CIS 419/519: Quiz 10

November 25, 2019

1. Given the joint probability table below:

P(smart, study, prepared)	smart&study	¬smart&study	smart&¬ study	¬ smart& ¬ study
prepared	0.3	0.1	0.05	0.05
¬ prepared	0.2	0.1	0.1	0.1

   1) Is "smart" conditionally independent of "prepared", given "study"?

   2) Is "study" conditionally independent of "prepared, given "smart"?

   Answer: No, No

2. Which of the following statement about Naive Bayes classifiers are correct (multiple answers may be correct)?

   Assume V is the set of output labels and each data instance X is represented as $x = [x_1, x_2, ..., x_n]$ using n features.

   Answer:

   1) To determine the NB prediction $v \in V$ on $x = [x_1, x_2, ..., x_n]$ it is necessary to estimate the conditional probability $P(x_i|v)$ for all $i = 1...n$

   2) To determine the NB prediction $v \in V$ on $x = [x_1, x_2, ..., x_n]$ it is necessary to estimate the probability $P(v)$ for all $v \in V$

3. Suppose $y_1, y_2, ..., y_n$ are i.i.d random variables, each having the probability density function:

   $$f_Y(y) = \theta y^{\theta-1}, 0 < y < 1$$

   What is the correct MLE of θ?

   Answer: $\hat{\theta}_{MLE} = -\frac{n}{\sum log(y_i)}$

   Solution: Due to i.i.d assumption, we multiply all n random variables to get our total probability function. Then, take the derivative w.r.t θ and find the θ that maximizes this function:

$$\prod_{i=1}^{n} f_Y(y_i) = \theta^n \prod_{i=1}^{n} y_i^{\theta-1}$$

$$log(\prod_{i=1}^{n} f_Y(y_i)) = nlog(\theta) + (\theta - 1)\sum_{i=1}^{n} log(y_i)$$

$$\frac{dlog(\prod_{i=1}^{n} f_Y(y_i))}{d\theta} = \frac{n}{\theta} + \sum_{i=1}^{n} log(y_i)$$

$$0 = \frac{n}{\theta} + \sum_{i=1}^{n} log(y_i)$$

$$\hat{\theta}_{MLE} = -\frac{n}{\sum_{i=1}^{n} log(y_i)}$$

4. Suppose $z_1, z_2, ..., z_n$ are i.i.d random variables, each having a Poisson distribution with a probability density function:

$$P_Z(z) = \frac{e^{-\lambda}\lambda^z}{z!}$$

What is the correct MLE of λ?

Answer: $\hat{\lambda}_{MLE} = \frac{\sum z_i}{n}$
Solution: Due to i.i.d assumption, we multiply all n random variables to get our total probability function. Then, take the derivative w.r.t λ and find the λ that maximizes this function:

$$\prod_{i=1}^{n} P_Z(z_i) = e^{-n\lambda} \prod_{i=1}^{n} \frac{\lambda^{z_i}}{z_i!}$$

$$log(\prod_{i=1}^{n} P_Z(z_i)) = -n\lambda + log(\lambda)\sum_{i=1}^{n} z_i - \sum_{i=1}^{n} log(z_i!)$$

$$\frac{dlog(\prod_{i=1}^{n} P_Z(z_i))}{d\lambda} = -n + \frac{1}{\lambda}\sum_{i=1}^{n} z_i$$

$$0 = -n + \frac{1}{\lambda}\sum_{i=1}^{n} z_i$$

$$\hat{\lambda}_{MLE} = \frac{\sum_{i=1}^{n} z_i}{n}$$

5. Say we have two features A and B and a set of possible labels $v \in V$. If we know that the values of A and B are conditionally independent given the label v, what can you say about the independence of the variable A and B?

   Answer: Not enough information