

I have finished gathering and processing all the necessary data for my project. I have created some different plots that compare features based on their classification (nonlake or lake) and I have created a pair plot with 16 different features. I have a basic neural network working, but since I need to figure out how many training and iterations to run, the real or optimal accuracy for my model is between 70-90 percent. I am not sure if trials with more training iterations that produce a higher accuracy are overfitting the data or not. I have prepared a cross-validation algorithm to run with the neural network to solve this. I haven't tested this code yet because I encountered some issues last week with the cloud-computing environment in which my work is stored, but I plan on having it tested and working within the week of 11/24-11/30. I will not have time to test multiple neural network architectures, but I am expecting that my simple custom network will provide sufficient results. My time estimates have been fairly accurate, and I am hoping that by the end of this work week, I will have a complete and functioning code. Thus the last step will be to work on the deliverables, including cleaning up the code and improving its modularity.