**GROUP – 15**

Final Report

SUBMITTED BY – ADITYA SHIVAJI PATIL (1001995431).

RISHABH JAIN (1001967271).

AKSHAY SADAGELKAR (1001944713).

SUBMITTED TO – DR. HABEEB OLUFOWOBI.

Logo, company name

Description automatically generated

Project Overview –

From a resale point of view, this service calculates how much your car is worth. It will be provided via a web-based platform which will be hosted in the cloud and can be

accessed by anyone. To compute the value of a car, we will use machine learning algorithms that will utilize certain factors such as Manufacturing year, Make, Model, Variant Transmission, Fuel Type, Body Type, Condition. In addition, web scraping will be used to find similar cars posted on other sites like Carvana and Carmax as well for comparative analysis of their prices.

Technology Used

**Web Components**

* Flask
* Html
* CSS

**ML Components**

* scikit-learn
* NumPy
* Pandas

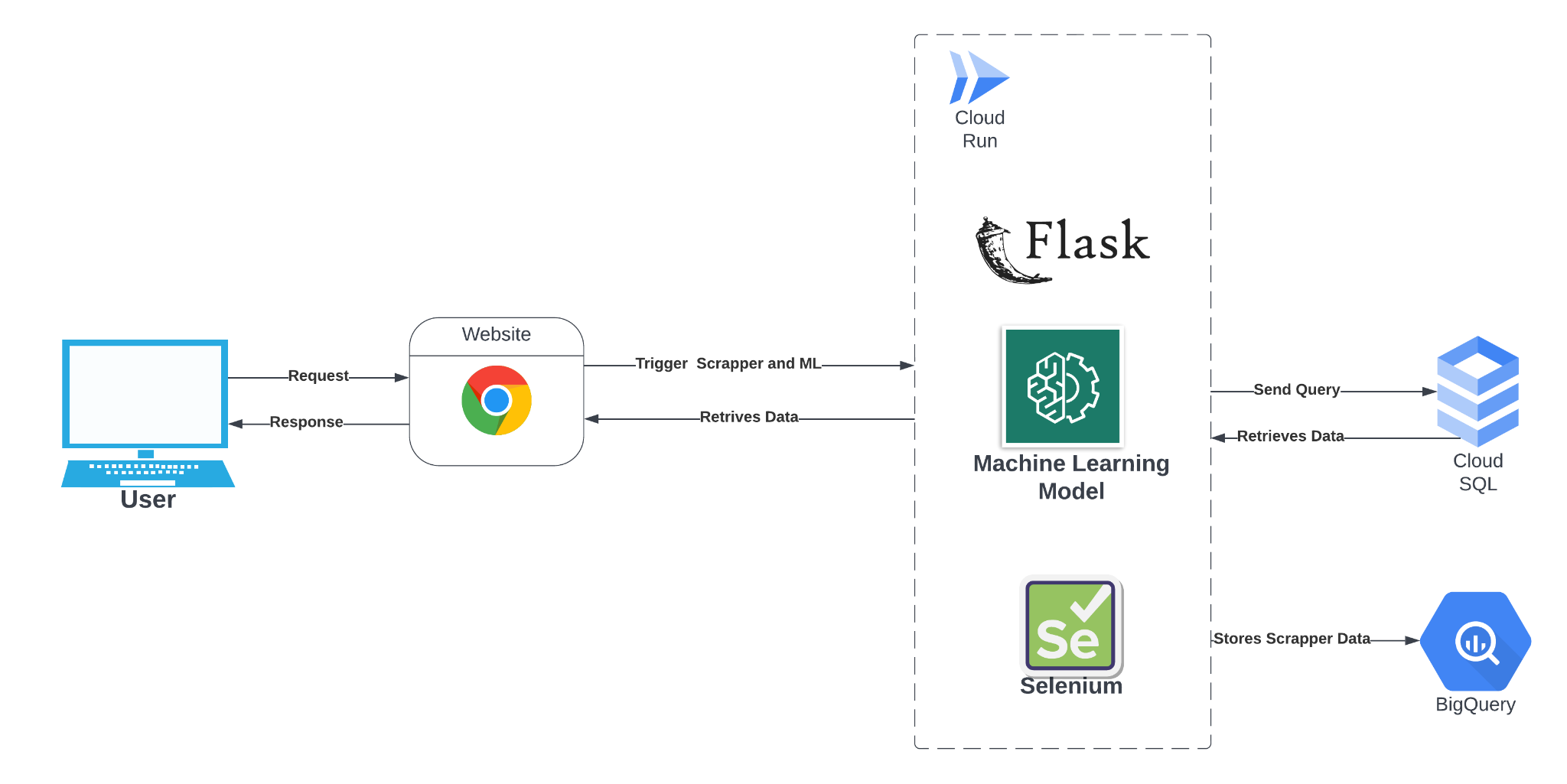
**Scrapper Components**

* Selenium
* Beautiful Soap
* Python
* Regex

**Google Cloud Components**

* Cloud SQL Storage
* Compute Run
* Big Query
* Containers
* Docker
* Google cloud container registry

