Exercise 3

Question 1 (5pt)

Use your own words to explain what is overfitting and underfitting? Assume you are developing a machine learning model, how should you identify and address them?

Question 2 (10pt)

Given the loss function of softmax regression:

$$L(W, b) = -\frac{1}{m} \sum_{i=1}^{m} \sum_{k=1}^{K} \left(\mathbb{1}\{y^{(i)} = k\} \log \frac{e^{W^{(k)}x^{(i)}+b^{(k)}}}{\sum_{j=1}^{K} e^{W^{(j)}x^{(i)}+b^{(j)}}} \right), \tag{1}$$

compute its derivative $\frac{\partial L}{\partial w}$, $\frac{\partial L}{\partial b}$.