Final Capstone Project Proposal Outline

I. Introduction:

The proposed project aims to develop a machine learning-based web application that predicts the price of a laptop based on its specifications. This app will assist users in making informed purchasing decisions.

II. Project Objective:

The primary objective of this project is to develop a machine learning model capable of accurately predicting laptop prices based on their specifications. In tandem with this, an intuitive web application will be created, allowing users to input various laptop specifications and receive real-time price predictions.

III. Data Description:

Sources: E-commerce websites such as Amazon.

Attributes: Brand, processor, RAM, storage, screen size, GPU, operating system, and additional features (e.g., touch screen, IPS).

Data Quantity: Aim for a dataset comprising at least 1,000 laptop entries.

IV. Methodology:

Data Cleaning: Handle missing values, outliers, and inconsistencies.

Feature Engineering: Convert categorical variables to numerical values, normalize data, and create new features if necessary.

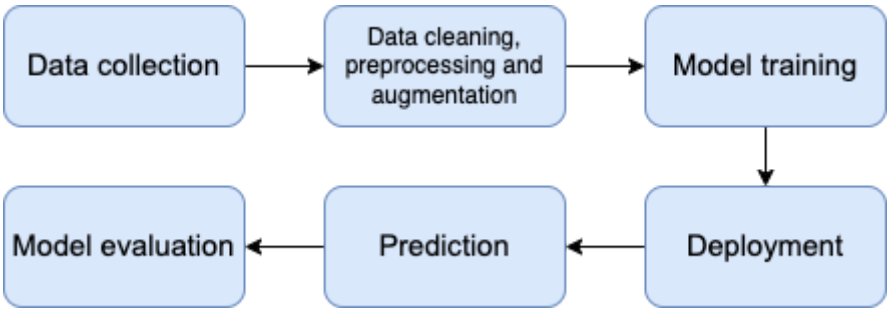
Data Splitting: Divide data into training, validation, and test sets.

V. Expected Deliverables:

Algorithms: Evaluate multiple algorithms such as Linear Regression, Decision Trees, Random Forest, and Gradient Boosting Machines.
Model Training: Train models using the training dataset and fine-tune using hyperparameter optimization.

Evaluation: Assess models based on accuracy, mean squared error, and other relevant metrics.

VI. Timeline and Tasks:



VII. Outcome:

- A functional web application that predicts laptop prices based on user-input specifications.
- A robust machine learning model with high prediction accuracy.
- A valuable tool for consumers to make informed laptop purchasing decisions.

VII. Potential Challenges:

- **Handling Missing Values:** Dealing with missing or incomplete data can be complex, requiring careful imputation or exclusion methods.
- **Feature Engineering:** Identifying and creating relevant features that accurately represent the data and improve model performance can be time-consuming and require domain expertise.
- **Model Validation:** Thoroughly validating the model to ensure its predictions are accurate and reliable across different scenarios and user inputs.
- **Maintenance:** Regularly updating the model with new data to maintain its accuracy and relevance, and handling any issues that arise post-deployment.