Reflection: Understanding a Linear Regression Model on Advertising Data (HW2)

In this homework, I analyzed a linear regression model to predict sales based on advertising spending in three channels: TV, radio, and newspaper. The code included feature scaling using StandardScaler, which standardized the input features so that each one had a mean of 0 and a standard deviation of 1. This step is important to ensure that features with larger ranges do not dominate the model.

By examining the model's coefficients, I learned that TV and radio advertising had strong positive effects on sales, while newspaper advertising had very little impact. The strongest predictor was TV advertising. For example, increasing TV ad spending by $1 is expected to increase sales by approximately 0.044 units, based on the model's parameters and scaling.

The model performed well, achieving an R^2 score of 0.91 on the training set and 0.86 on the test set, indicating it explains most of the variance in the data and generalizes well.

Although I did not write the code myself, analyzing and understanding it gave me valuable insights into how linear regression works, how to interpret the results, and the importance of preprocessing like feature scaling. It also helped me learn how to evaluate model performance using metrics like MAE, MSE, RMSE, and R^2 score.