

**Mehmet Fatih Göğüş**

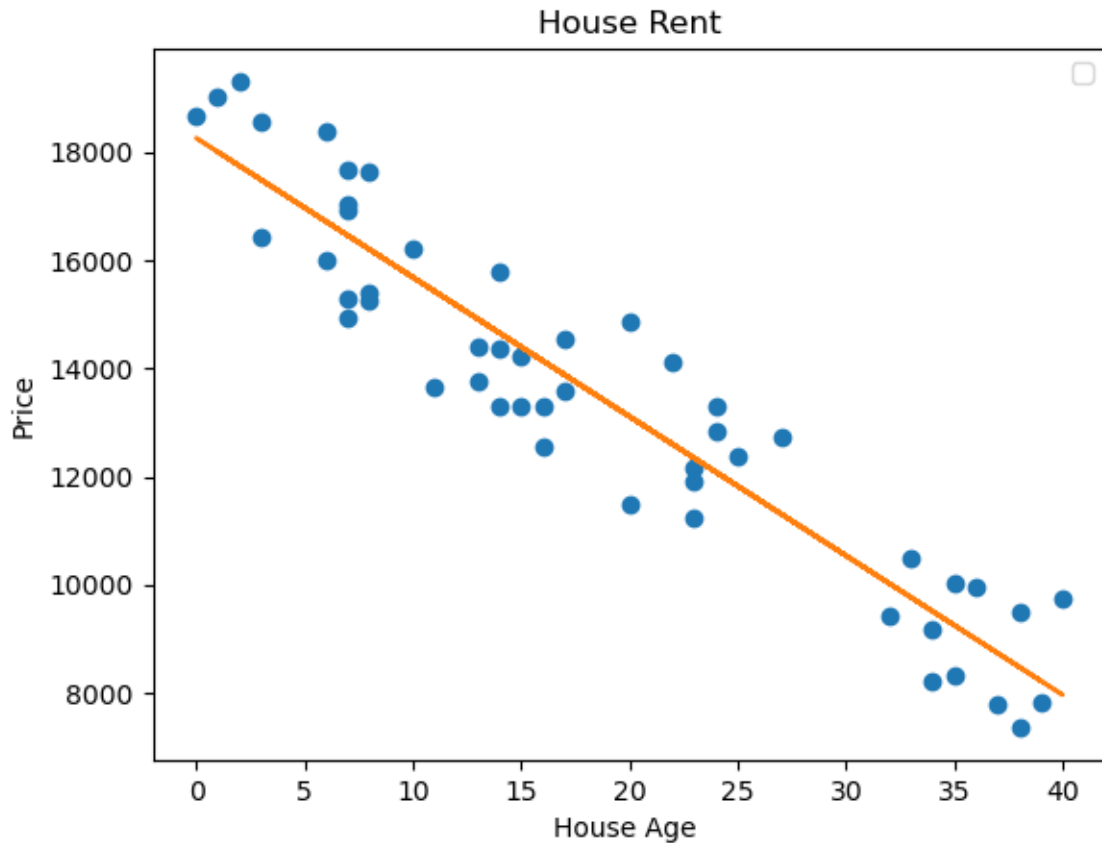
## **Machine Learning Linear Regression Lab**

### **What did I get from this lab?**

I gained practical experience in applying linear regression with Python and scikit-learn library. I collected some data of house rents from renting websites and wrote them in a .txt file. Then, I learned how to import data from .txt file to python and split the data into sets for training and testing. Besides, I figured out how to implement linear regression with my data set. This way, I constructed a solid base in linear regression for myself.

