Homework set 1 TDA231 Algorithms for Machine Learning Inference

Kevin Jaquier 921119T334
Sandra Viknander 9003012482

March 2018

1 Theoretical problems

1.1 Bays Rule

We can think of the process of testing as having four superposition's, the test can come back positive when you are sick making it a true positive it could also come back negative while you are sick which is a false positive. Likewise if you are healthy the test can come back either positive or negative making a false negative and true negative respectively. This is also depicted in figure 1. The probability of testing positive if you are infected and testing negative if you are healthy are $P(\text{Positive Test If Cancer}) = 0.99$ and $P(\text{Negative Test If Healthy}) = 0.99$ respectively. This is a rare type of Cancer that only affect 1 in 10000 thus $P(\text{Cancer}) = 0.0001$.

We are now interested in the probability that we are sick if the test shows positive that is $P(\text{Cancer}|\text{Test is positive})$ which can be calculated using bays rule shown in equation 1. We now need to find the probability of getting a positive test $P(\text{Positive Test})$, we do this by finding both the probability of true positives and of false negatives, and add them together as in equation 2. Now if we put in the values of P(Positive Test If Cancer), P(Negative Test If Healthy) and P(Cancer) we get the probability that we in fact have Cancer. The calculations give us $P(\text{Cancer}|\text{Test is positive}) = 9.9 * 10^{-5}$

$$P(\text{Cancer}|\text{Test is positive}) = P(\text{Test is positive}|\text{Cancer}) * \frac{P(\text{Cancer})}{P(\text{Positive Test})} \tag{1}$$

$$\begin{aligned} P(\text{Positive Test}) = & P(\text{Positive Test if Cancer}) * P(\text{Cancer}) \\ & + (1 - P(\text{Negative Test if Healthy})) * (1 - P(\text{Cancer})) \end{aligned} \tag{2}$$

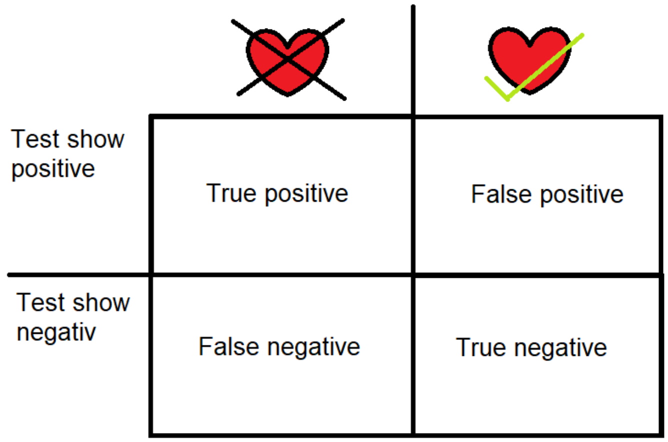

Figure 1: Figure displaying the different outcomes that tests can give given the health state of the patient. In the left boxes the patient are sick and in the right boxes the patients are healthy. Thus if a test comes back positive it could either be a true positive or a false positive depending on the actual state of the patient. Likewise if the test is negative it could be a false negative or true negative.

1.2 Correlations and Independence

Covariance is defined as follow:

$$Cov(X,Y) = E[XY] - E[X]E[Y] \tag{3}$$

Given a uniform distribution X on the interval $[-a,b]$, and the functions f and g defined over that interval where f is even and g is odd, we know:

$$E[f(X)] = \frac{f(b) - f(a)}{2} \tag{4}$$

$$E[g(X)] = 0 \tag{5}$$

X is a uniform distribution on the interval $[-1,1]$. Because of the property given by equation 5, we have $E[X] = 0$ and thus $E[X]E[Y] = 0$. Also, $Y = X^2$ which gives $E[XY] = E[X^3]$. X^3 is an odd function, and since X is an uniform distribution on the interval [-1,1], X^3 will also be a uniform distribution on that interval and thus $E[X^3] = 0$. So $E[XY] = 0$ which leads to $Cov(X,Y) = 0$.

1.3 Setting hyperparameters

For a Beta distribution θ we have that the expectation value $E[\theta] = \frac{a}{a+b}$ and the variance $Var[\theta] = \frac{ab}{(a+b)^2(a+b+1)}$. To scale a and b so that the expectation value and the variance can take constant values m and v respectively we use substitution for b as shown in 6. Then we can solve for a as shown in equation 7. Then by substituting a with the answer we got from 7 in to the expression for b in equation 6 we get an expression for b in terms of m and v as shown in equation 8. Thus $a = \frac{m^2(1-m)-vm}{v}$ and $b = \frac{m(1-m)^2+v(m-1)}{v}$ written in terms of the constants m and v. It is important that m and v are only valid in the interval (0,1) since θ is a distribution on the interval [0,1].

$$
\begin{aligned}
&\frac{a}{a+b} = m \implies \frac{a(1-m)}{m} = b \\
&\frac{a^2\frac{1-m}{m}}{(a+\frac{a(1-m)}{m})^2(a+\frac{a(1-m)}{m}+1)} = v \\
&\frac{\frac{a^2(1-m)}{m}}{\frac{a^2(a+m)}{m^3}} = \frac{m^2(1-m)}{a+m} = v
\end{aligned} \tag{6}
$$

$$
\begin{aligned}
\frac{m^2(1-m)}{a+m} &= v \\
\frac{m^2(1-m)-vm}{v} &= a
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
\frac{a(1-m)}{m} &= b \\
\frac{(m^2(1-m)-vm)(1-m)}{vm} &= \\
\frac{m(1-m)^2+v(m-1)}{v} &= b
\end{aligned} \tag{8}
$$

2 Practical problems

See the completed Jupyter notebook.