**2.1 Learning Regression Coefficients using Gradient Descent (60 points)**

**A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated**

**A screenshot of a computer

Description automatically generatedA screenshot of a cell phone

Description automatically generated**

**A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated**

**3. Least Squares Regression using Normal Equations (30 points)**

**A screenshot of a computer

Description automatically generated**

**4. Deriving Normal Equations for Univariate Regression(30 points)**

Consider the linear regression model for the case of a single input and output variable:

y = fw(x) = w0 + w1x

Assume that you have a training dataset consisting of N observations (xi),yi) ∀i ∈ {1,2,...,N}. Find the values of w0 and w1 that have the minimum sum of squares error on the dataset.

Answer:

A picture containing bird

Description automatically generated

**5.1 Polynomial Regression using Normal Equations(40 points)**

**A close up of a device

Description automatically generated**

**A screenshot of a cell phone

Description automatically generated**

**A close up of a screen

Description automatically generatedA screenshot of a cell phone

Description automatically generated**

**7. Programming Ridge Regression(50 points)**

**A screenshot of a cell phone

Description automatically generated**

**A screenshot of a cell phone

Description automatically generated**

**A close up of a piece of paper

Description automatically generated**

**A screenshot of a cell phone

Description automatically generated**

**7.1 Interpretation(20 points)**

Do you see a difference in the behavior of regularization of the two synthetic datasets we created in the two problems? Explain your results in detail, keeping in mind that a zero value for λ corresponds to the solution without regularization i.e., linear regression.

Answer:

**8. Maximum Likelihood For Univariate Normal (20 points)**

Consider N samples {x1, x2 . . . , xN }, generated from a univariate normal distribution:

A picture containing object, clock

Description automatically generated  
formulate the log-likelihood for the N samples, and derive the maximum likelihood estimate for

the mean of the distribution μML.

Answer:

A picture containing bird

Description automatically generated