**Real Time Projects**

1. Customer Shopping Dataset - Retail Sales Data

Source: <https://www.kaggle.com/datasets/mehmettahiraslan/customer-shopping-dataset>

Description: The Dataset contains shopping information from 10 different shopping malls between 2021 and 2023. We have gathered data from various age groups and genders to provide a comprehensive view of shopping habits in Istanbul. The dataset includes essential information such as invoice numbers, customer IDs, age, gender, payment methods, product categories, quantity, price, order dates, and shopping mall locations.

1. Medical Insurance Premium Prediction

Source : Machine Learning with R by Brett Lantz

Description : The data includes 1,338 examples of beneficiaries currently enrolled in the insurance plan, with features indicating characteristics ('age', 'sex', 'bmi', 'children', 'smoker', 'region') of the patient as well as the total medical expenses ('charges' ) charged to the plan for the calendar year.

1. Predicting Credit Risk

Source: <https://www.kaggle.com/datasets/kabure/german-credit-data-with-risk>

Description: In this dataset, each entry represents a person who takes a credit by a bank. Each person is classified as good or bad credit risks according to the set of attributes (Checking account, Saving accounts, Age,Sex, Job, Housing, Credit amount, Duration, Purpose, Risk)

1. Restaurant Revenue prediction

Source : <https://www.kaggle.com/c/restaurant-revenue-prediction>

Description: TFI has provided a dataset with 137 restaurants in the training set, and a test set of 100000 restaurants. The data columns include the open date, location, city type, and three categories of obfuscated data: Demographic data, Real estate data, and Commercial data. The revenue column indicates a (transformed) revenue of the restaurant in a given year and is the target of predictive analysis.

1. House price Prediction

Description: Predict the house price based on Dist\_Taxi, Dist\_Market, Dist\_Hospital, Carpet Builtup, Parking, City\_Category, Rainfall.

1. Customer Segmentation and Market Basket Analysis

Source: <https://www.kaggle.com/code/mgmarques/customer-segmentation-and-market-basket-analysis>

Description: On-line Retail data to explore customer segmentation and apply association rule mining approach to find interesting rules and patterns in this transaction database. Based on InvoiceNo, StockCode, Item Description, Quantity, InvoiceDate, UnitPrice, CustomerID, Country.

1. Crop Yield Prediction

Source: <https://www.kaggle.com/datasets/patelris/crop-yield-prediction-dataset>

Description: The objective of crop yield prediction using machine learning is to develop models

that can accurately predict the yield of a crop given various environmental and farming factors.

1. Gold-Price-Prediction

Source: https://www.kaggle.com/datasets/sid321axn/gold-price-prediction-dataset

Description : The historical data of Gold ETF fetched from Yahoo finance has 7 columns, Date, Open, High, Low, Close, Adjusted Close and Volume, the difference between Adjusted Close and Close is that closing price of a stock is the price of that stock at the close of the trading day. Whereas the adjusted closing price takes into account factors such as dividends, stock splits and new stock offerings to determine a value. We would use Adjusted Close as our outcome variables which is the value we want to predict.