

AI LAB 10: Data Pre-processing and Feature Engineering Report

1. Project Overview

The objective of this lab is to prepare the raw Laptop Dataset for the Laptop Price Prediction machine learning model. This process involves Data Cleaning, handling missing values, and converting categorical features into numeric format.

2. Data Loading and Initial Cleaning

Key actions include importing libraries, loading the dataset, and dropping unnecessary identifier columns (Unnamed: 0, id).

3. Data Cleaning and Type Conversion

Object-type numerical fields were cleaned: - RAM: 'GB' removed and converted to integer. - Weight: 'kg' removed and converted to float.

4. Null Value Handling (Imputation)

Categorical values filled with Mode; Numerical values filled with Mean.

5. Feature Selection and Encoding

Target variable defined as Price, remaining columns treated as features. Categorical features encoded using pandas factorize().