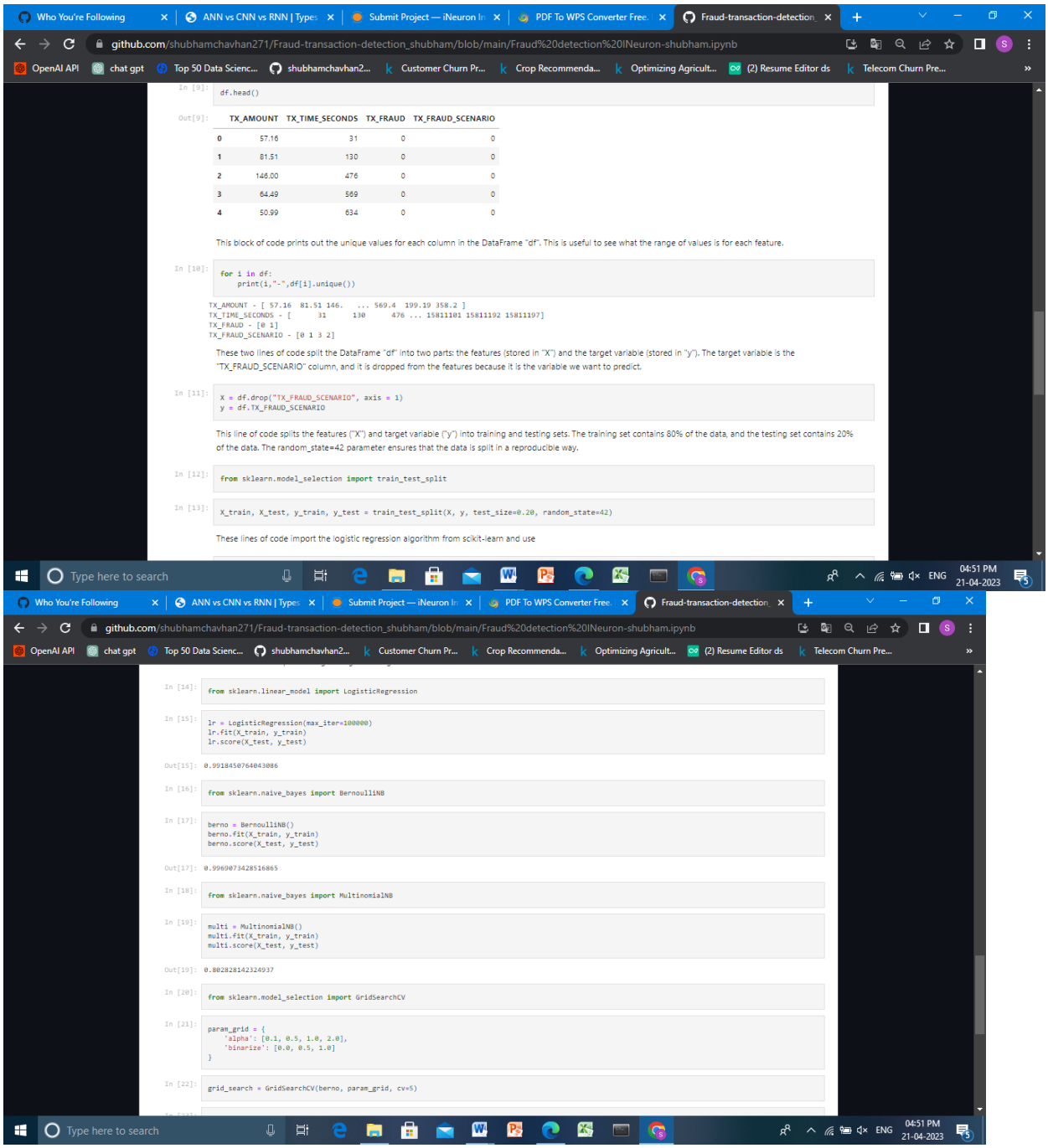
Dataset

The dataset given by INeuron



The data set consists of various data types from integer to floating to object as shown in Fig.

