

Machine Learning, Winter 2023
Project 1

Due date is October 25, 2023 at 11:59 PM
Submitted in groups of maximum 2

In this project, you will explore and implement linear and polynomial regression techniques using the provided dataset. The main objective is to gain a practical understanding of these regression methods and their applications.

1 Dataset

This dataset contains information about used cars. It is divided into two .csv files, train and test. This dataset contains details about used cars, including their name, manufacturing year, selling price, kilometers driven, fuel type, seller type, transmission, and owner history. All of these attributes are used to be able to predict the selling price of the used cars on the market to help sellers on setting a reasonable price, and for buyers to find the best value.

2 Linear Regression

The first part of the project focuses on linear regression. Two approaches will be implemented:

- Multi-Feature Equation: Implement linear regression using the multi-feature equation approach.
- Gradient Descent: Implement linear regression using the gradient descent approach.

Both approaches will be applied to the provided dataset to obtain standardized regression coefficients β . After implementing the two approaches, compare the accuracy of these models using **RMSE**, **MSE**, and **MAE** to assess model performance.

3 Polynomial Regression

Using the same data, implement Polynomial Regression Models with different degrees.

- Implement polynomial regression with degrees ranging from 1 to 10.
- Calculate RMSE values **for each degree**
- Create a plot to visualize the relationship between model **degree** and **RMSE**.

4 Report

Create a .pdf file that contains the following points:

- For the linear regression part
 - Document the **residual errors** and **RMSE values** for both the Multi-Feature Equation and Gradient Descent approaches.
 - Compare their performance and document which approach performed better according to your results and mention why.
- For the polynomial regression part
 - Maintain a table or list that records the **residual errors** and **RMSE values** for each degree, along with the corresponding degree of the polynomial model
 - Write what is the optimal degree according to your results and mention why.

PLAGIARISM IS NOT TOLERATED AND COPIED WORK WILL BE AWARDED 0 POINTS FOR BOTH TEAMS INVOLVED! There will be an individual evaluation for each team