Architecture

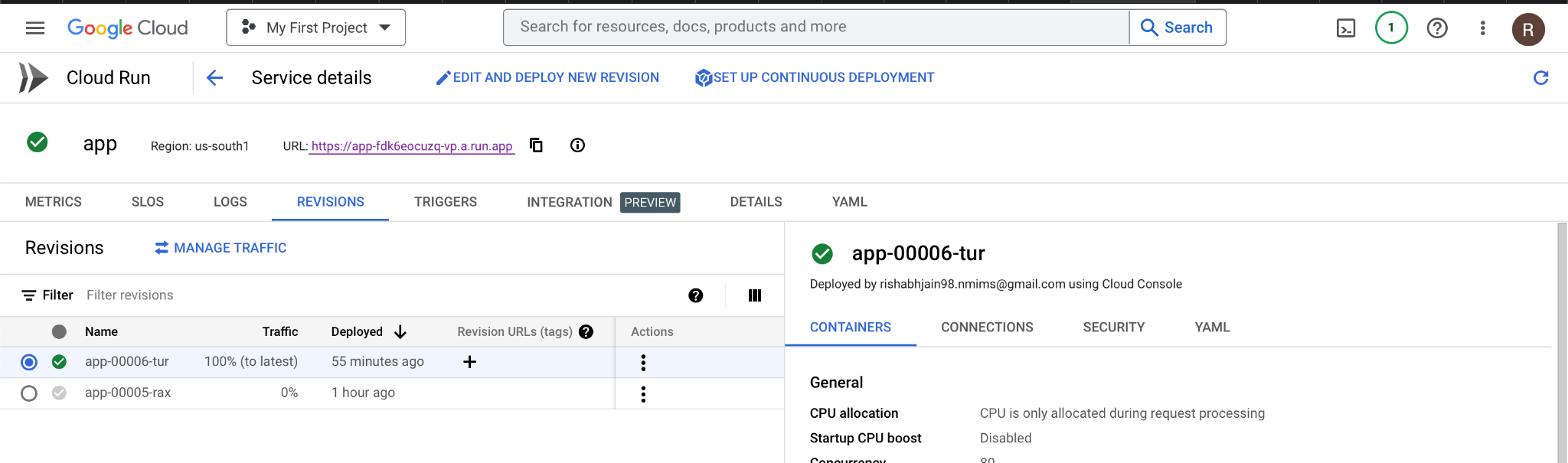


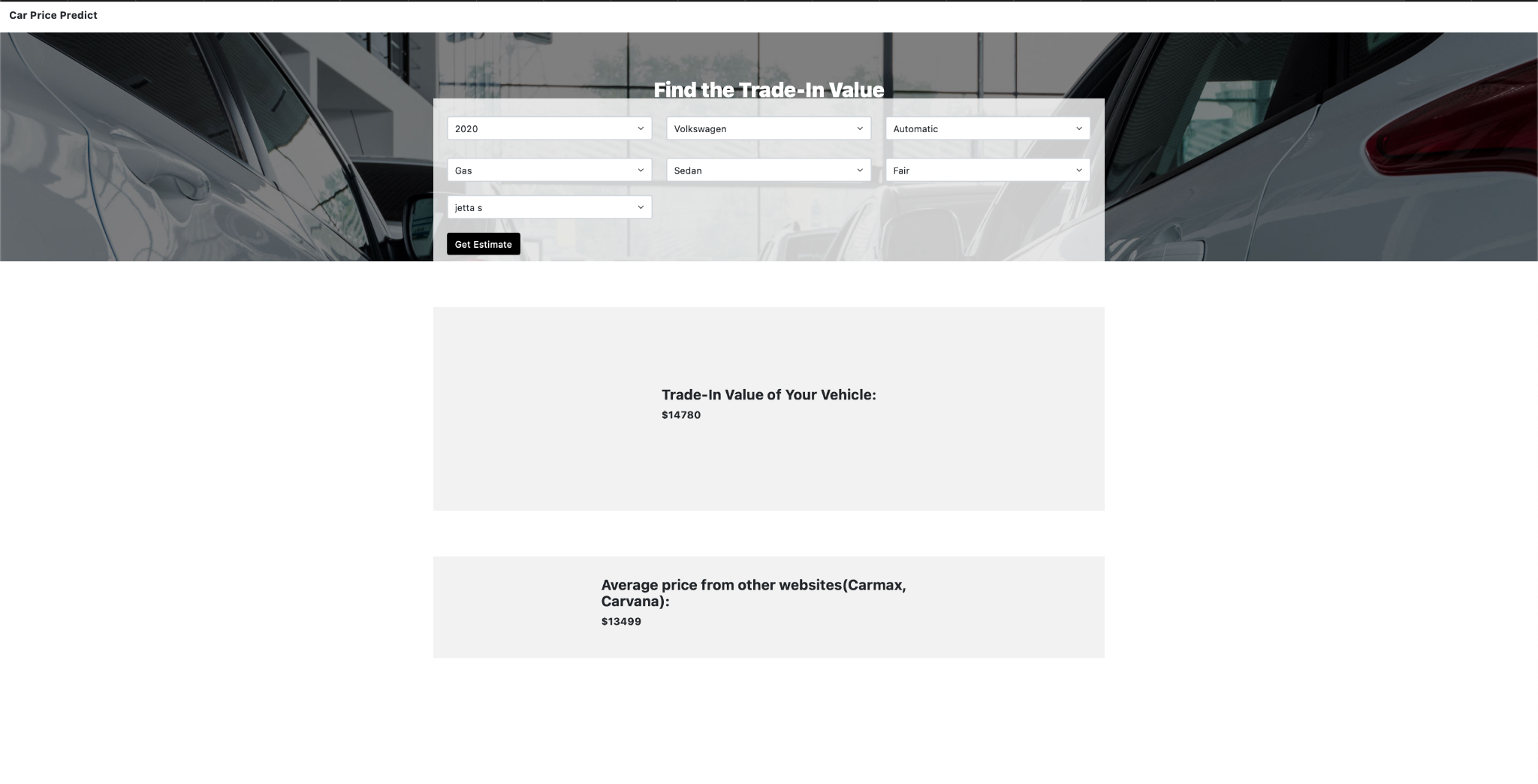