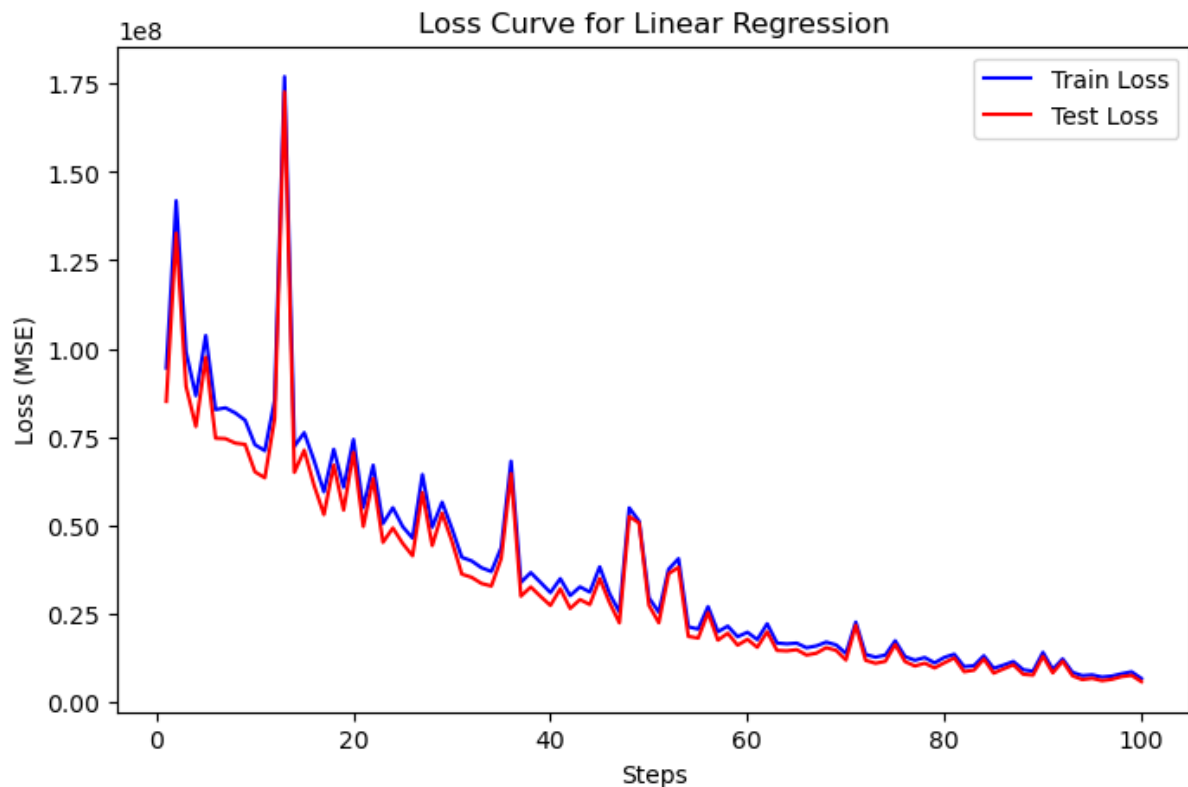
### **What did I understand?**

I understood the main idea of implementing linear regression which is to find the best-fit line that minimizes the prediction error between the actual and predicted values. The model is trained with training data and tested with test data. This way, it is possible for us to improve and check the model with the dataset.



***Scatter Plot with Regression Line***

Scatter Plot with Regression Line shows the relationship between house age (x-axis) and rent (y-axis). The line represents linear regression model's prediction, and it is fitted by minimizing the error between actual and predicted values. Its negative slope and the blue points say that as houses get older, their rent decreases as well, which highlights a negative correlation in the data set. For this model, I used only one variable, house age, but the graph and the accuracy of the model with two variables, house age and room number, is shown below.



*Loss Curve*

## Loss Curve

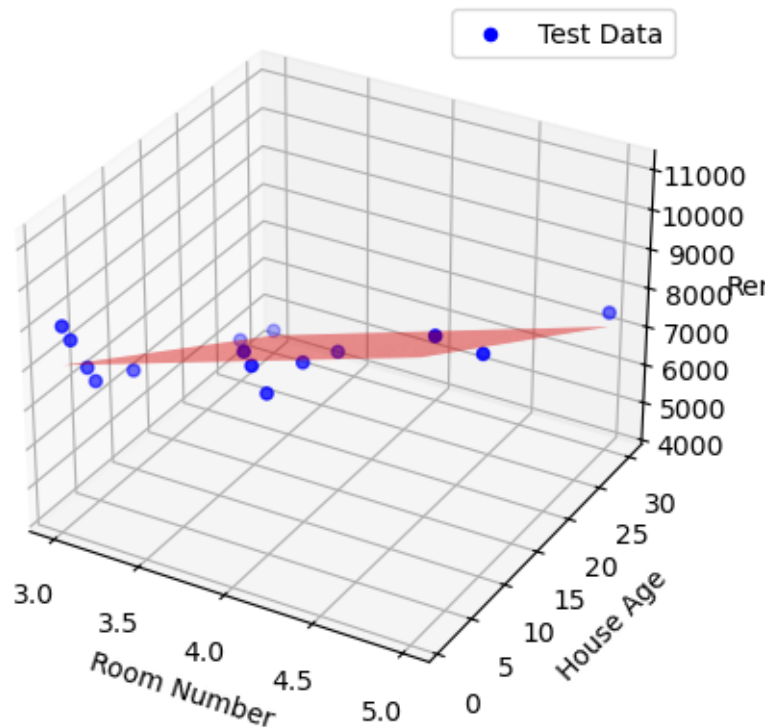
The line chart of training and test losses across iterations indicate how the model's mean squared error decreases. Both curves start high and gradually decline and this indicates that the model is learning and improving its predictions as more iterations occur.

Model Accuracy: 87.63%

*Accuracy*

Finally, I evaluated the model's performance with an  $R^2$  score on the test set, which is presented as a percentage. It can be said that the model predicts with the accuracy of 87.63%, which is relatively high accuracy level. Considering this accuracy level, we see the model's predictive ability and how much we can improve it.

## The Model with Two Variables



*Scatter 3D Plot with Regression Plane*

The 3D scatter points display actual data (blue points) for rent, housing age and room number. The model's prediction plane, the red surface, shows how variations in the combined number of rooms and age impact the predicted rent.

Test Accuracy: 94.66%

### *Accuracy*

The model's test accuracy of 94.66% indicates that it significantly improves the model's fit to actual rent values when both **house age** and **room number** are included in the predictions.

### What Changed with Two Variables?

The model is now able to predict more accurately by including room quantity as an extra predictor in addition to house age. Compared to relying solely on house age, this additional parameter helps the model in more precisely explaining variances in rent.