# ADS2002- Individual Portfolio

Week 2

This week we first investigated google colab in which we will be conducting most of our coding. It is relatively similar to the Jupyter notebooks with which we have been working for the last few semesters. It seems slightly more difficult to actually access the dataframes as they need to be reuploaded into colab each time, however it is easier to integrate with github and so lends itself to more group collaboration which will be useful with the project. We also had our first introduction to our group members and fully explored the full project outline. We were tasked with coming up with meaningful research questions which we could ask our group mentor, Simon, in next week's class. We have explored the project introduction and basic overview of the task and were overall just brainstorming as a group possible approaches to the final task.

Week 3

This week we investigated the use of github, committing changes and pulling branches. Hopefully this will be useful further down the line with the group project work as in previous semesters projects collating all of the programming work at the end of the semester from a variety of jupyter notebooks has been difficult. Simon came into this week's class to clarify and give real details about the projects we will be undertaking. He explained the nature of the data as well as giving some recommendations for how to work with it. It will be difficult due to the size of the image data being so much larger than anything we have had to work with in the past but hopefully this leads to more learning and useful skills being developed. We decided to do some simple exploratory data analysis and form a basic understanding of the types of catheters involved as there are some technical terms which we were unfamiliar with. Following on from this, the only changes made to our group github were the creation of our own individual branches within our collaborative repository as well as the uploading of some trial programs in which we were just familiarising ourselves with the use of google colab and github.

Week 4

This week we were introduced to the concept of gradient descent. We viewed a video which was pretty informative and it seems to be how most of the regression models we have been using function. Following this, we finally managed to import the proper train dataset into colab to begin having a basic browse over the data and continue with further exploratory data analysis. There doesn't seem to be too much information within this dataset, just the type of catheter and whether or not it is present. However, the real challenge seems to lie within accessing the full array of image data within colab which is something we haven't had to deal with previously. We further analysed the train dataset and it became evident that most of the catheters were in the CVC - Normal definition. As a group we have also been trying to gain a better understanding of programs such as tensorflow and neural networks which may be useful in our image analysis later on but at this phase they seem a little out of scope of our current ability. We are also having difficulty accessing the full image dataset but we have communicated with Simon and this hopefully be fixed soon.
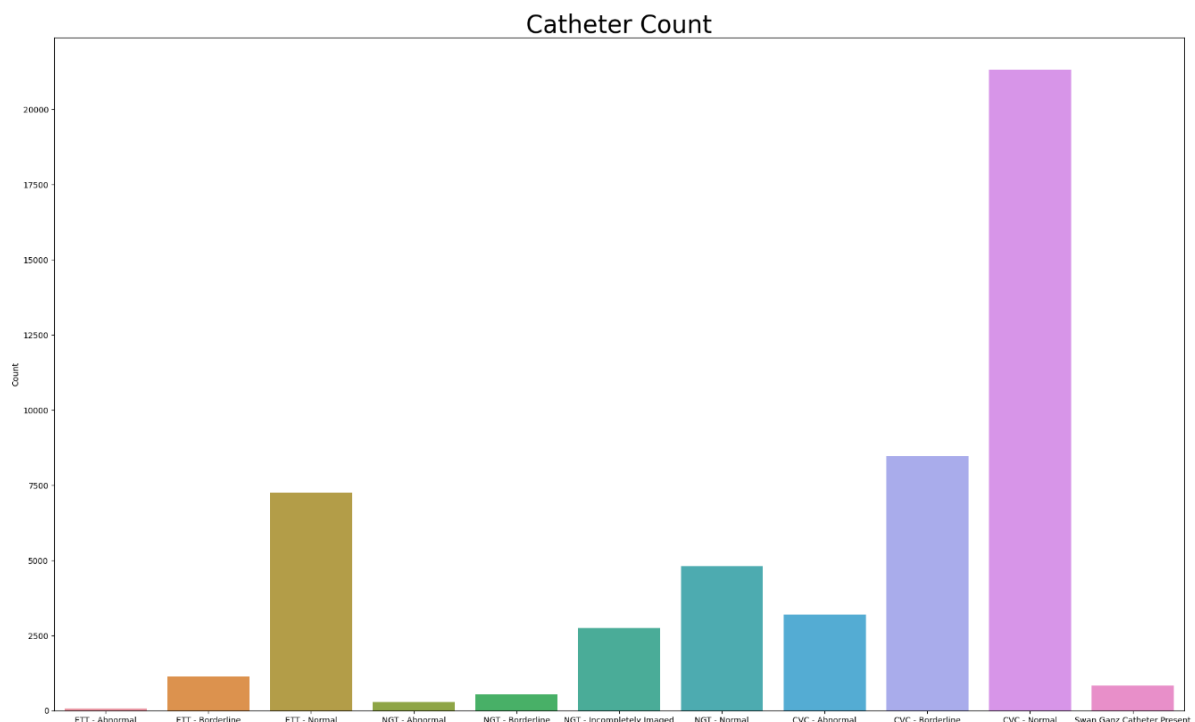
Week 5