Application Working

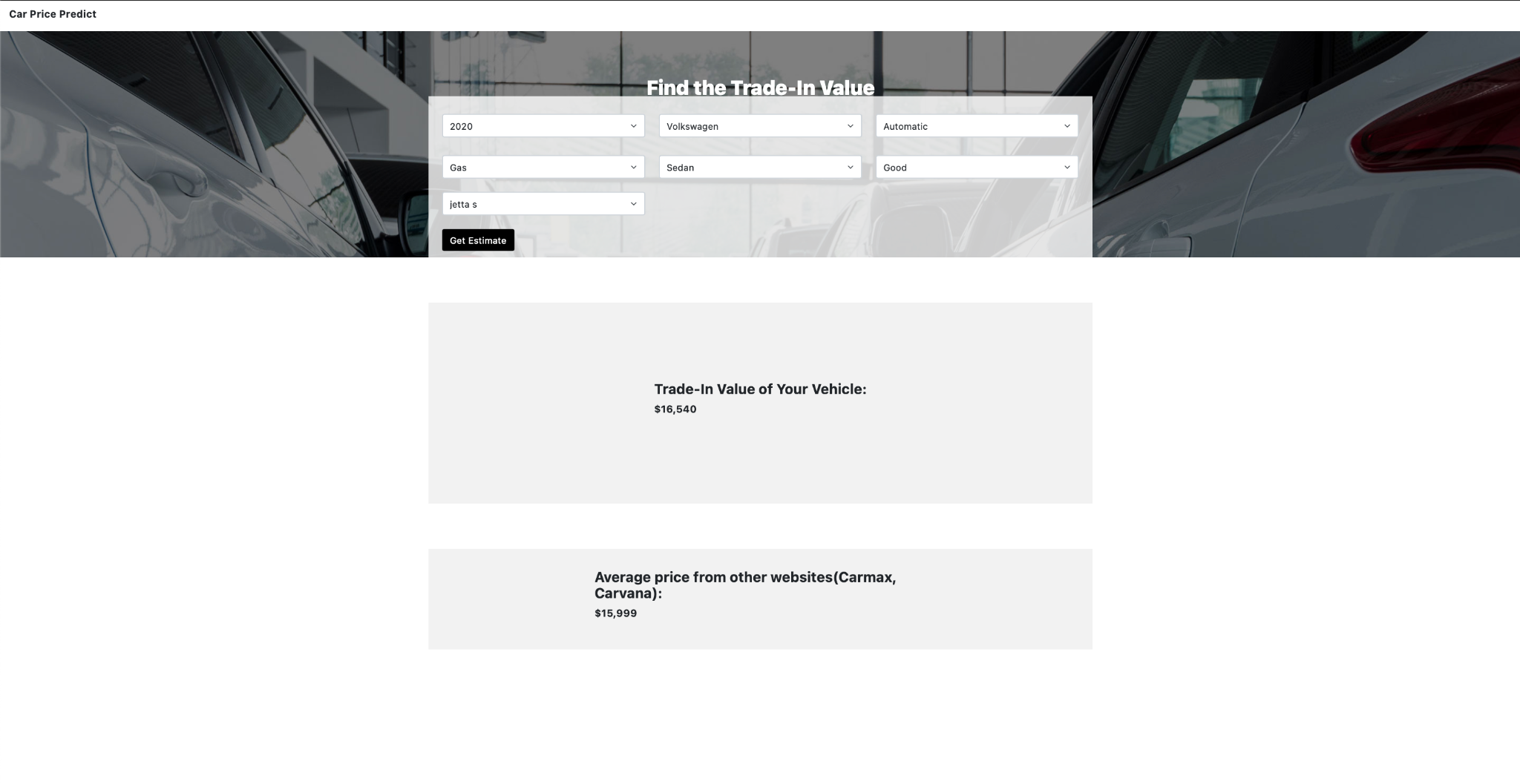
* Selected the craigslist dataset to run the ML model on approx. 450k records Table

