Name: Jason Zhou
Mentor: Dr. Dongjin Song
Status Report #: 24
Time Spent on Research This Week: 7.5
Cumulative Time Spent on Research: 184.25
Miles Traveled to/from Mentor This Week: 0
Cumulative Miles Traveled to/from Mentor: 0
===

Monday, March 28th, 2022: (0.5 Hrs)

To start the week, I had my weekly meeting with my mentor. I started off by sending him a PDF file of my research article draft (intro, methods, and references) and asking him to look it over when he had the chance.

Afterward, we talked about the performance of my neural network. Specifically, last week, I had been having problems with the accuracy of my neural network which had only managed to predict around 50% of the samples correctly. He explained to me that for these types of tasks, accuracy usually is not the best evaluation metric to measure the performance of a neural network. Instead of accuracy, he suggested that I try AUC (area under the curve) and pAUC (partial area under the curve) to evaluate the model. Although I had a rough idea of what AUC and pAUC were from my previous studies earlier in the year, he later sent me a slideshow that explained these concepts in further depth.

Besides meeting with my mentor, I also went to the Reading and Writing Center to have my paper checked. Although a gave a couple of grammar suggestions, they said it looked fine and gave me a completion mark.

Wednesday, March 30th, 2022: (1 Hrs)

On Wednesday, Ms. Pintavalle gave my class time to work on our papers. As such, I spent the first half of class reading through the suggestions that the staff at the Reading and Writing Center had given me and made some changes to my paper.

I spent the second half of class looking through the evaluation metrics slideshow that my mentor had sent me the other day. Although AUC is similar to accuracy, they are not quite the same. For example, instead of saying that something is right, it is broken down into two labels: true positive[1] (TP) and true negative (TN). Conversely, if something is wrong, it can be broken

---

[1]In the context of my project, if the model predicts positive, this means that it predicted a normal sound. If it predicted a negative, this means it predicts an abnormal sound. The word "true" simply means that the model's prediction was correct.

down into false positive[2] (FP) and false negative (FN). In order to solidify these concepts in my head, I found a useful graphic:
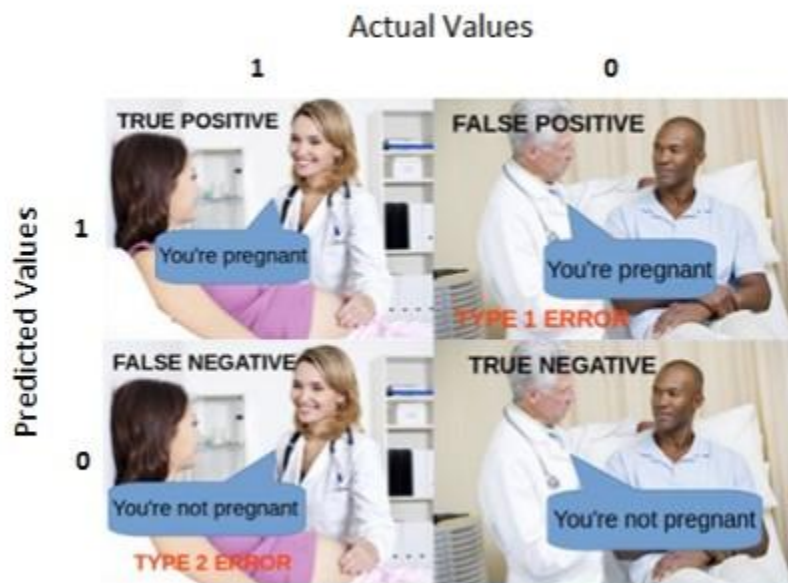
(This is an image outlining the concept of TP, TN, FP, and FN. In this case, being pregnant is positive, and vice versa.)

Friday, April 1st, 2022: (2 Hr)

Taking a break from AUC and pAUC, on this day, I investigated the abnormally high loss values that I was getting from the training of my model. I expected the loss values for normal samples to be low, but they were actually very high (which means that they are reconstructed poorly). My first thought was that the model was overfitted[3].

Thus, I retrained the model on fewer epochs (scaling from 100 to 50, then from 50 to 25, and so on). However, after several attempts at lowering the epochs and measuring the loss, it seems to have no noticeable effects on the unexpectedly high loss values I was getting.

---

[2] If it is a false positive, this means that the model thought the data was normal but its prediction was wrong. If it is a false negative, this means that the model thought the data was abnormal but its prediction was wrong.

