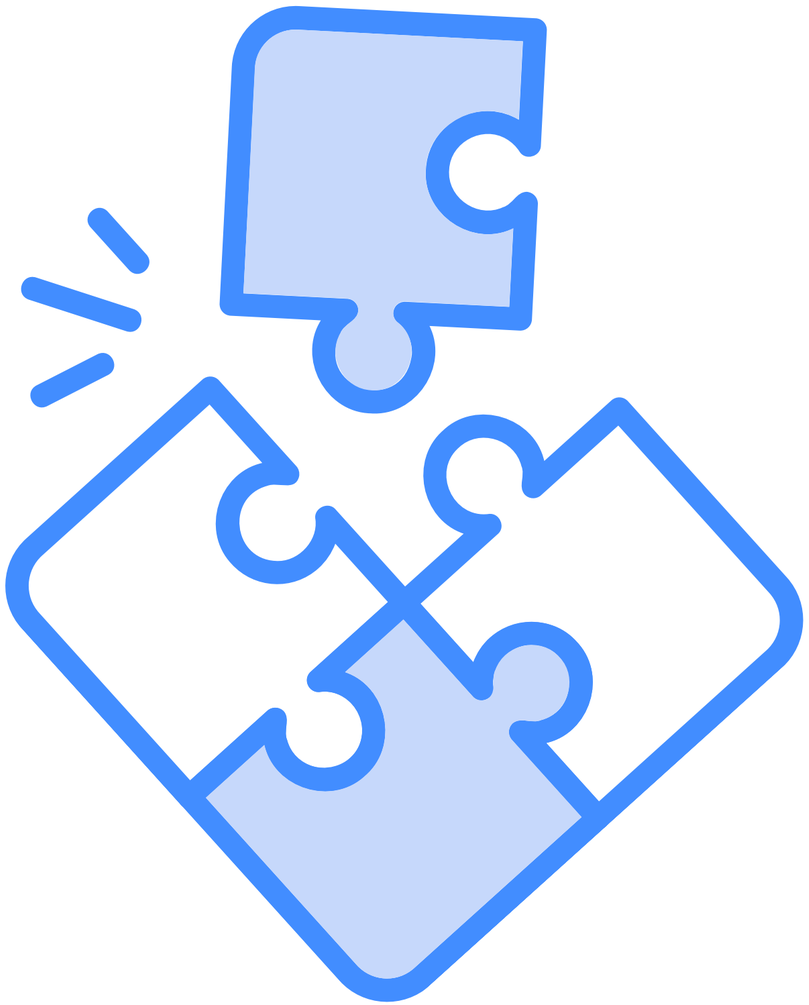
Buckman
Chemistry, connected.

Investment Decision Recommendation System Project

Jananee M - 21PW07



UNDERSTANDING

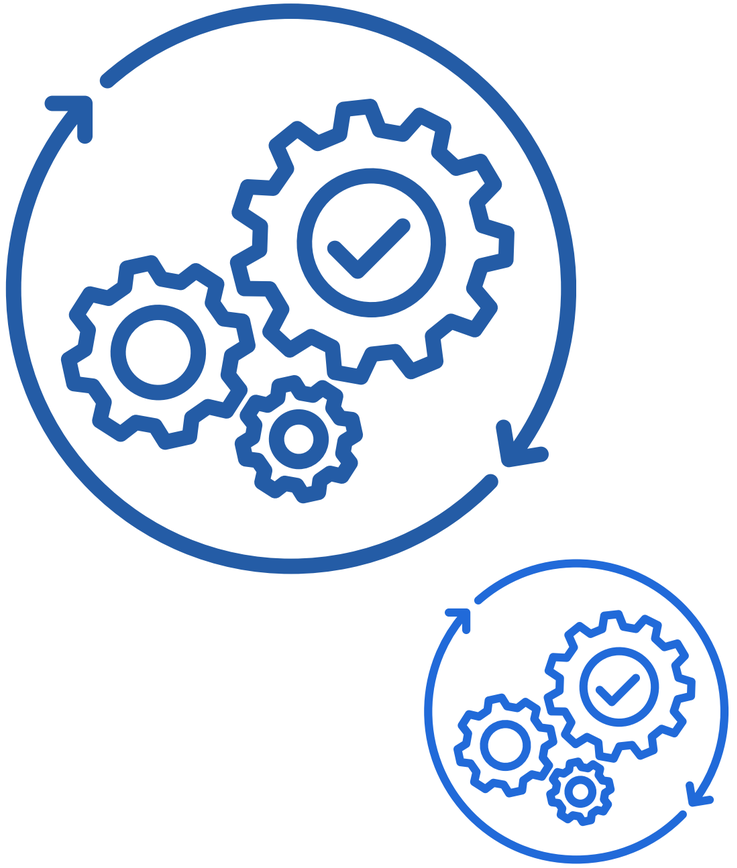
My understanding of the problem statement