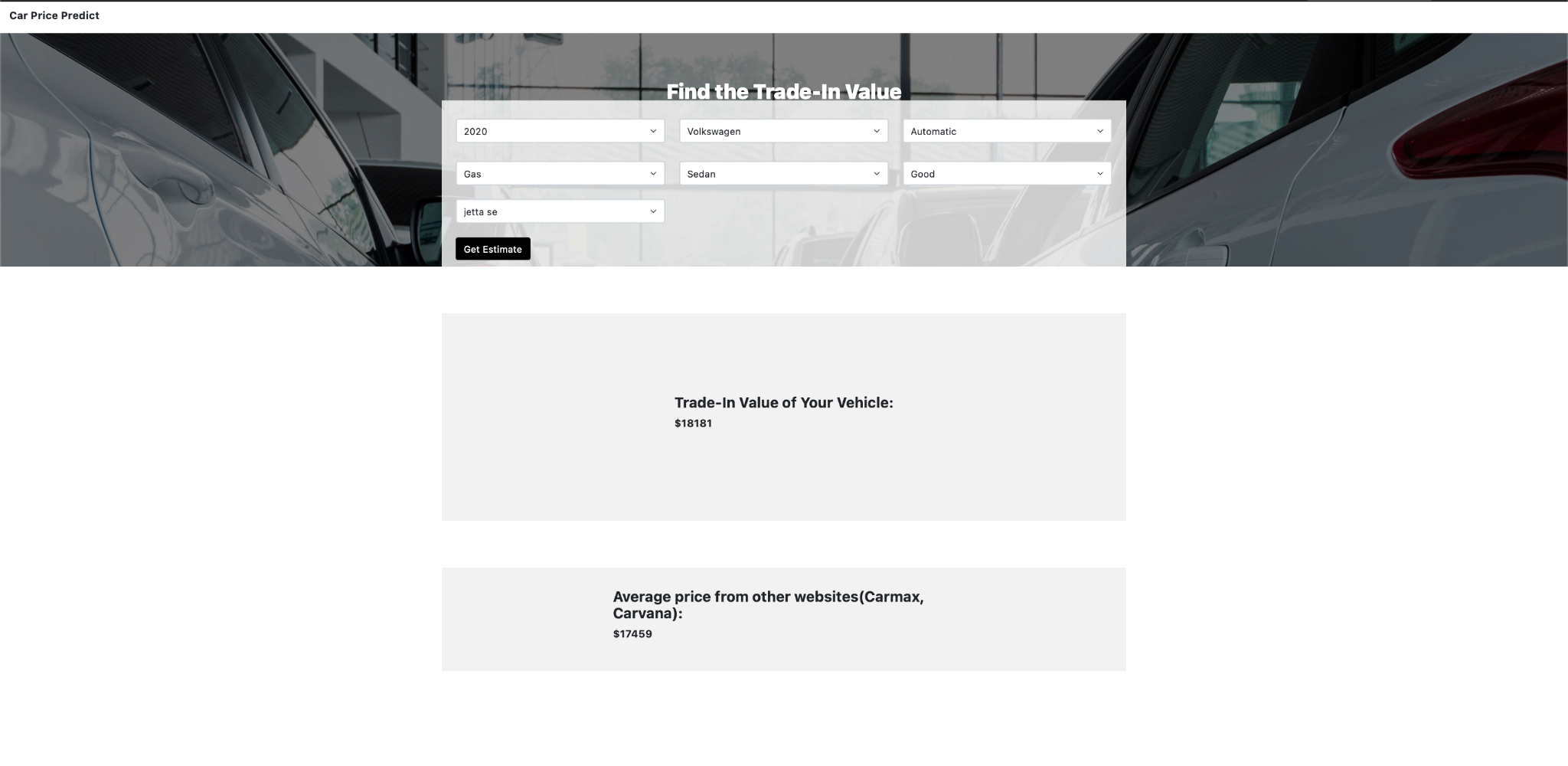
  Description automatically generated
* After preprocessing , we concluded to use random regression to estimate car price for the ML model.
* Machine Learning Algorithm: for calculating prices using scikit-learn, data preprocessing using python NumPy and pandas.
* Initial run test provided with an accuracy of approx. 80% on training data.Text

