

FACULTY OF COMPUTING

**Assignment-1**

Submitted By:

Areeba Khan (24267)

Abdul Rafay (25601)

SECTION: 6-A

# SUBJECT: AI Lab

BS-CS

SPRING 2023

INSTRUCTOR(S)-IN-CHARGE: Sir Aleem Raheem

Date: 28/05/2023

GitHub Link

<https://github.com/ABRJrocks/Abdul_Rafay_and_Areeba_Khan_Assignment>

1. Problem Description

The problem is to predict the price of gift items based on various attributes provided in the dataset.

3. Data Preprocessing

Before training a model, we need to preprocess the data. This involves handling missing values, encoding categorical variables, and splitting the data into training and testing sets.

4. Model Training

In this stage, we'll train a linear regression model using the preprocessed data. Linear regression is suitable for solving regression problems, where the goal is to predict a continuous target variable based on input features.

5. Model Evaluation

We'll evaluate the trained model's performance using appropriate evaluation metrics such as RMSE (Root Mean Squared Error) and R2 score. Lower RMSE and higher R2 score indicate better model performance.

6. Predictions and Further Analysis

We can utilize the trained model to make predictions on new gift items and perform further analysis on the dataset if needed.

Linear regression is a statistical modeling technique that can be used to analyze the relationship between a dependent variable and one or more independent variables. In the context of your dataset, linear regression can help in predicting the price of a gift item based on its attributes.

Here's how linear regression can be applied to solve this problem:

1. \*\*Identify dependent and independent variables:\*\* In your dataset, the "price" column represents the dependent variable that we want to predict. The other columns such as "gift\_type," "gift\_category," and "volumes" can be considered as independent variables that may have an impact on the price.

2. \*\*Data preprocessing:\*\* Before applying linear regression, it's important to preprocess the data by handling missing values and encoding categorical variables (if any) as discussed earlier.

3. \*\*Splitting the data:\*\* Once the data is preprocessed, it needs to be divided into training and testing sets. The training set will be used to train the linear regression model, and the testing set will be used to evaluate its performance.

4. \*\*Training the linear regression model:\*\* The training set is used to fit the linear regression model to the data. The model will learn the coefficients for each independent variable, representing the relationship between the independent variables and the dependent variable (price).

5. \*\*Model evaluation:\*\* After training the model, it is important to evaluate its performance using the testing set. Common evaluation metrics for linear regression include mean squared error (MSE), root mean squared error (RMSE), and R-squared.

6. \*\*Prediction:\*\* Once the linear regression model is trained and evaluated, it can be used to make predictions on new data. Given the attributes of a gift item (independent variables), the model can predict the corresponding price (dependent variable).

It's worth noting that linear regression assumes a linear relationship between the independent variables and the dependent variable. If the relationship is non-linear, other regression techniques or non-linear models may be more appropriate. Additionally, feature selection and feature engineering techniques can be applied to improve the model's performance.

Overall, linear regression provides a simple and interpretable approach for predicting the price of gift items based on their attributes.