My aim is to build a recommendation system that can predict the investment risk level for new data based on important factors, that will let user to make best decisions.

Language	Editor
Python 3.12	VS Code

PROCESS

My Process



01

Data Preprocessing

Data preprocessing involves encoding categorical variables and removing unwanted fields.

02

Factor Selection Process

Next, to identify factors that contribute to making the best investment decision.

03

Train the model

Random Forest Classifier has been trained on the data using the selected features.

04

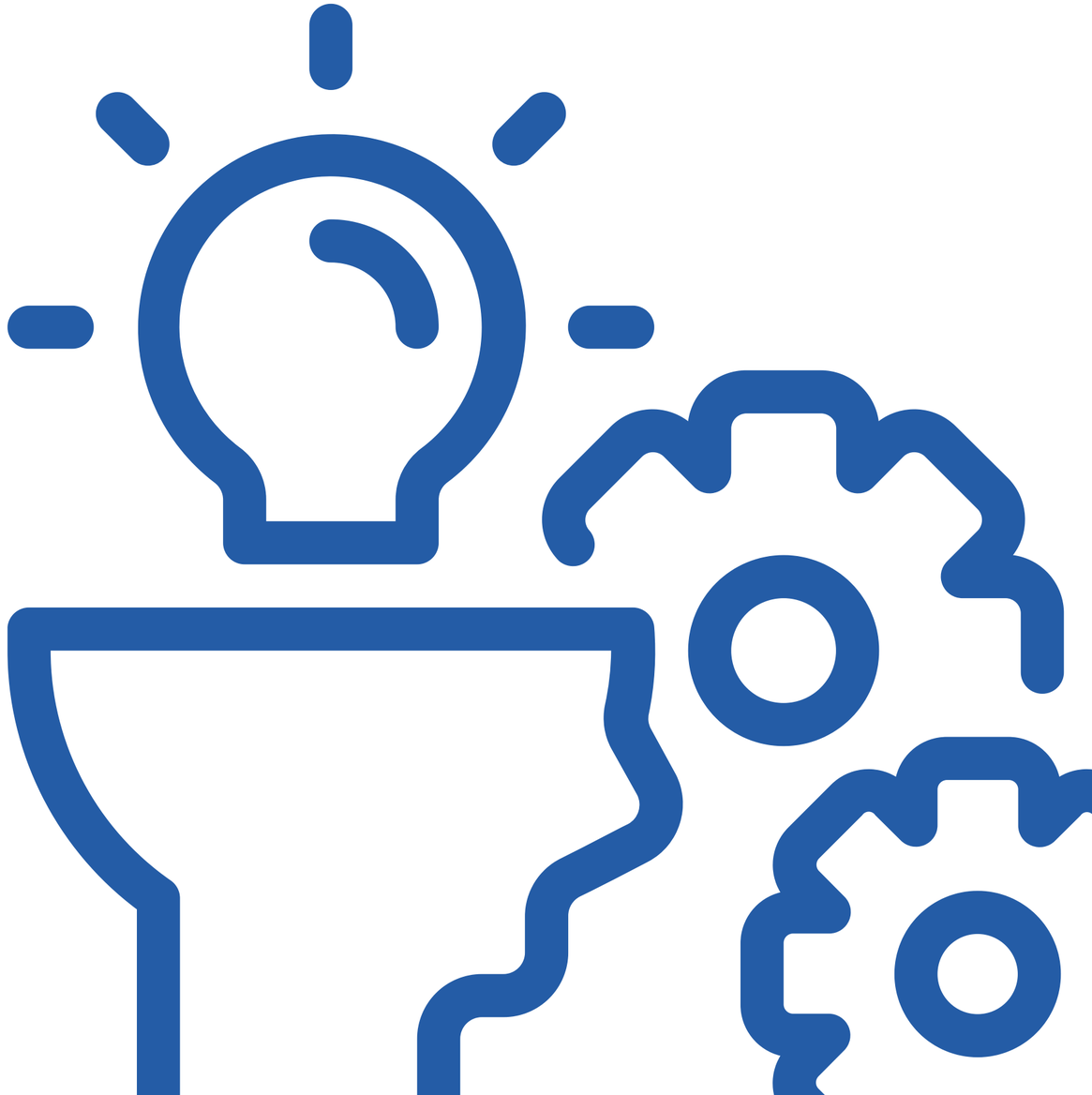
Testing

Through a simple user interface, users will input the values for the important features and risk factor is predicted.

MODEL SELECTION

Why Random Forest?

- Random Forest is applied due to its robustness against overfitting, ability to handle large datasets with high dimensionality, and capability to provide feature importance.
- Random Forest is also preferred for its capability to handle both numerical and categorical data without requiring extensive preprocessing.



SOLUTION

My Solution

- The developed recommendation system accurately predicts investment risks based on individual attributes, aiding in informed decision-making.
- The trained model is capable of predicting investment risks accurately, aiding users in decision-making.
- The Flask web application allows users to input their characteristics and receive personalized risk predictions, displayed along with a correlation matrix visualization.