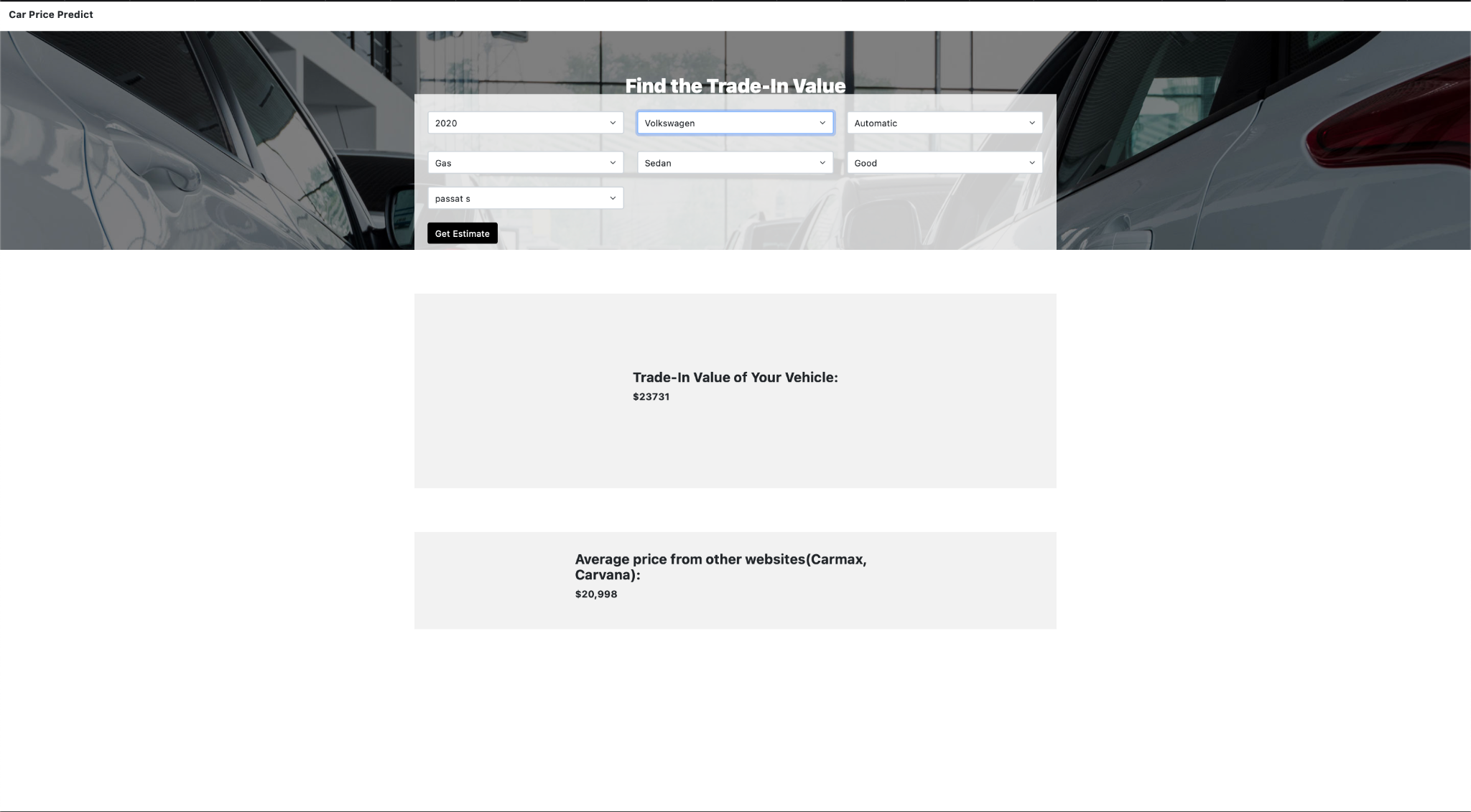
  Description automatically generated with low confidence
* Value evaluator: A Python program that uses a machine learning model which will be called every time the user requests the car price.
* For aggregated results, third-party websites for comparative analysis such as Carvana.com, CarMax.com.
* Web Scraper used Selenium, Beautiful Soap and regex to fetch price using parameters Year, Make, Model,Variant Transmission, Fuel Type, Body Type, Condition.
* Website is launched for user interface utilizing the HTML, and CSS Template which is rendered using the Flask application.
* Once the user requests to fetch information, the query is searched in the cloud sql database to check if the price is previously fetched and return results with a faster response time rather than going through the whole process of running ml model and scrapper.
* If the data is not available in cloud sql, the scrapper and ml model is run and the result is generated.
* Every time the scrapper is run, the result is stored in the BigQuery database for data collection. Reason being the data will be used to recalibrate the ML model to increase its accuracy.











Insights

In response to changes in the parameters, prices change according to market trends. There is a change in condition of the car in the first two screenshots, resulting in a difference of about $2000. The third screenshot with the same car condition but different variant further increases the price. In the last screenshot, a different car model is used, which showcases a new price accordingly.

Future Work

* Using Data in BigQuery to recalibrate the ML model to increase accuracy.
* Integrating more websites for scrappers to fetch data.
* Improve UI for better visual representation and display statistics.

Conclusion

We have implemented a Car price estimator for all individuals who want to get an estimate of their car’s price and at the same time compare it with the leading car resellers all at a single place using the parameters. The program ingests seven input values about the car, including specifications and conditions, and generates a result. In addition to predicting the price, the website shows an average selling amount on various websites using web scrapers.

