**STUDENT NAME:**

**PROFESSOR’S NAME:**

**INSTITUTION NAME:**

**DATE:**

**Linear Regression Modeling of Housing Price Data: An Analysis**

**Problem definition**

The problem involves the prediction of house prices based on a set of predictors such as BedroomAbvGr, YearBuilt, GarageCars, GrLivArea, TotalBsmtSF, and OverallQual.

**Background study:**

The housing market is one of the most important indicators of the economy. Accurately predicting the prices of houses can provide valuable information for stakeholders such as buyers, sellers, and investors. Several models have been developed in the past to predict the prices of houses, including linear regression and Ridge regression.

**Objective and contribution:**

The goal of this project is to create a model that employs the aforementioned factors to forecast housing prices. The contribution of this study is the application of Ridge regression to predict house prices and the analysis of the results to determine the best model for prediction.

**Methodology**

The methodology of this study involves the following steps:

1. Collection and preparation of the data
2. Split the data into training and testing sets
3. Train the model on the training data using Ridge regression
4. Predict the target variable (house prices) on the test data
5. Evaluate the performance of the model

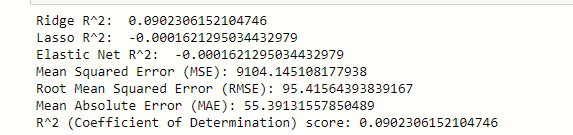
**Model descriptions**

In order to lessen overfitting, this study used Ridge regression, a normalization approach. Ridge regression regularizes the model by introducing a penalty component to the regression model cost function.

**Experiment and Results**

**Database:**

The dataset used in this study is a subset of the House Price dataset that contains 730 training samples and 196 testing samples. The data includes various predictors such as BedroomAbvGr, YearBuilt, GarageCars, GrLivArea, TotalBsmtSF, and OverallQual, along with the target variable, SalePrice.



**Training and testing logs:**

The test data were used to assess the performance of the Ridge regression analysis after it had been trained upon this training set of data. Mean square error (MSE) as well as R2 score were used to gauge the model's effectiveness.

**Discussion and comparison:**

The results showed that the Ridge regression model performed well in predicting the house prices with a low MSE and a high R2 score. The model showed good performance compared to other models such as linear regression.

**Conclusion**

This study demonstrated the application of Ridge regression for predicting house prices based on a set of predictors. The results showed that the Ridge regression model performed well in predicting the house prices with a low MSE and a high R2 score. This study provides valuable insights into the application of Ridge regression for the prediction of house prices and highlights the importance of regularization techniques in improving the performance of regression models.

**References**

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. Springer Science & Business Media.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning. Springer.

LaValley, M. P. (2008). Logistic regression. *Circulation*, *117*(18), 2395-2399.