This week through the notebook we were investigating the feature selection and how it can be implemented. While this was interesting to know it doesn't seem as relevant to our project which is more focused towards image analysis for the different types of catheters. Moving on, we can now fully access all of the images and I figured out how to load them into google colab and plot them as well. I attempted to implement some PCA on the images which we have been provided at the recommendation of my group members. Ideally this process would reduce the dimension of each of the images making them easier to process as there is a very large amount of data we need to work with. It proved difficult however to load an entire folder of images into the notebook and I am still researching how this can be done the way we need. Ultimately, I managed to get some reduction working on a single x-ray image but it also raises concerns of how accurate our model will end up being with this reduced detail as the actual catheters are already pretty difficult to distinguish in the x-rays.
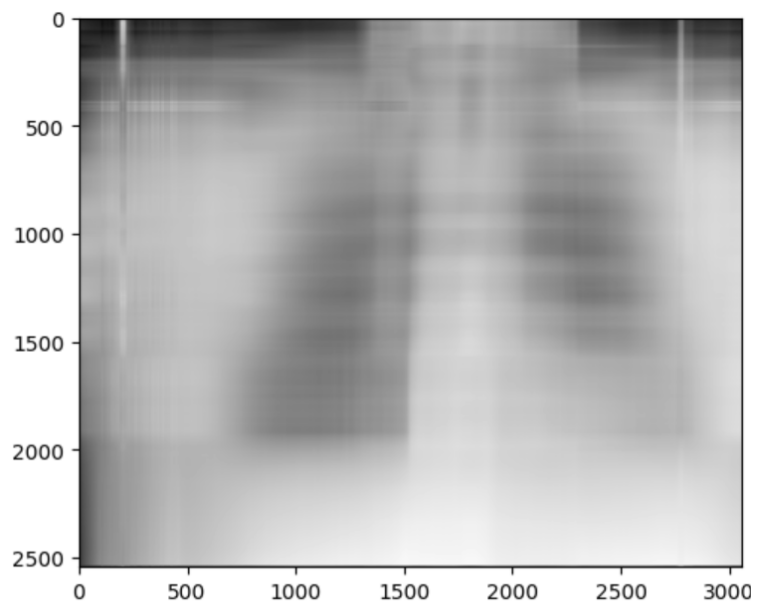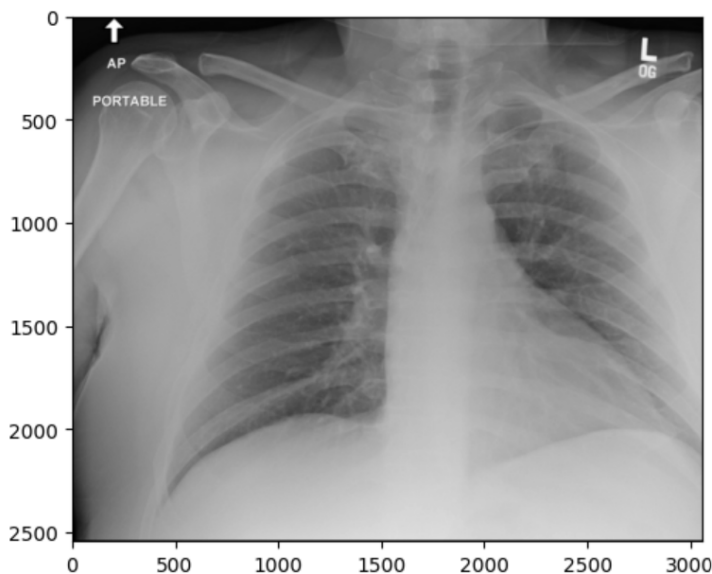
Week 6

This week we spend the first half of the class time working on the imputation activity for the abalone dataset. In terms of our group project we don't have any values in the dataset that need imputation however it may be useful further in the future. In regards to the group project I began attempting to implement a k-means classification model on some of the images from the dataset. This is proving to be difficult as when researching online the majority of examples appear to be classifying sets of images with only two options, such as dog or a cat. In the data we have been provided there are eleven different types of catheter which can be found, oftentimes multiple of them being within the same image. In order to simplify the starting of this process I think it would be appropriate to create a subset of the data to work on, only trying to decide whether a catheter is CVC - Normal or not. Additionally, all of the x-ray images are not of the same dimension which means they don't work with the kmeans model so this is another issue that needs to be corrected.

1st Reflection

So far, working on this catheter group project has presented quite a challenge both in terms of the physical work as well as the group dynamics. It has been challenging managing people's expectations of one another as well the timeframe for the entire project. While it was stressful to begin with, I believe that we have developed more of an understanding of one another to produce a less stressful environment more conducive to success. Additionally, we faced challenges when two of our original group members left the class altering the team interactions and workload required. Overall, I have mostly been focused on some of the basic image analysis, like kmeans, which will hopefully provide a solid foundation upon which we will be able to develop our project and more complex image analysis models. I believe that I could have better implemented github into more work at the start of the project so that all group members are working from the same initial point, preventing issues later on. Overall, I believe we are on track to understanding and developing successful outcomes for this project.

Week 7

This week we first investigated the use of perceptrons. It was interesting to finally be working with something we haven't encountered before as most of the previous concepts I was at least somewhat familiar with, although not completely fresh in my memory. Continuing on from this, I was further attempting to implement a kmeans model on the subset of images that i chose last week. I finally had some success with importing all of these images from their folder on google drive into colab. The command I was using also had the added feature of being able to resize the images which is useful. As well as this we also needed to make sure that the images were brought in grayscale so that there was not a third dimension of RGB value which was negatively affecting the model. It was still proving difficult to actually get the kmeans model up and running however but there is still time to fix these issues. It will be a difficult week to work further on the group project as there are multiple mid semester exams as well as the essay due this week so I don't expect too much group work to be conducted.

Week 8

Week 9

Week 10

Week 11

Week 12