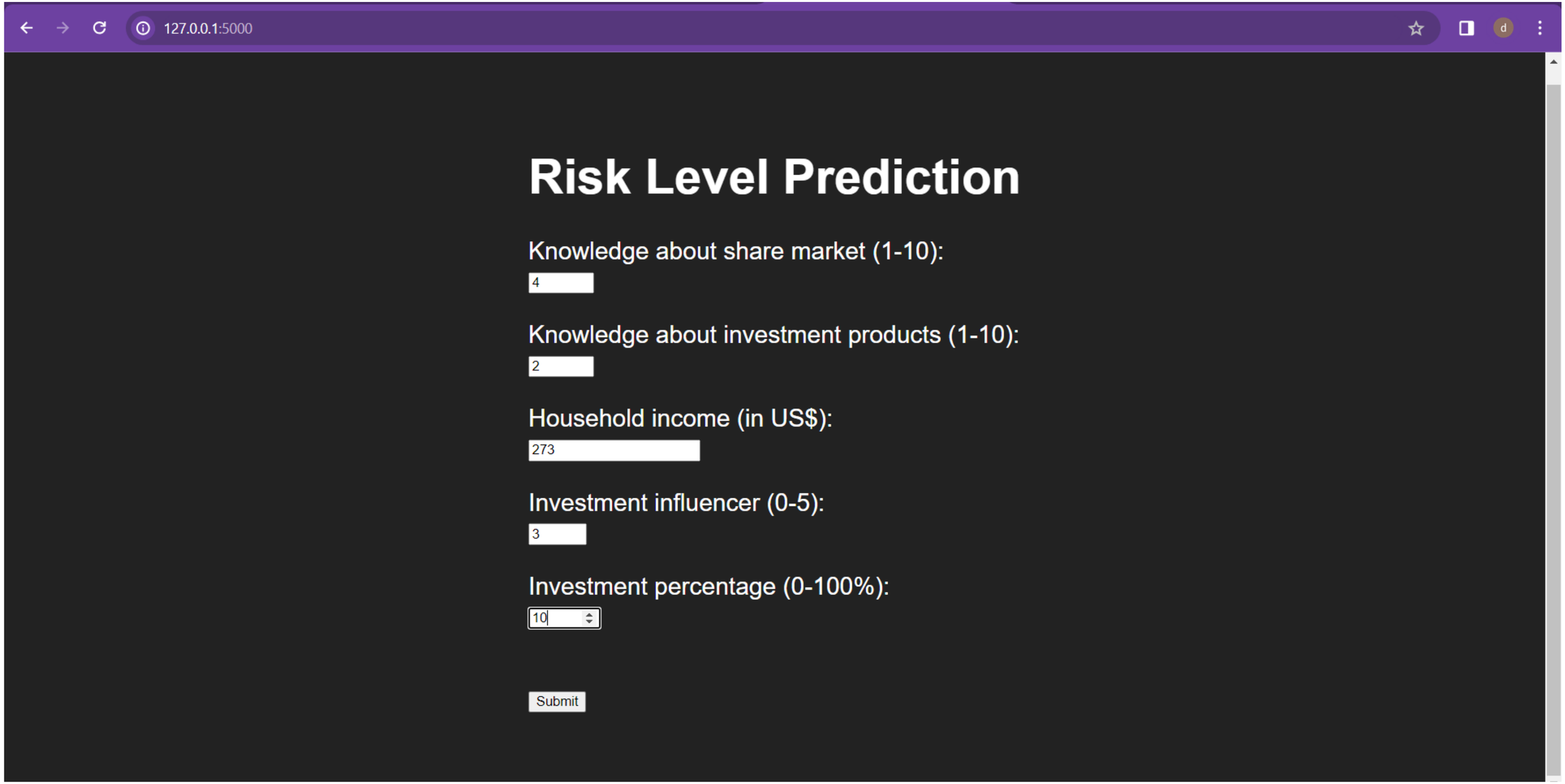
ACCURACY

```
Accuracy: 0.3148148148148148
Feature Importances:
Knowledge level about sharemarket                    0.080007
Knowledge level about different investment product   0.077997
Household Income                                     0.071166
Investment Influencer                                0.069223
Percentage of Investment                             0.068094
Source of Awareness about Investment                 0.067008
Reason for Investment                                0.062476
Return Earned                                        0.062208
Knowledge about Govt. Schemes                        0.061407
Role                                                 0.059231
Education                                            0.058038
Age                                                  0.056456
Number of investors in family                        0.055832
City                                                 0.052687
Investment Experience                                0.052316
Marital Status                                       0.022977
Gender                                               0.022878
dtype: float64
```

FEATURE SELECTION

```
BEST PARAMS...
best params: {'max_depth': None, 'min_samples_split': 2, 'n_estimators': 50}
 x:  Index(['Knowledge level about sharemarket',
       'Knowledge level about different investment product',
       'Household Income', 'Investment Influencer',
       'Percentage of Investment'],
      dtype='object')
Model Trained and Ready for Testing...
```

Home Page


127.0.0.1:5000
Risk Level Prediction
Knowledge about share market (1-10):
4
Knowledge about investment products (1-10):
2
Household income (in US$):
273
Investment influencer (0-5):
3
Investment percentage (0-100%):
10
Submit

