



Intel Unnati Industrial Training Program 2024

Problem Statement-12

By InnovateHers
(Deepsikha Das
Ankita Raj
Diya Manna)

Knowledge Representation and Insights Generation from Structured Datasets



Unique Idea Brief (Solution)

The purpose of Data Insight Generator is a web app that aims to simplify the whole process of data analysis and insight generation at users' end. Through allowing users to upload their datasets and question their data, it offers a summary of statistics, visualization as well as machine-learning predictions. This provides users an opportunity they otherwise wouldn't have had in undertaking analytics since they do not possess advanced skills required but henceforth they can still make informed decisions based on available information it presents.

Features Offered



Dataset Upload:

Users can upload datasets in CSV format.



Data Cleaning:

Automatic handling of missing values and encoding of categorical variables.



Statistical Summaries:

Provides mean, median, standard deviation, and other statistical metrics.

Features Offered



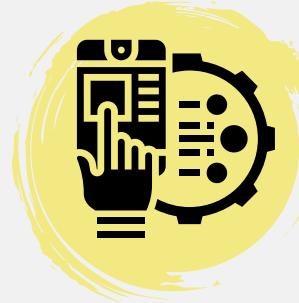
Data Visualizations:

Generates various types of plots, including scatter plots, line graphs, and histograms



Machine Learning Models:

Supports clustering (K-Means), classification (Random Forest), and regression (Random Forest).



User-Friendly Interface

Easy-to-use web interface for data upload and question submission.



Interactive Responses:

Returns insights in the form of text summaries and visual plots.

Process Flow

01. **User Uploads Dataset:**

The user navigates to the home page and uploads a dataset in CSV format.

02. **Data Preprocessing:**

- The uploaded dataset is loaded, and missing values are handled.
- Categorical variables are encoded, and numerical variables are scaled.

03. **User Asks Questions:**

The user submits questions about the dataset through a form.

Process Flow

04.

Processing Questions:

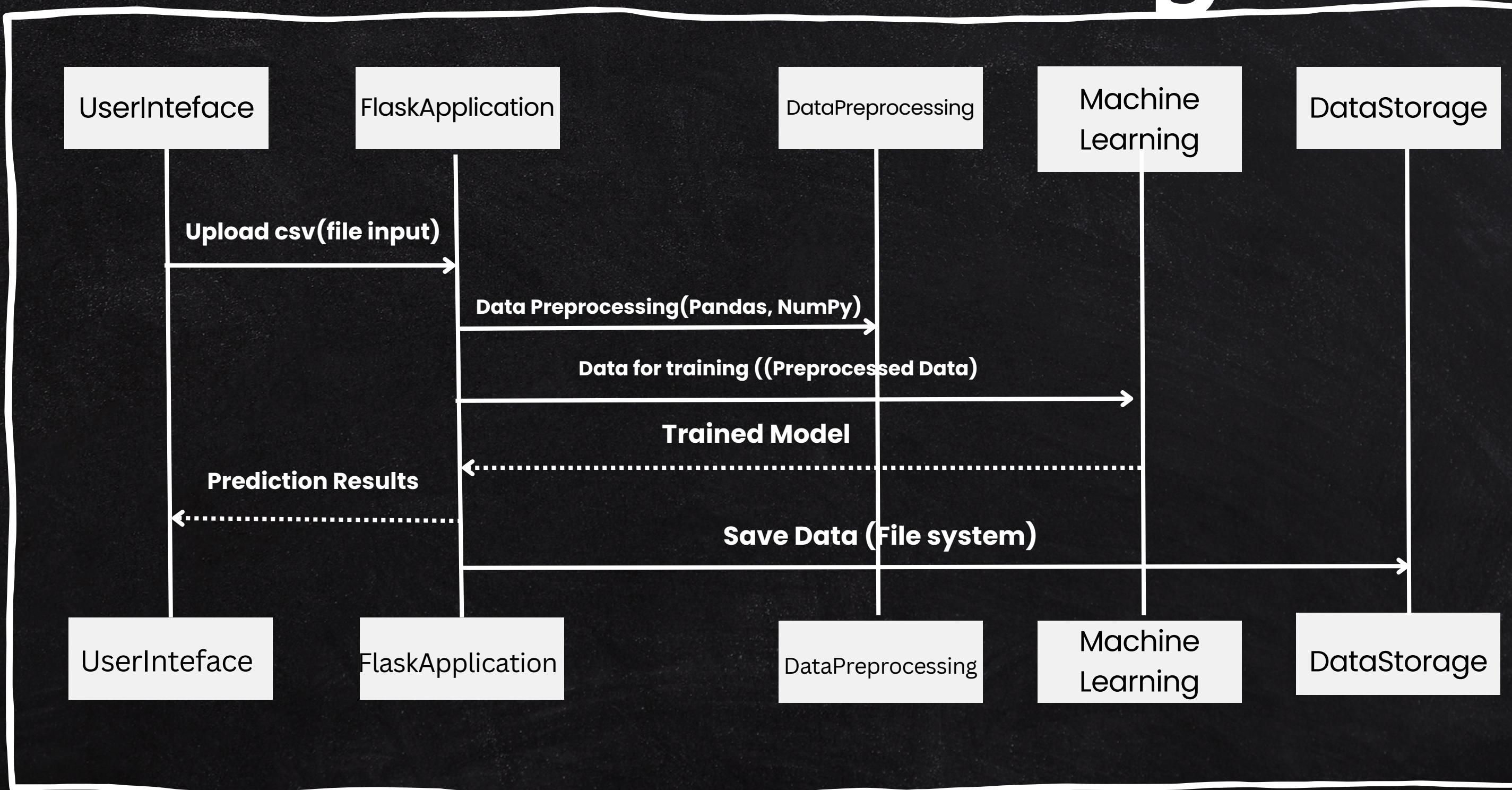
The application processes the questions, generating statistical summaries, visualizations, or applying machine learning models as needed.

