Weekly Learning Report

(First Three Sessions, Colab & Notebook Review)

Name: Irfan Hayat

Submission date: 21 Aug 2025

TYM AI and ML Fellowship

Notebook filename (.ipynb): Linear_Regression_Model

Session 1:

In session 1 we pretty much covered the Basics of AI and ML and there basic terminologies and some core concepts like different types of learning, Feature engineering and thing such as overfitting and Cross validation and a bit of its history that I have learned and remembered completely.

Session 2:

In session 2 I learned about how different Algorithms represents different strategies for learning, they maybe linear, rule based or by using ensemble.

Also that if we want better interpretability we choose Linear regression on the other hand for higher predictive power we use forests.

I learned about Regression model matrix i.e how far off are we from our numerical predictions in this we have multiple one like Mean absolute error, Mean squared error etc.

Also learned about confusion matrix i.e Accuracy, Recall and precision etc.

And things such as Accuracy Paradox and Precision and Recall trade-off.

Session 3:

IN session 3 we learned about Neural networks i.e how they are brain inspired to learn complex, Hierarchical patterns how they work i.e learning and shallow learning and deep learnings difference , different layers i.e hidden layer, input and output layer etc.
Also learned about concepts of forward and backward propagation with the former being how a network makes a decision and the latter being the way to adjust the weights of different inputs or attributes in decision making . Also learned about the cheat like transfer learning concept and activation functions used to bring Non-linearity in the mix.

Collab and Notebook Review:

In this Assignment we basically learned about ETL and how we used a dataset of used cars in Pakistan to show linear regression in the car prices.
We first uploaded the data, then we handled missing values, we first try dropping rows with missing values but in that we lose too much data so we filled the missing values with the most

occurring values of that column.
Then we started the data Transformation stage, first we encoded the label and then we standardized the data e.g numerical features with too high a number were rescaled to have 0 as a mean and a standard deviation of 1.
Then we split the data one part consisting of 80% used to teach the model and the other 20% used as testing data.

Then we used the cleaned data and tried to find patterns among them. After cleaning the data, the first step was to get a quick overview, which confirmed the average car price and that the Toyota Corolla is a very popular model. To find deeper connections, they discovered a clear trend: cars with bigger engines have higher prices. This relationship was made obvious using a color-coded heatmap. Finally, various charts like histograms and box plots were used to visualize the data's patterns and easily spot any outliers, such as cars with unusually high prices.