

Q1. Explain what is linear and logistic regression, make sure to emphasize the proper context for each one. Your comparison must cover data type modelling, cost function and target application.

Logistic and linear regression are supervised learning techniques, the first one is used for classification while the second is used for numerical prediction that is given a set of numerical data and the curve that best fit such data predict the value of a new element in this set.

Categorical data and numerical data are respectively the possible datatypes capable of being modelled by each of these, logistic regression cost function is called log-likelihood, can also be the negative of this function, and is defined as follows

$$C(\theta) = \sum_i y_i \log h_{\theta}(x_i) + (1 - y_i) \log(1 - h_{\theta}(x_i))$$

the cost function for linear regression is

$$C(w, b) = \frac{1}{2n} \sum_i ((wx_i + b) - y_i)^2$$

finally we can consider the application of each one, logistic regression can be used for problems such as skin cancer detection, mail and call spam as well other myriad of applications. Linear prediction can be used to find cost of housing, insurance value prediction and others.