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.

$\hat{P}$ = θ*π + (1-θ)

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

Substituting $\theta = \frac{1}{2}$, we get $\hat{P}$ = 1/2 * π + (1/2), which when we solve for π we get π = 2$\hat{P}$ - 1.

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.

g(x') = ϕ_0 which menas that the model is constant and there is no meaningful relationship between the features and the output.

Relative contribution of feature i = $\phi_i x_i{}'$ whhich can be interpreted as how much the prediction changes compared to the previous expression.

[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.

```
chebychev <- function(vector1, vector2) {
  return(max(abs(vector1 - vector2)))
}

nearest_neighbors <- function(data, point, k, distance_function) {
  distances <- sapply(data, function(other_point) distance_function(point,
other_point))
  nearest_neighbor_indices <- order(distances)[1:k]
  return(nearest_neighbor_indices)
}

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

```
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, 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 prediction is classified as virignica but the actual classification was setosa.

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.

Student Answer