Hw 7

Sofia Zhang

12/3/2024

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 $\hat{P} = 2\hat{\pi} - \frac{1}{2}$ where $\hat{\pi}$ 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 $\hat{\pi}$.

$\hat{\pi}$ equals to the sum of the first category, proportion of people who flipped tail and head so they have to say yes, and the second category, proportion of people who flipped head first and said "Yes" which is a ground truth. For the first category, the proportion equals to $\theta(1-\theta)$. For the second category, the proportion equals to $\theta(\hat{P})$. Sum the two categories up we can get $\hat{\pi} = \theta(1-\theta) + \theta(\hat{P})$. Therefore, $\hat{P} = \frac{\hat{\pi}}{\theta} - (1-\theta)$.

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

When $\theta = \frac{1}{2}$, $\hat{P} = \frac{\hat{\pi}}{\theta} - (1-\theta) = \frac{\hat{\pi}}{\frac{1}{2}} - (1-\frac{1}{2}) = 2\hat{\pi} - \frac{1}{2}$.

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.

```r
#student input
#chebychev function
cheby <- function(a,b){
  max(abs(a-b))
}
```

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

```r
#nearest_neighbors function
nearest_neighbors = function(x, obs, k, dist_fn){
  dist = apply(x, 1, dist_fn, obs)
  distances = sort(dist)[1:k]
  neighbors = which(dist %in% sort(dist)[1:k])
  return(list(neighbors, distances))
}


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

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])
  return(groups[groups == max(groups)])
}

#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])
obs[,1:4]
knn_classifier(x[ind,], 'Species')
obs[,'Species']
```

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?

The KNN classifier assigned the label "virginica" to the given observation, relying on the 5 nearest observations from the x-dataset, based on Chebyshev's distance. This classification is correct. However, although we set $K = 5$, the resulting output dataframe includes 7 observations. This happens because the nearest_neighbor function selects all indices matching the smallest distances. In this case, two additional observations had the same distance as the fifth closest, resulting in 7 observations in the final output.

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.

the company managing the healthcare software should have access to the data, but only in a way that ensures it does not cause harm to the patient. Since the insurance company may possibly deny care, they should not be able to access that.

Firstly, the harm principle suggests that individuals should not be subjected to harm by the actions of others, especially when it involves personal and sensitive information. If healthcare data is transferred or shared in ways that could potentially lead to harm—such as insurance companies using this data to deny care—it directly violates this principle. The possibility of algorithmic models determining someone's eligibility for care based on risk factors could unfairly disadvantage individuals in marginalized groups, particularly if those algorithms are not fully transparent or accountable. Second, patient autonomy and data privacy are key ethical considerations. Every individual has the right to control who has access to their personal health information. Since Informed consent is a critical factor, patients must be fully informed about how their data will be used, who will have access to it, and what the potential consequences could be. This includes understanding that their data might be used for purposes beyond their direct care. Insurance companies might use patient data to deny coverage or make decisions that harm the patient's health and well-being, despite having not given explicit consent for such uses.

The role of the company is primarily to manage, process, and apply data in the service of improving patient outcomes. However, this access must be bounded by strict ethical guidelines and transparency. For example, the company could use the data to refine its algorithms or improve the quality of care provided to individuals, but it must not use the data to make decisions that could result in discrimination, harm, or exploitation. Any use of data should be in strict alignment with the patient's informed consent, ensuring that patients understand and agree to how their data will be utilize.

I have described our responsibility to proper interpretation as an *obligation* or *duty*. How might a Kantian Deontologist defend such a claim?

A Kantian Deontologist would defend the claim that we have an obligation to proper interpretation by grounding it in the core principles of Kantian ethics—particularly the categorical imperative, which emphasizes duty over consequences and the inherent dignity of rational agents. From this perspective, the responsibility to interpret information correctly is not merely a preference or a goal, but a strict moral duty.

First, a Kantian would argue that it is impossible to universalize the practice of advancing claims that cannot be substantiated. If individuals were to regularly make unsupported or misinterpreted claims, it would erode trust in the communication process, making the very act of making claims self-defeating. The failure to provide truthful, accurate interpretations would contradict the moral law, as such actions would undermine the universal value of truth. This aligns with Kant's notion that moral actions must be capable of being consistently universalized without contradiction.

Second, a Kantian would emphasize that failing to interpret properly treats others as mere means to an end, rather than as ends in themselves. Whether the audience consists of fellow researchers or the public, advancing claims without proper substantiation or interpretation uses them for one's own purposes—such as advancing a particular argument or agenda—while disregarding their autonomy and rational capacity. Proper interpretation respects the dignity of others by ensuring that they are engaged in a way that supports their ability to reason and act autonomously. Misinterpreting or distorting information would violate this respect for persons by undermining their ability to make informed decisions.