ML HW#2

Joshua
Shapiro
10/15/16

1)

1) Write down an expression for the likelihood of D in terms of π, p, & n.

$$\boxed{\mathcal{L}(\pi|D) = \pi^{p}(1-\pi)^{n}}$$

2) By differentiating the log-likelihood L, find the value of π that maximizes the likelihood.

1) Find log-likelihood:

$$\log(\mathcal{L}(\pi|D)) = \log(\pi^{p}(1-\pi)^{n}) = p\log(\pi) + n\log(1-\pi)$$

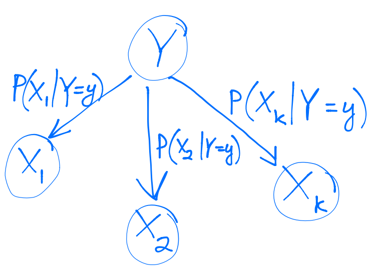
2) Differentiate the log-likelihood & set = to 0

$$\frac{d}{d\pi}\left(\log(\mathcal{L}(\pi|D))\right) = \frac{p}{\pi} + \frac{n}{1-\pi}$$

$$\frac{p}{\pi} + \frac{n}{1-\pi} = 0 \Rightarrow \frac{(p\pi - p) - n\pi}{\pi(\pi-1)} = 0 \Rightarrow p\pi - p - n\pi = 0 \Rightarrow \pi(p-n) - p = 0 \Rightarrow$$

$$\pi(p-n) = p \Rightarrow \boxed{\pi = \frac{p}{p-n}} \leftarrow \text{value of } \pi \text{ that maximizes log-likelihood}$$

3) Assume we add k boolean variables that describe each sample, and suppose these vars are conditionally independent of eachother given Y. Draw the bayes net.



4) Write down the likelihood for the data using the following additional notation

$$\mathcal{L}(\alpha_i, \beta_i, p_i^+, n_i^+, p_i^-, n_i^- | D) = \prod_{i}^{k}\left((\alpha_i)^{p_i^+}(1-\alpha_i)^{n_i^+}(\beta_i)^{p_i^-}(1-\beta_i)^{n_i^-}\right)$$

5) Find α_i & β_i that maximize the log likelihood

$$\log\left((\alpha_i)^{p_i^+}(1-\alpha_i)^{n_i^+}(\beta_i)^{p_i^-}(1-\beta_i)^{n_i^-}\right) = p_i^+\log(\alpha_i) + n_i^+\log(1-\alpha_i) + p_i^-\log(\beta_i) + n_i^-\log(1-\beta_i)$$

$$\frac{\partial}{\partial\alpha_i}\left(p_i^+\log(\alpha_i) + n_i^+\log(1-\alpha_i) + p_i^-\log(\beta_i) + n_i^-\log(1-\beta_i)\right) = \frac{p_i^+}{\alpha_i} + \frac{n_i^+}{1-\alpha_i}$$

$$\Rightarrow \frac{p_i^+}{\alpha_i} + \frac{n_i^+}{1-\alpha_i} = 0 \Rightarrow \frac{(p_i^+\alpha_i - p_i^+) - n_i^+\alpha_i}{\alpha_i(\alpha_i - 1)} = 0 \Rightarrow p_i^+\alpha_i - p_i^+ - n_i^+\alpha_i = 0 \Rightarrow \alpha_i(p_i^+ - n_i^+) - p_i^+ = 0$$

$$\Rightarrow \boxed{\alpha_i = \frac{p_i^+}{p_i^+ - n_i^+}}$$

$$\frac{\partial}{\partial\beta_i}\left(p_i^+\log(\alpha_i) + n_i^+\log(1-\alpha_i) + p_i^-\log(\beta_i) + n_i^-\log(1-\beta_i)\right) = \frac{p_i^-}{\beta_i} + \frac{n_i^-}{1-\beta_i}$$

Following the same logic shown above, $\boxed{\beta_i = \frac{p_i^-}{p_i^- - n_i^-}}$

3) c) Find the error rate of your Naive Bayes algorithm on the test set

My algorithm incorrectly labeled 16 points, yielding an error rate of 20.0%

Did Not do 4.