Hw 7

Pranav Kallem

11/23/2024

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 $\hat{\pi}$.

For a differentially private mechanism for such a coin, the observed proportion of affirmative responses is $\hat{\pi}$. Using the probability θ, the generalization of $\hat{P}$ can be derived by considering two components. First, the probability of answering "yes" due to actual incrimination is given by θP, where θ is the probability of the coin landing heads. Second, the probability of answering "yes" due to the randomization mechanism, irrespective of actual incrimination, is $(1-\theta)(1/2)$. These components together form the basis for estimating $\hat{P}$.

Thus, the observed proportion of affirmative responses $\hat{\pi}$ is given by

$$\hat{\pi} = \theta P + (1-\theta)\frac{1}{2}$$

Which can be rearranged for P

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

This is the generalized formula for the estimated proportion of incriminating observations.

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

When $\theta = 0.5$, the biased coin becomes a fair coin. Substituting $\theta = 0.5$ into the formula from the previous question, we get

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

This result matches the expression derived in class, where $\hat{P} = 2\hat{\pi} - \frac{1}{2}$ for a fair coin. The generalized formula for $\hat{P}$ accounts for the bias in the coin flip mechanism using the parameter θ. For the special case of a fair coin, the formula simplifies to the result discussed in class.

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

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(vec1, vec2) {
  if (length(vec1) != length(vec2)) {
    stop("Vectors must have the same length.")
  }
  max(abs(vec1 - vec2))
}

#nearest_neighbors function
nearest_neighbors <- function(data, target, k, distance_func) {
  if (!is.matrix(data) && !is.data.frame(data)) {
    stop("Data should be a matrix or data frame.")
  }
  if (length(target) != ncol(data)) {
    stop("Target point should have the same dimensions as the data points.")
  }

  # Calculate distances
  distances <- apply(data, 1, function(row) distance_func(row, target))

  # Identify the indices of the k smallest distances
  neighbor_indices <- order(distances)[1:k]

  list(indices = neighbor_indices, distances = distances[neighbor_indices])
}

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

data <- matrix(c(1, 2, 3,
                 4, 5, 6,
                 7, 8, 9),
               nrow = 3, byrow = TRUE)
target <- c(5, 5, 5)
nearest_neighbors(data, target, 2, chebychev)
```

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.

```
library(class)
df <- data(iris)
#student input
knn_classifier <- function(neighbors, class_col) {
  class_labels <- neighbors[, class_col]
  return(names(sort(table(class_labels), decreasing = TRUE))[1])
}

#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, chebychev)[[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 classification result indicates that the observation in question, corresponding to the species virginica, was correctly classified by the KNN algorithm. The predicted class, based on the 5 nearest neighbors, also matched virginica, demonstrating that the classifier performed accurately in this instance.

The output, however, includes 7 observations instead of the expected 5. This discrepancy arises due to ties in the Chebychev distance. When multiple observations share the same distance to the target point, all tied points at the boundary of the k-th nearest neighbor are included in the result. For example, if the 5th closest point shares the same distance as additional observations, these tied observations are also included, resulting in more than k neighbors being returned.

In summary, while the classifier achieved the correct classification, the inclusion of additional points highlights a nuance in handling ties within the distance metric. This behavior can be addressed by breaking ties using additional criteria, random selection, or preprocessing the data to avoid tied distances.

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.

Sensitive healthcare data, such as that used by Google's DeepMind for acute kidney injury management, should be handled with strict ethical oversight. Access should be limited to those directly involved in patient care or research explicitly aimed at improving health outcomes, and only with informed patient consent.

From a deontological perspective, transferring patient data, especially during company acquisitions, violates the duty to respect autonomy and confidentiality unless explicit consent is obtained. Universalizing this principle ensures no individual is treated as a means to an end. Meanwhile, a utilitarian perspective might justify limited data use for societal benefits, but the risk of harm, such as denial of care by insurance companies using the data to assess actuarial risks, outweighs these potential gains.

Thus, I argue that data transfer during company acquisitions should be prohibited without patient consent, and sensitive data should never be made available to entities like insurance companies whose motivations may conflict with patient well-being. Safeguarding patient privacy and trust is paramount and must take precedence over other interests.

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 duty to proper interpretation by emphasizing the principle that we must treat individuals as ends in themselves, never merely as means to an end. Misinterpreting data for personal gain or convenience instrumentalizes others, as it manipulates their ability to make informed decisions based on accurate information. Additionally, proper interpretation aligns with the idea of duty as rooted in respecting the dignity of others. By ensuring honest and accurate representation, we fulfill our moral obligation to respect others as rational agents capable of making decisions based on truthful data. This duty is absolute and does not admit exceptions, as violating it would compromise the ethical foundation of communication and decision-making.