Response Generation:

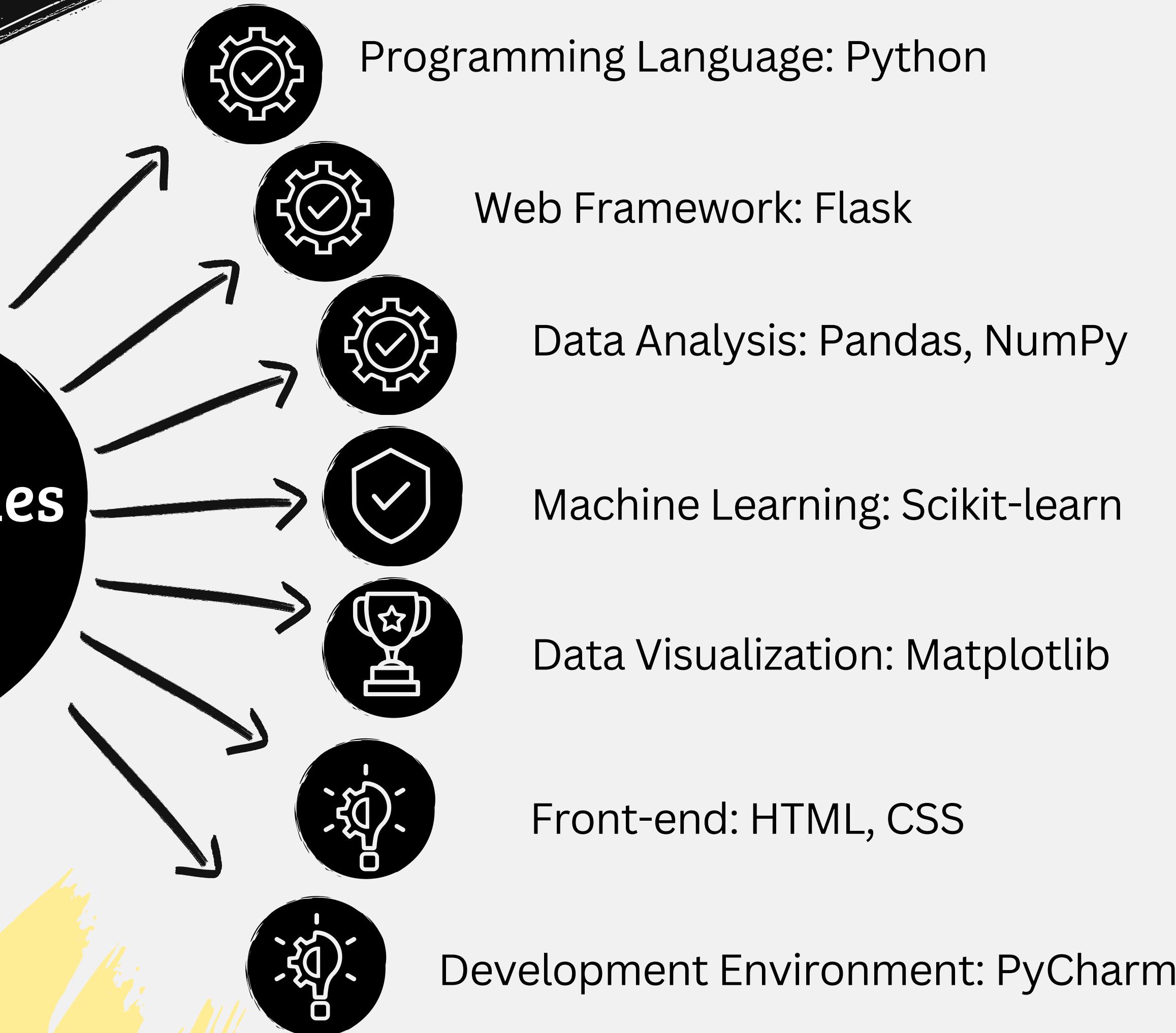
05.

The application generates the appropriate response (text or plot) and displays it to the user.

Architecture Diagram



Technologies Used



Team

Deepsikha Das



Team Lead

The Team Lead oversaw the project, coordinated tasks, and ensured timely delivery. Supervised data preprocessing, and machine learning with KMeans, RandomForestClassifier, and RandomForestRegressor, also facilitated communication between frontend and backend teams, ensuring seamless integration and efficient workflow throughout the project.

Ankita Raj



Backend Manager

designed a scalable and secure architecture, integrated machine learning models, and managed data preprocessing. Ensured performance optimization and additionally, they managed the database, tested and debugged issues, and collaborated with frontend developers for seamless integration and comprehensive documentation.

Diya Manna



Frontend Manager

designed the user interface, implemented responsive layouts, and ensured a seamless user experience. Integrated forms for data upload and question input, displayed results dynamically, and coordinated with the backend team, also ensured cross-browser compatibility and optimized performance for a smooth, interactive experience.

Conclusion

The Data Insight Generator is a comprehensive solution designed to make data analysis accessible to a broader audience. By providing a platform for users to upload datasets, ask questions, and receive insightful responses, the application empowers users to leverage data for decision-making. With a user-friendly interface and robust backend capabilities, the Data Insight Generator stands as a valuable tool for individuals and businesses alike. Future enhancements could further extend its capabilities, making it even more versatile and powerful.