

HW4

211482559 207253899

1.

על פי ההגדרה:

$$Pr[Y = 1|X = x, w] = \Phi(w^T x), \quad \Phi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

בנוסף על פי הגדרת log likelihood נקבל:

$$l(w) = \sum_{i=1}^n [y_i \cdot \log(\Phi(w^T x_i)) + (1 - y_i) \cdot \log(1 - \Phi(w^T x_i))]$$

2.

נגדיר את הBCE:

$$BCE(w) = - \sum_{i=1}^n [y_i \cdot \log(\Phi(w^T x_i)) + (1 - y_i) \cdot \log(1 - \Phi(w^T x_i))]$$

למצוא את המקסימום של $l(w)$ שקול למציאת המינימום של $BCE(w)$

3.

נחשב את הנגזרת של $BCE(w)$:

נגדיר:

$$\phi(x) = \frac{d}{dx} \Phi = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

ונחשב:

$$\nabla_w BCE(w) = - \sum_{i=1}^n \left[\frac{y_i}{\Phi(w^T x_i)} - \frac{1-y_i}{1-\Phi(w^T x_i)} \right] \cdot x_i \cdot \phi(w^T x_i)$$

על מנת למצוא את הערך, נבחר w רנדומלי ולאחר מכן נריץ את החישוב עד להתכנסות כך שבכל פעם:

$$w = w - LearninRate \cdot \nabla_w BCE(w)$$

.4

אין צורך להוסיף ביאס כי כלל השורות מתחילות במקדם (פיצ'ר) זהה 1.

.5

```
#1, G: [-1.59576912 -6.38307649 -3.9894228 -5.58519193 -3.9894228 ], L: 5.545177444479562
#2, G: [-0.47103524 -4.89271261 -2.74715295 -3.87662579 -0.72281722], L: 4.681374238526555
#3, G: [-0.04238906 -4.08249951 -2.03185723 -3.14498851 0.43763178], L: 4.258822937834527
#4, G: [ 0.13856614 -3.54204677 -1.54208438 -2.76098277 0.87479489], L: 3.971489704554526
```

.6

```
Predictions: [1, 0, 1, 1, 0, 1, 1, 1]
Classification Errors: [3, 4, 5]
Accuracy: 0.625
```