Siyuan Zou 111639762 CSE353 Hw2

Theory Portion

I. A) $z = \mathrm{w_0} + \mathrm{w_1 x_1} + \cdots + \mathrm{w_n x_n}$ or $\langle \mathrm{w}, \mathrm{x} \rangle$, $sigmoid(z) = \frac{1}{1+e^{-z}}$, $y = \begin{cases} 1, & sigmoid(z) \\ 0, & 1 - sigmoid(z) \end{cases}$

B) $P(y = 1|x) = \frac{P(x|y = 1) * P(x)}{P(y=1)}, P(y = 0|x) = \frac{P(x|y = 0) * P(x)}{P(y=0)}$

C) $P\left(x_j \middle| d_i\right) = \frac{1}{\sigma_{ji}\sqrt{2\pi}} \exp\left(-\frac{\left(x_j - \mu_{ji}\right)^2}{2\sigma_{ji}^2}\right)$

D) $P(y = 1|x) = \frac{P\left(x_j \middle| d_i\right) * P(d_i)}{P(y=1)} \propto P\left(x_j \middle| d_i\right) * P(d_i)$

E) huh?

II. Cross-Verification Failure:

Since we know that the $P(y = 1) = P(y = 0) = \frac{1}{2}$ $or\ the\ loss_D(h) = \frac{1}{2}$ is used when choosing the label, and that the learning predictor:

$$h(x) = \begin{cases} 1, if\ number\ of\ 1s\ in\ training\ set\ labels\ is\ odd \\ \qquad\qquad 0, otherwise \end{cases}$$

Let say that the number of 1s in training set labels is odd, then you leave-one-out, thus making the predictor constantly output 0, so since the output is all the same then the $loss_S(h) = 1$. And vice versa this is true when training set labels is even. So $loss_s(h) - loss_D(h) = 1/2$

III. Decision Tree:

Labels = {0, 1}, $and\ data\ points\ of\ x_i\ i \in (1, \ldots, d)$. Decision tree is a recursive algorithm that build the tree based on test or question ($is\ x_i = 0$) we recursively find best feature x and split on it, since there is d numbers of x we split on it d+1 because each split we would remove the used feature.