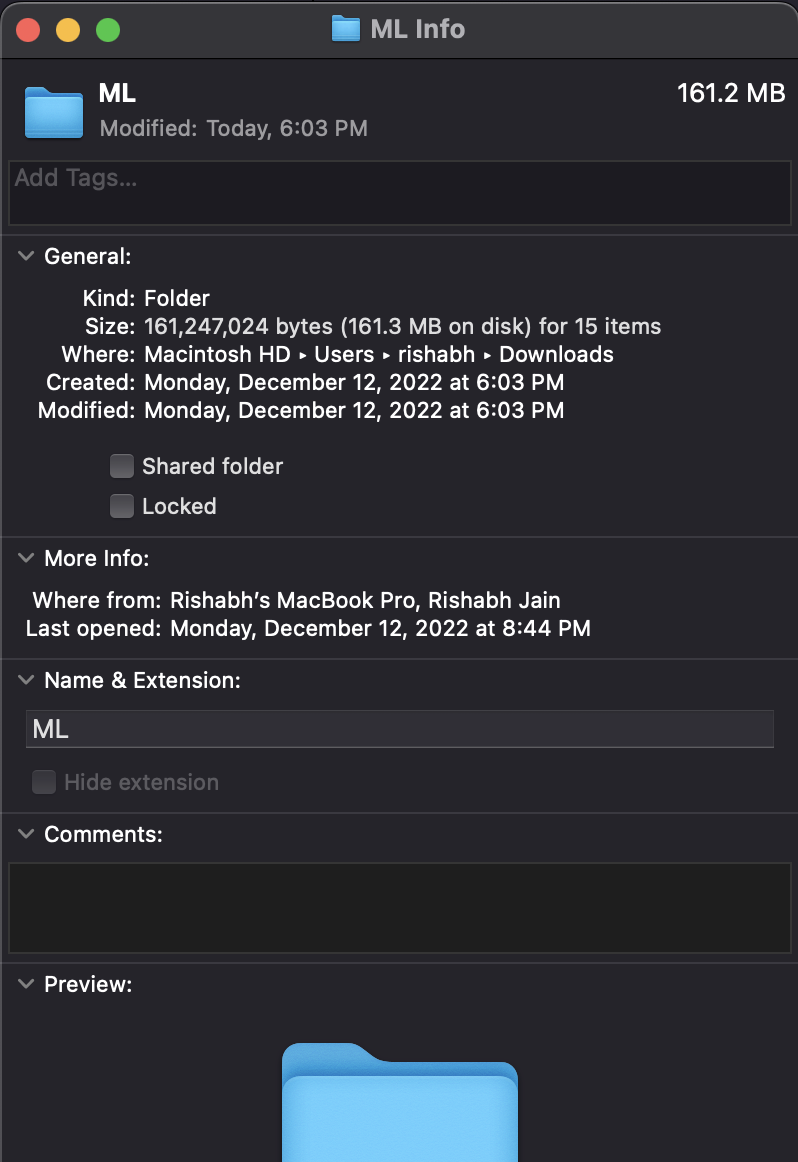
Project Contribution

| **Team Members** | **Contribution** |
| --- | --- |
| AKSHAY SADAGELKAR | 1. Collected the Dataset for ML Model 2. Designed the ML architecture 3. Calibrated the ML model for 80% accuracy 4. Designed & Developed Flask Application |
| RISHABH JAIN | 1. Designed and Programmed the Web Scraper 2. Running selenium on cloud run 3. Integrated all the components on Google Cloud Platform 4. Hosted the application on Cloud Run |
| ADITYA SHIVAJI PATIL | 1. Design and Created the SQL and BigQuery Databases 2. Integrated Database with the web application 3. Designed the Web Interface 4. Testing the application and fixing bugs |

References

* <https://cloud.google.com/run/docs>
* <https://www.roelpeters.be/how-to-deploy-a-scraping-script-and-selenium-in-google-cloud-run/>
* <https://medium.com/xebia-france/how-to-deploy-your-own-ml-model-to-gcp-in-5-simple-steps-bf2b5898c1ab#6151>
* <https://www.selenium.dev/documentation/>
* <https://docs.docker.com/engine/reference/commandline/run/>
* <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

Github Link : Since the file size is more than 100 mb, we are unable to upload the project on github. We have attached a working link of cloud run below.l Kindly check the link for reference.



Drive Link for Code - https://drive.google.com/drive/folders/1x5Rx9pNcWvDAFxvAld5Yfv8A7WCEiNAq?usp=sharing

Access Link : <https://app-fdk6eocuzq-vp.a.run.app/>