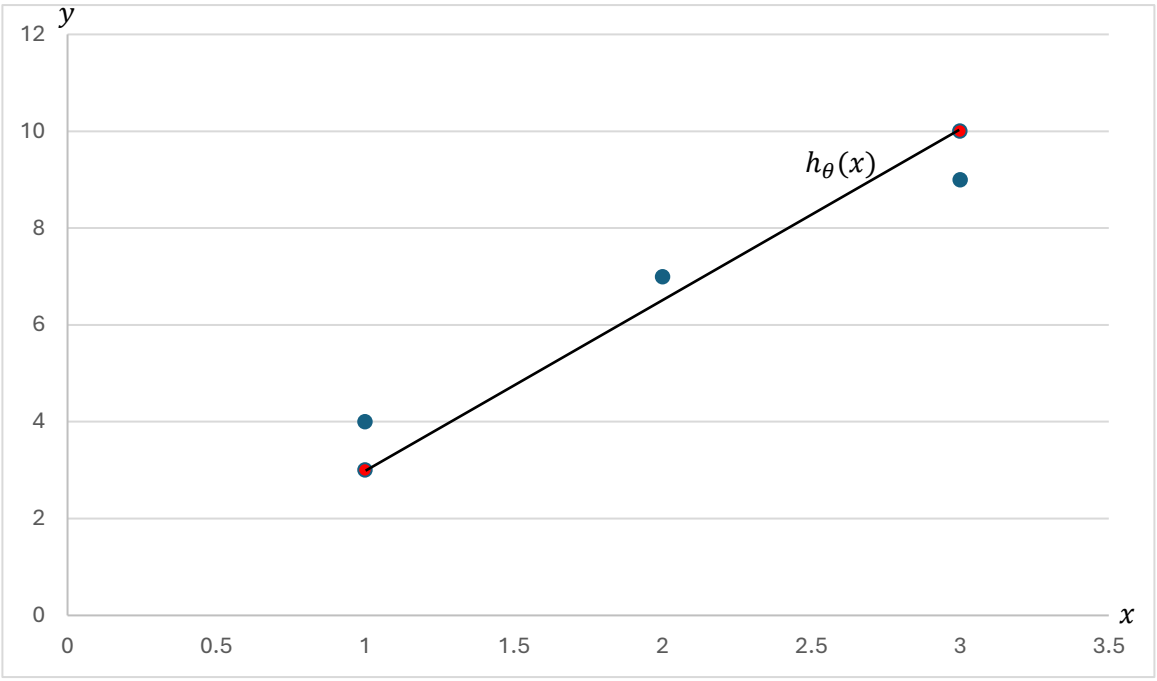
CS 335: Introduction to Large Language Models

Habib University

Activity Sheet 01

In the diagram provided, the hypothesis line $h_\theta(x)$ is plotted along with three data points: (1,4), (2,7), and (3,9).



(a) The hypothesis line $h_\theta(x) = \theta_1 x + \theta_0$ represents the linear regression model used to approximate the relationship between the input x and the target output y.

The line passes through the points (1,3) and (3,10). Find the values of the parameters θ_1 and θ_0.

(b) Calculate the predicted values $h(x_i) = \hat{y}_i$ for each data point x_i.

(c) A loss function $L(y, \hat{y})$ quantifies the gap between the actual data and the model.

$$L(y, \hat{y}) = \frac{1}{n}\sum_{1}^{n}(y_i - \hat{y}_i)^2$$

Using results from (b) calculate the loss.

(d) The hypothesis line does not perfectly fit the data points. Is it desirable for a model to perfectly fit the training data?