Alperen Karaoglu N13012011 ak4410

Homework Assignment 1
Perceptron and Logistic Regression

Lecturer: Kyunghyun Cho

1. In the perceptron algorithm we try to make the weight vector w to be aligned more positively with $M^*(x)$. It is necessary to add a one when the perceptron has another input, the bias so that we can treat it just as another input. This allows us to shift the function curve along the input axis while leaving the slope determined by the weights unchanged.
For example when we consider the scenario provided in class with the robbery detection system, we might want our algorithm to be more likely to detect a robbery even if there isn't any (when it can't be 100% accurate) in order to diminish the total cost, in this case we would need to add a bias and adjust for it by augmenting the input vector x with an extra 1.

2. When $D(M^*(x),M,x) = 0$ we assume that the prediction is correct but there is another case in which we could get 0 as a value for the distance function and that is when the weight vector is 0. In this case we have to make sure that we don't account for the Distance function as being a correct prediction since the result was due to an anomaly and not an actual correct prediction. We can remedy this problem by ensuring that whenever the weight vector contains a zero element we set the distance function to something positive (1) and not 0 so that we don't account for this anomaly as a correct prediction in our final result.

3. On page 2

3.

$D(y^*, w, x) = -(y^* \log M(x) + (1-y^*) \log(1-M(x)))$

let $\alpha = M(x) = \sigma(w^T x)$ (sigmoid)

from class we know that $\sigma(w^T x) = \frac{1}{1+e^{-w^T x}}$

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

$\log(1-\alpha) = -w^T x - \log(1+e^{-w^T x})$

$\frac{\partial}{\partial w} \log \alpha = \frac{x e^{-w^T x}}{1+e^{-w^T x}} = \underline{x(1-\alpha)}$

$\frac{\partial}{\partial w} \log(1-\alpha) = -x + x(1-\alpha) = \underline{-\alpha x}$

plugging in the above results we get:

$\nabla_w D(y^*, w, x) = -(y^* x(1-\alpha) - (1-y^*)\alpha x)$

$= -(y^* x - y^* x\alpha - \alpha x + y^* x \alpha)$

$= (\alpha - y^*) x$

which is equivalent to the solution from class notes

$(\alpha - y^*) x = -(M^*(x) - M(x)) \overline{x}$