Machine Learning CSE 4020

Assignment 1

Date: April 15, 2019

Question 1: What is the effect of different initialization on the perceptron and Logistic Regression algorithm? Answer along with examples.

Question 2: How can we convert linear regression into logistic regression without using a link function (logit)?

Question 3: Take any dataset and perform Logistic regression. Plot the parameter space explored during the training.

Question 4: How does the normalization of the data help in training?

Question 5: Find the least no of perceptron that we can use to solve XOR.

Question 6: Given a linearly separable data set, is there only one solution (hyperplane) that a perceptron model can find?

Question 7: Solve the problem of logistic regression on any dataset by searching the parameters in the parameter space. Use any optimal heuristic to reduce the number of iteration.

Question 8: Use the following activation function and check the performance metrics:

a) Sigmoid

$$f(x) = \sigma(x) = \frac{1}{1 + e^{-x}}$$

b) Tanh

$$f(x) = \tanh(x) = \frac{(e^{x} - e^{-x})}{(e^{x} + e^{-x})}$$

c) Soft Clipping

$$f(\alpha, x) = \frac{1}{\alpha} \log \frac{1 + e^{\alpha x}}{1 + e^{\alpha(x-1)}}$$

d) Softmax

$$f_i(\vec{x}) = \frac{e^{x_i}}{\sum_{j=1}^{J} e^{x_j}} \quad \text{for } i = 1, \ldots, J$$

e) Elliot Sigmoid

$$f(x) = \frac{x}{1 + |x|}$$

Question 9: For the simple linear regression Y = A + BX, prove the following identity:

a) $\sum_{i=1}^{n}(y - \hat{y}) = 0$

b) $\sum_{i=1}^{n}(y) = \sum_{i=1}^{n}(\hat{y})$

c) $\sum_{i=1}^{n} x(y - \hat{y}) = 0$

d) $\sum_{i=1}^{n} \hat{y}(y - \hat{y}) = 0$

Question 10: If we use linear regression of form Y = AX, which of the above four identities will hold true?