Hw 7

Alyson Longworth

4/18/2024

Recall that in class we showed that for randomized response differential privacy based on a fair coin (that is a coin that lands heads up with probability 0.5), the estimated proportion of incriminating observations $\hat{P}$ [1] was given by $\hat{P} = 2\pi - \frac{1}{2}$ where π is the proportion of people answering affirmative to the incriminating question.

I want you to generalize this result for a potentially biased coin. That is, for a differentially private mechanism that uses a coin landing heads up with probability $0 \leq \theta \leq 1$, find an estimate $\hat{P}$ for the proportion of incriminating observations. This expression should be in terms of θ and π.

$\hat{P} = \frac{\pi-(1-\theta)\theta}{\theta}$

Next, show that this expression reduces to our result from class in the special case where $\theta = \frac{1}{2}$.

$\hat{P} = \frac{\pi-(1-0.5)0.5}{0.5}$ $\hat{P} = \frac{\pi-(0.5-0.25)}{0.5}$ $\hat{P} = \frac{\pi-(0.25)}{0.5}$ $\hat{P} = \frac{\pi}{0.5} - \frac{(0.25)}{0.5}$ $\hat{P} = 2\pi - \frac{1}{2}$

Consider the additive feature attribution model: $g(x') = \phi_0 + \sum_{i=1}^{M} \phi_i x_i'$ where we are aiming to explain prediction f with model g around input x with simplified input x'. Moreover, M is the number of input features.

Give an expression for the explanation model g in the case where all attributes are meaningless, and interpret this expression. Secondly, give an expression for the relative contribution of feature i to the explanation model.

Don't Answer

Part of having an explainable model is being able to implement the algorithm from scratch. Let's try and do this with `KNN`. Write a function entitled `chebychev` that takes in two vectors and outputs the Chebychev or L^∞ distance between said vectors. I will test your function on two vectors below. Then, write a `nearest_neighbors` function that finds the user specified k nearest neighbors according to a user specified distance function (in this case L^∞) to a user specified data point observation.

[1] in class this was the estimated proportion of students having actually cheated

```r
cheby <- function(x,y){max(abs(x-y))}
nearest_neighbors = function(x, obs, k, dist_func){
  dist = apply(x, 1, dist_func, obs) #apply along the rows
  distances = sort(dist)[1:k]
  neighbor_list = which(dist %in% sort(dist)[1:k])
  return(list(neighbor_list, distances))
}


x<- c(3,4,5)
y<-c(7,10,1)
cheby(x,y)
```

```
## [1] 6
```

Finally create a `knn_classifier` function that takes the nearest neighbors specified from the above functions and assigns a class label based on the mode class label within these nearest neighbors. I will then test your functions by finding the five nearest neighbors to the very last observation in the `iris` dataset according to the `chebychev` distance and classifying this function accordingly.

```r
library(class)
df <- data(iris)
knn_classifier = function(x,y){

  groups = table(x[,y])
  pred = groups[groups == max(groups)]
  return(pred)
}


#data less last observation
x = iris[1:(nrow(iris)-1),]
#observation to be classified
obs = iris[nrow(iris),]

#find nearest neighbors
ind = nearest_neighbors(x[,1:4], obs[,1:4],5, cheby)[[1]]
as.matrix(x[ind,1:4])
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width
## 71           5.9         3.2          4.8         1.8
## 84           6.0         2.7          5.1         1.6
## 102          5.8         2.7          5.1         1.9
## 127          6.2         2.8          4.8         1.8
## 128          6.1         3.0          4.9         1.8
## 139          6.0         3.0          4.8         1.8
## 143          5.8         2.7          5.1         1.9
```

```
obs[,1:4]
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width
## 150          5.9           3          5.1         1.8
```

```
knn_classifier(x[ind,], 'Species')
```

```
## virginica
##         5
```

```
obs[,'Species']
```

```
## [1] virginica
## Levels: setosa versicolor virginica
```

Interpret this output. Did you get the correct classification? Also, if you specified $K = 5$, why do you have 7 observations included in the output dataframe?

This output shows us the nearest neighbors to the last observation in the dataset according to the cheby function and then predicts the classification of the last observation. I did get the correct classification, with the last observation being virginica. Although I did specify K=5, 7 observations were included in the output dataframe most likely because there were ties in the nearest neighbors of this observation.

Earlier in this unit we learned about Google's DeepMind assisting in the management of acute kidney injury. Assistance in the health care sector is always welcome, particularly if it benefits the well-being of the patient. Even so, algorithmic assistance necessitates the acquisition and retention of sensitive health care data. With this in mind, who should be privy to this sensitive information? In particular, is data transfer allowed if the company managing the software is subsumed? Should the data be made available to insurance companies who could use this to better calibrate their actuarial risk but also deny care? Stake a position and defend it using principles discussed from the class.

Anyone should be privy to the sensitive information included in the retention of sensitive health care data as long as it is proven and necessary to fully benefit the patient and is only used for the purpose of benefiting the patient alone. From both a deontological and virtue ethics perspective, as long as a company only has the intention of bettering those using their services, the use of any sensitive information, assuming informed consent, is moral. Because the company, in this situation, is performing in a selfless way to help others, they would be acting with virtuous intentions and only for the improvement of another's life. If the company managing the software is subsumed, on the other hand, data transfer should only be allowed if it can be proven that the new company has the same intentions with the sensitive information as the original. This logic continues when deciding whether the data should be made available to insurance companies who could use it to better calibrate their actuarial risk. However, with an insurance company, it would only be moral to deny care to someone or raise the price for insurance if the company does not expect to have the resources available to be able to assist without compromising another patient's potential resources. In this situation, a consequentialist approach would be necessary to determine which decision would benefit the most people.