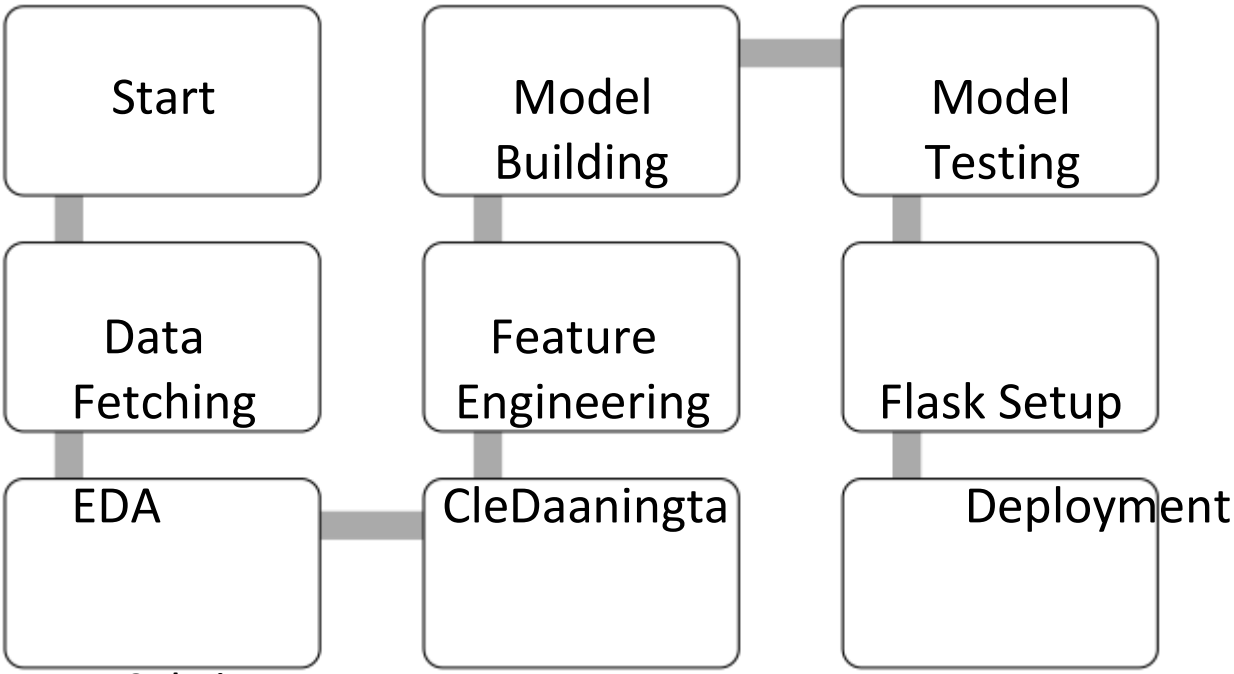




4. **Proposed Solution**

We will use performed EDA to find the important relation between different attributes and will use a machine-learning algorithm to estimate the cost of expenses. The client will be filled the required feature as input and will get results through the web application. The system will get features and it will be passed into the backend where the features will be validated and preprocessed and then it will be passed to a hyperparameter tuned machine learning model to predict the final outcome.

5. **Architecture**



5.1 **Data Gathering**

Data source: https://github.com/shubhamchavhan271/Fraud-transaction-detection_shubham/blob/main/Fraud%20detection%20INeuron-shubham.ipynb

Dataset is stored in .csv format.

5.2 **Raw Data Validation**

After data is loaded, various types of validation are required before we proceed further with any operation. Validations like checking for zero standard deviation for all the columns, checking for complete missing values in any columns, etc. These are required because the attributes which contain these are of no use. It will not play role in contributing to the estimating cost of the premium.

5.3 **Exploratory Data Analysis**

Visualized the relationship between the dependent and independent features. Also checked relationship between independent features to get more insights about the data.

5.4 **Feature Engineering**

After pre-processing standard scalar is performed to scale down all the numeric features. Even one hot encoding is also performed to convert the categorical features into numerical features. For this process, pipeline is created to scale numerical features and encoding the categorical features.

.

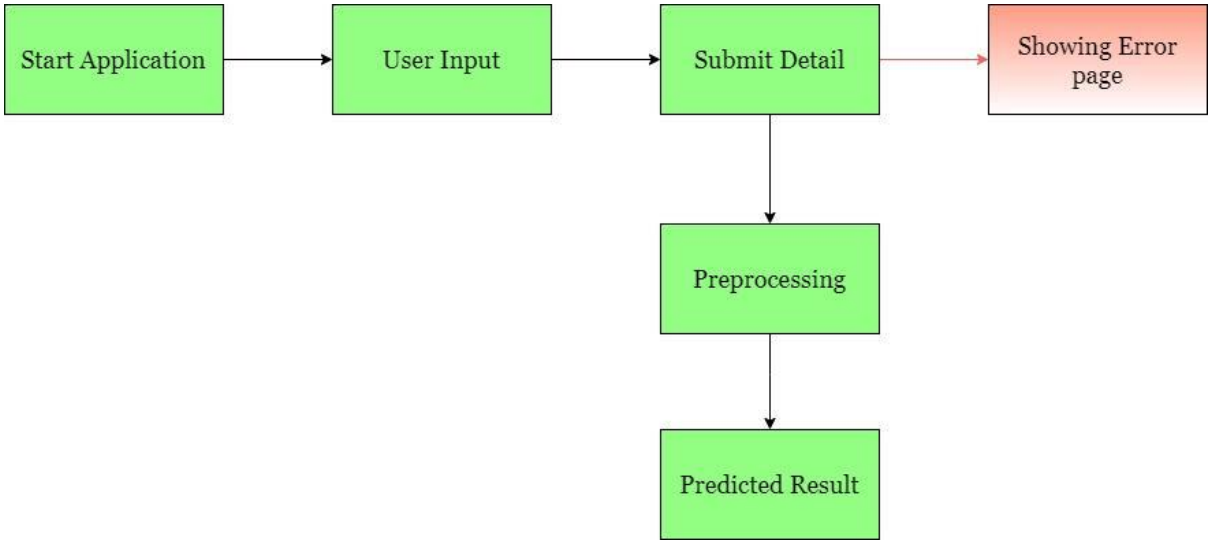
5.8 **GitHub**

The whole project directory will be pushed into the GitHub repository.

5.9 **Deployment**

The project was deployed from GitHub into the Heroku platform.

6. **User Input / Output Workflow.**


Start Application
User Input
Submit Detail
Showing Error page
Preprocessing
Predicted Result