[3] This is a term that refers to when a model is over trained. Essentially, this means that it gets so good at predicting a certain dataset that it becomes worse at predicting other, new data. Although it is not quite the same, one could view this as a student who puts all of their time into studying math. Because they've spent so much time training to be good at math, they lack proficiency in other subjects such as science, English, or history.
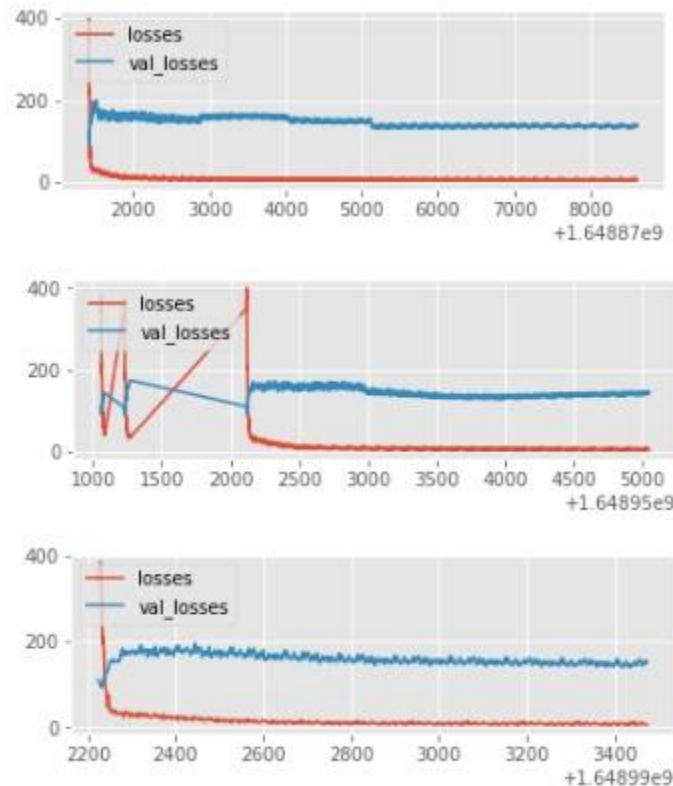
(This is multiple graphs of the loss values for each model I trained. Each with decreasing training time, and fewer epochs. What is important in this image is the blue line. Notice that it seems to stay relatively constant across all the graphs, regardless of its appearance in the beginning. I could lower the epochs even further; however, based on these trends, it would likely still result in loss values that are too high for real-world use)

Beyond the loss graphs, I worked on implementing the AUC evaluation metrics that my mentor suggested I use. After some research, I discovered that a library called SKLearn had the functionality to do this; however, I would need to modify my code in order to give the right inputs that the SKLearn function needed to construct the AUC-ROC[4] curve. Thus, I spent the rest of my time doing this.

Sunday, April 3rd, 2022: (4 Hrs)

On the last day of the week, I continued to work on implementing a successful AUC-ROC curve to model the performance of my model. Speaking honestly, most of my time was spend trying to comb through compile time and runtime errors while coding. I had a lot of trouble successfully converting from data type to data type and also had difficulty working with multi-dimensional arrays to prepare the data.

---

[4] AUC-ROC is a graph that models performance. AUC is the actual measurement or numerical score of the performance.

However, after a couple of hours, I finally got everything working properly. Here is the result:
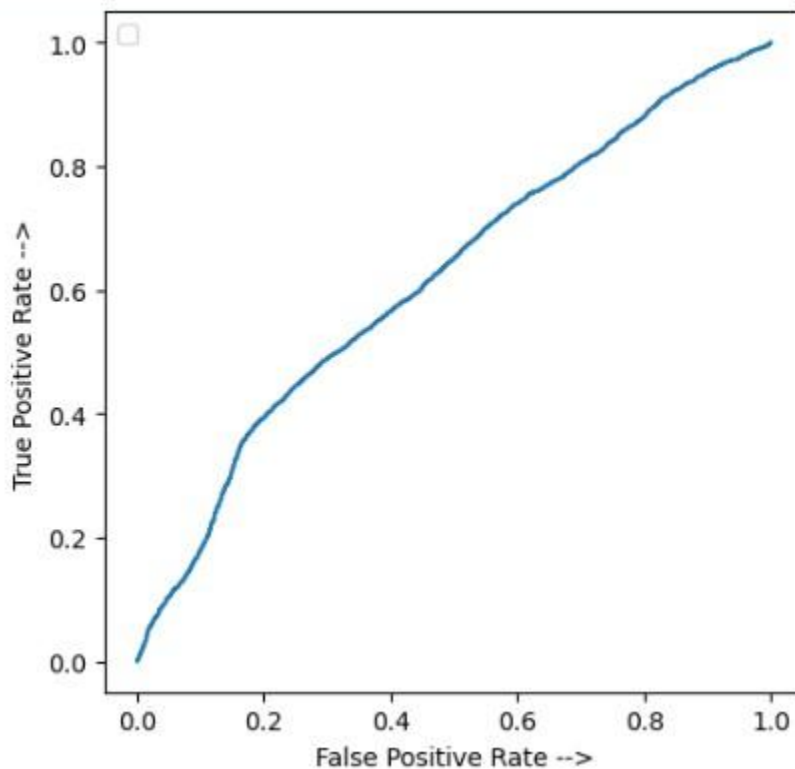
(A graph of the AUC-ROC curve I was talking about above. For reference, the closer the line is to y = x, a straight line, the worse the score is. As such, this can be considered a bad AUC score)

Because I am not getting good results, I am going to try to change the architecture of my model to hopefully improve its performance.

References

Sklearn.metrics.roc_curve. (n.d.). Scikit learn. Retrieved April 4, 2022, from

https://scikit-learn.org/stable/modules/generated/sklearn.metrics.roc_curve.html