Ryan Langlois

CS 5110 Data Privacy section A

Differentially Private Gradient Descent Writeup

For my final project I chose to implement differentially private gradient descent with a new dataset from the University of California, Irvine. I found a dataset consisting of survey responses from Portuguese high school students, as well as their grades, and attempted to predict whether they had a high or low grade in their class based on their survey responses. The survey consisted of questions about their home life, school experience, and data from their previous classes. It is split into two files – one for a math class and one for a Portuguese language class.

The first step of my solution is loading in the dataset, which was more difficult than anticipated. I needed to transform it by converting the binary columns to ones and zeros and ones and one hot encoding the categorical columns. I also changed the column encoding their academic performance to binary, so the model only needs to predict high or low. After this I mostly followed along with the in-class exercise of the week of October 28th. I trained a model with the built in sklearn functions, then recreated this manually. Once defining a loss function, a gradient function, and an average gradient function, I was able to fully implement gradient descent. The final step was adding the noise. To do this

accurately, I clip the gradients then add noise with a sensitivity of the clipping parameter. Once this was completed for the first dataset, I was able to do the same thing for the second dataset easily because the two have the same format.

I also performed an analysis of the model constructed to examine its accuracy. Without noise the model performs at about 65% accuracy, which is relatively poor considering there are only two categories to guess so an accuracy of 50% could be achieved simply by guessing. This means that it's difficult to determine the academic performance of a student given this data. The model performs worse with noise added and varies with each run due to randomness. The models for the math class and the Portuguese class also perform similarly. Additionally, by examining the coefficients of the models, I can guess as to which factors are most effective in raising a student's grades. Some notable examples include how often they go out with friends, with those who do performing worse, as did those who have failed previous classes or required supplemental support in school. The most important factor, however, was their desire to pursue a higher education. Overall, I'm a bit disappointed with how these models performed, but this is likely due to the small size of the dataset. In the future I'd like to supplement the data or choose a more robust dataset entirely.