ROB 313 Assignment 3

Q 2. A. $\hat{f}(\underline{x}^{(i)}; \underline{w}) = 1 \quad , \quad y^{(i)} = 0$

$$\log Pr(\underline{y} \mid \underline{w}, \underline{X}) = \sum_{i=1}^{N} y^{(i)} \log[\hat{f}(\underline{x}^{(i)}; \underline{w})] + (1 - y^{(i)}) \log(1 - \hat{f}(\underline{x}^{(i)}; \underline{w}))$$

$$= \sum_{i=1}^{N} \underbrace{y^{(i)}}_{0} \underbrace{\log(1)}_{0} + \underbrace{(1 - y^{(i)})}_{0} \underbrace{\log(0)}_{-\infty}$$

$$= -\infty$$

This is not reasonable behaviour since the log-likelihood blows-up to negative infinity and could cause numerical issues.

B. Prior: $Pr(\underline{w}) = \mathcal{N}(\underline{w} \mid \underline{0}, \sigma^2 \underline{I})$

GD

$$\underline{w}_{k+1} = \underline{w}_k - \eta \nabla_{\underline{w}}(\log(Pr(\underline{y} \mid \underline{w}, \underline{X})) + \log(Pr(\underline{w})))$$

$$= \underline{w}_k - \eta \sum_{i=1}^{N} (y^{(i)} - \hat{f}(\underline{x}^{(i)}; \underline{w}) [1, x_1^{(i)}, \ldots, x_D^{(i)}]^T) + \nabla_{\underline{w}} \log((\tfrac{1}{\sqrt{2\pi\sigma^2}})^D \exp(-\sum_{i=1}^{N} \tfrac{1}{2\sigma^2} \underline{w}^T$$

$$= \underline{w}_k - \eta \sum_{i=1}^{N} (y^{(i)} - \hat{f}(\underline{x}^{(i)}; \underline{w}) [1, x_1^{(i)}, \ldots, x_D^{(i)}]^T + \sum_{i=1}^{N} (\tfrac{1}{2\sigma^2}) \underline{w}$$

$$= \underline{w}_k - \eta \sum_{i=1}^{N} ((y^{(i)} - \hat{f}(\underline{x}^{(i)}; \underline{w}) [1, x_1^{(i)}, \ldots, x_D^{(i)}]^T) + (\tfrac{1}{2\sigma^2}) \underline{w})$$

SGD since predicting on only the random sample per

$$\underline{w}_{k+1} = \underline{w}_k - \eta ((y^{(t)} - \hat{f}(\underline{x}^{(t)}; \underline{w}) [1, x_1^{(t)}, \ldots, x_D^{(t)}]^T) + (\tfrac{1}{2\sigma^2}) \underline{w})$$