* How can you detect overfitting or underfitting?

Generally, Overfittying can be detected if the error rate improves for a while then starts getting worse again. Training loss < Validation loss is not necessarily a sign of overfitting.

The model is underfitting if training loss is much larger than validation loss, which indicates that the model is too simplistic.

* Describe how gradient decent works.

Gradient descent is basically taking the derivative of the optimization function (usually the mean squared error) of the model, to identify the direction that helps us reach the global minima of the function or a point close to it. Global minima in this case is the point in which the error is minima, the derivative is taken with respect to the parameters that we are trying to learn (weights).

* What is the goal of regression?

Making predictions on a problem with continuous domain, for 1D linear regression this is done by finding the optimal value for the parameters of the line (the slope and the y-intercept) such that the average squared distance between the scattered datapoints (available data) and the line of best fit is minimized.