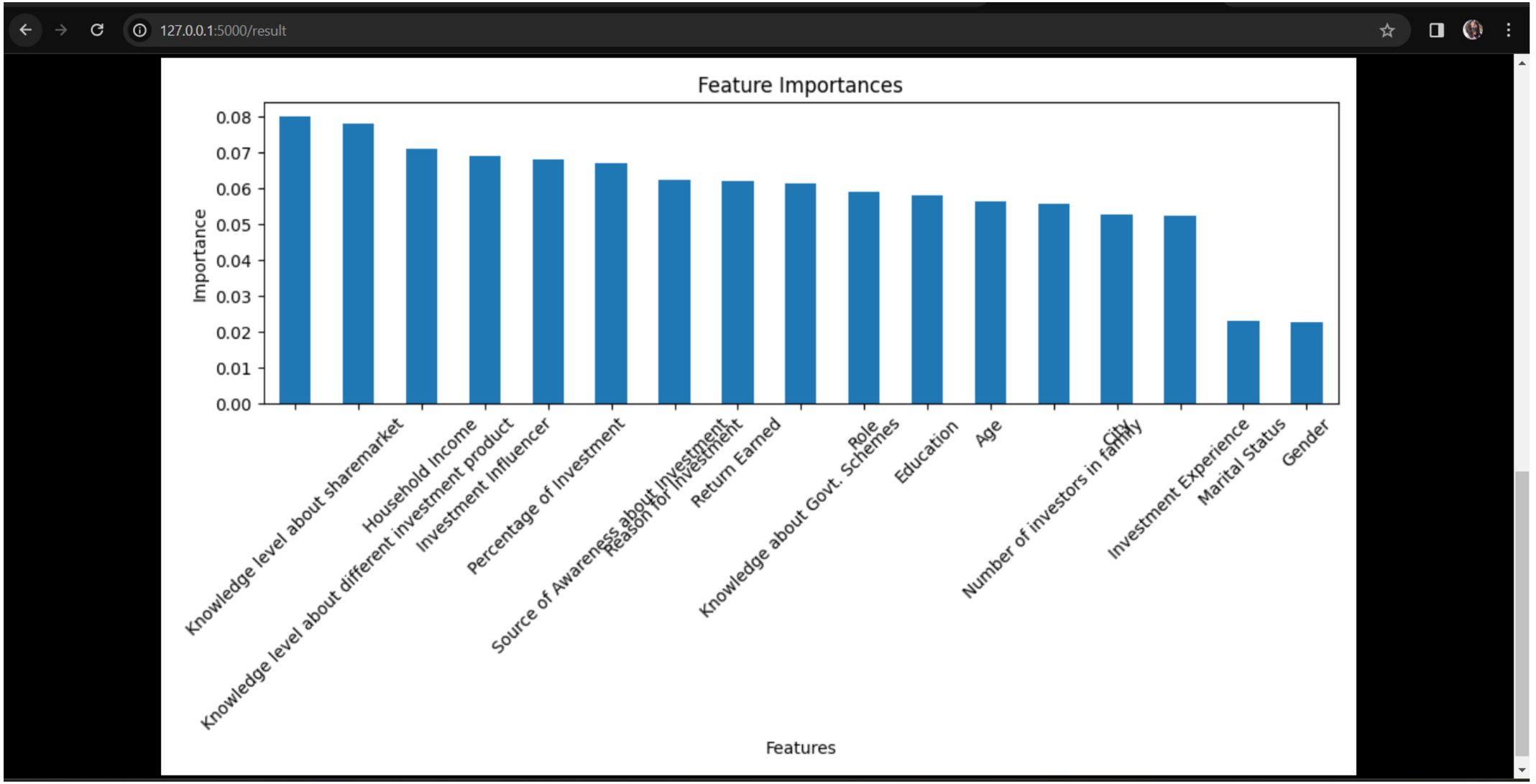
Result Page

Risk Level Prediction Result

Predicted Risk Level: MEDIUM RISK INVOLVED

Correlation Matrix of Selected Factors

Result Page


127.0.0.1:5000/result
Feature Importances
Importance
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
Knowledge level about sharemarket
Household Income
Knowledge level about different investment product
Investment Influencer
Percentage of Investment
Source of Awareness about Investment
Reason for Investment
Return Earned
Role
Knowledge about Govt. Schemes
Education
Age
City
Number of investors in family
Investment Experience
Marital Status
Gender
Features

