

## Project Design Phase-II Technology Stack (Architecture & Stack)

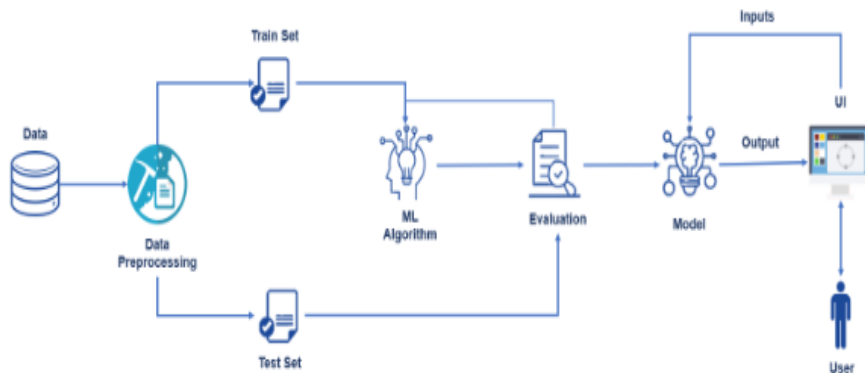
Date	15 October 2022
Team ID	PNT2022TMID21102
Project Name	Car Resale value Prediction
Maximum Marks	4 Marks

### Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2

### Car Resale Value Prediction

Reference: <https://developer.ibm.com/patterns/ai-powered-backend-system-for-order-processing-during-pandemics/>



#### Guidelines:

1. Include all the processes (As an application logic / Technology Block)
2. Provide infrastructural demarcation (Local / Cloud)
3. Indicate external interfaces (third party API's etc.)
4. Indicate Data Storage components / services
5. Indicate interface to machine learning models (if applicable)

**Table-1 : Components & Technologies:**

S.No	Component	Description	Technology
1.	Collect Dataset	ML is a data hunger technology, it depends heavily on data, without data, it is impossible for a machine to learn. It is the most crucial aspect that makes algorithm training possible.	We store the dataset as a Excel.
2.	Pre-Process the data	<ol style="list-style-type: none"><li>1. Handling the null values.</li><li>2. Handling the categorical values if any.</li><li>3. Normalize the data if required.</li><li>4. Identify the dependent and independent variables.</li><li>5. Split the dataset into train and test sets.</li></ol>	Python
3.	Model Building	<p>There are several Machine learning algorithms to be used depending on the data you are going to process such as images, sound, text, and numerical values. The algorithms can be chosen according to the objective. As the dataset which we are using is a REgression dataset so you can use the following algorithms</p> <p>Multi Linear Regression Random Forest Regression / Classification Decision Tree Regression / Classification K-Nearest Neighbors Support Vector Machine</p>	Python

4.	Application Building	After the model is built, we will be integrating it into a web application so that normal users can also use it to know the resale price of the care. In the application, the user provides the parameter values affecting the resale value.	HTML,CSS,Python flask
5.	File Storage	File storage is required to store the dataset(Excel)	Local Filesystem
6.	Machine Learning Model	Different regression models can be used to know the performance and choose whichever works better.	<ul style="list-style-type: none"> <li>• Multi Linear Regression</li> <li>• Random Forest Regression / Classification</li> <li>• Decision Tree Regression / Classification</li> <li>• K-Nearest Neighbors</li> <li>• Support Vector Machine</li> </ul>

**Table-2: Application Characteristics:**

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	<b>Flask:</b> It is a framework of microweb that is written in Python language and is classified as a microframework because it does not need any particular libraries and tools. Database abstraction layer, form validation and other such components with third-party libraries providing functionalities are all absent in flask.	Python Flask
2.	Security Implementations	All communications will be encrypted. This protects the application, data, and the personal information of the user from interception. The application will not store personal data of the user on the device. The application requires a valid SSL certificate be maintained at all times to allow	HTTPS,Encryptions,OWASP etc.

		trusted and secure communication.	
3.	Scalable Architecture	Our solution provides approximately 0.7 accuracy for training dataset and 0.6 accuracy for test dataset in the case of higher workloads also.	Train and test the dataset using python machine learning
4.	Availability	The system should always be available for access at 24 hours, 7 days a week. Also in the occurrence of any major system malfunctioning, the system should be available in 1 to 2 working days, so that business process is not severely affected.	Cloud infrastructure
5.	Performance	The system response time for every instruction conducted by the user must not exceed more than a minimum of 10 seconds. The system should have high performance rate when executing user's input and should be able to provide response within a short time span usually 50 second for highly complicated task and 20 to 25 seconds for less complicated task.	Machine Learning Techniques