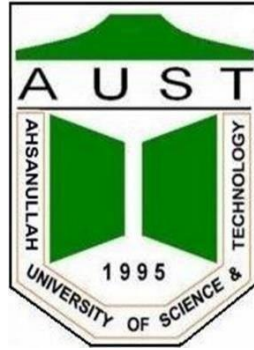# Ahsanullah University of Science and Technology

Department of Computer Science and Engineering

Program: Bachelor of Science in Computer Science and Engineering

Course No: CSE 4108

Course Title: Artificial Intelligence Lab

Project Name: **Laptop Price Prediction.**

Date of Submission: 4 September, 2022

Submitted to:

      Md. Siam Ansary
Lecturer, Department of CSE, AUST.

      Ms. Tamanna Tabassum
Lecturer, Department of CSE, AUST.

Submitted by:

    Md. Hujaifa Islam   Student ID: **18.02.04.021**

    Samiur Rahman Abir   Student ID: **18.02.04.008**

Introduction

We will predict camera price based on different useful features of a laptop of various brands. It will help the beginners to decide about which laptop they want to buy and how much money they need to buy their desired laptop. Feature-wise comparison of price can play significant role about decision making.

Dataset

We have constructed a dataset of about 240 laptop containing attributes like Brand Name, processor, processor type, ram, storage, storage type, display size, display type, graphic card and Price. About 192 instances will be used as train data and 48 instances will be used as test data. Our dataset contains various brands of laptop with the above mentioned features.

ML Models

1. **Multiple Linear Regression**:

   Multiple linear regression, also known as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of multiple linear regression is to model the linear relationship between the explanatory variables and response variables.

2. **Bayesian Ridge Regression**:

   Bayesian regression allows a natural mechanism to survive insufficient data or poorly distributed data by formulating linear regression using probability distributors rather than point estimates. The output or response 'y' is assumed to drawn from a probability distribution rather than estimated as a single value.

3. **Support Vector Machines Regression**:

   Support vector machine (SVM) analysis is a popular machine learning tool for classification and regression. SVM regression is considered a nonparametric technique because it relies on kernel functions. One of the main advantages of SVR is that itscomputational complexity does not depend on the dimensionality of the input space. Additionally, it has excellent generalization capability, with high prediction accuracy.

4. **Decision Trees Regression**:

   Decision tree builds regression or classification models in the form of a tree structure. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision nodes and leaf nodes. A decision node has two or more branches, each representing values for the attribute tested. Leaf node represents a decision on the numerical target. The topmost decision node in a tree which corresponds to the best predictor called root node. Decision trees can handle both categorical and numerical data.

5. **Nearest Neighbors Regression**:

    KNN regression is a non-parametric method that, in an intuitive manner, approximates the association between independent variables and the continuous outcome by averaging the observations in the same neighborhood. The size of the neighborhood needs to be set by the analyst or can be chosen using cross-validation to select the size that minimizes the mean-squared error.

Comparison of Performance Scores

Name of Algorithm	Mean Absolute Error	Mean Squared Error	Root Mean Squared Error	Mean Absolute Percentage Error
Multiple Linear Regression	17235.575	708022535.85	26608.692	15.922 %
Bayesian Ridge Regression	18559.774	779598350.9951	27921.288	17.527 %
Support Vector Machines Regression	55042.087	7551990673.11	86902.19	38.812%
Decision Trees Regression	12556.040	495726324.20	22264.912	9.96%
Nearest Neighbors Regression	14849.387	656114159.183	25614.725	15.911%

Discussion

The error rates are so high because we have trained the models using very small dataset containing only 240 instances. Decision Trees Regression model performs better, i.e., give less erroneous result among the models we have used. We can get better result from our models by using large dataset to train the models. We are keeping it as the future improvement scope of our project.

Members Contribution

Members Contribution

■ 180204008   ■ 180204021
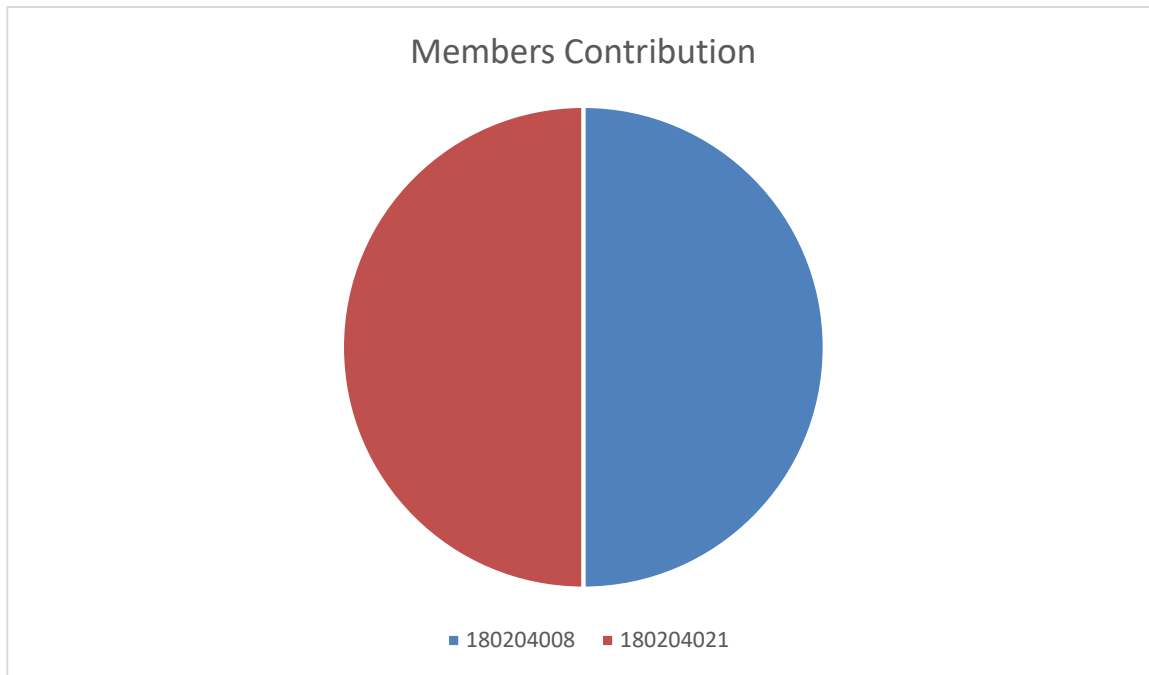
1. **18.02.04.021**: Dataset collection of laptop prices, Code implementation of Multiple Linear Regression, Bayesian Ridge Regression and user input.

2. **18.02.04.008**: Dataset collection of laptop, Code implementation of Support Vector Machines Regression and Decision Trees Regression.