DESIGN WORK

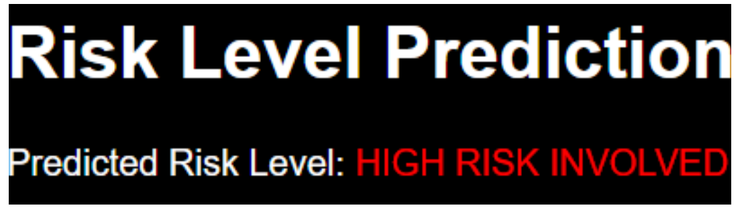
LOW RISK

Green indicates this person has a low risk investment.

MEDIUM RISK

Orange indicates this person has medium risk investment.

HIGH RISK



Red indicates this person has a high risk investment.

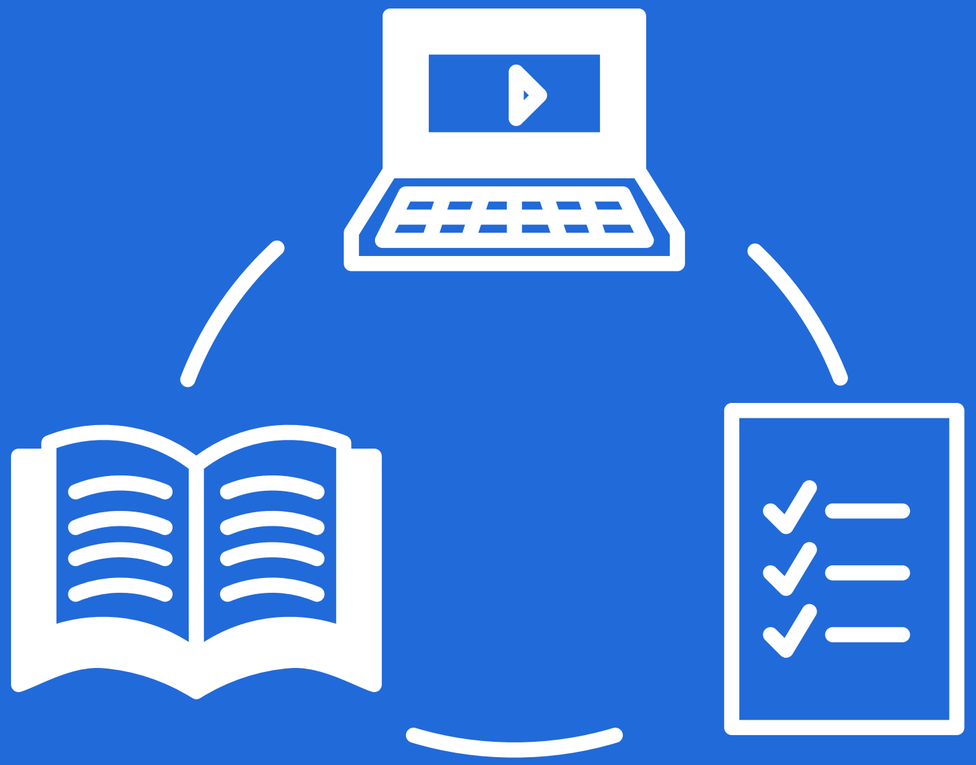
FUTURE

Further Development

Future developments may focus on enhancing model performance and expanding the application's features for broader financial analysis.

LEARNING

Conclusion



Technically...

By leveraging Random Forest, this project demonstrates an effective approach to building a recommendation system for investment decisions, showcasing the potential for data-driven solutions in finance to optimize decision-making processes and mitigate risks.

Non Technically...

I had an amazing experience working on this hackathon. The problem statement was interesting and more challenging. It helped me realize my potential of solving a hackathon in a short period.

THANK YOU