Self and Peer Review

By: Ashna Ali

Team Name: Data Divers

9/23/2024

Self-Review

As a team lead, I think I did a good job of directing our team on which tasks they needed to accomplish. However, I do think I can improve on setting strict deadlines on when things need to get done. I am someone who tries to get tasks done in a timely manner, sometimes much earlier than the deadline. But I forget that everyone works at a different pace. So assigning better deadlines outside of the deadlines the professor gives will be a good solution as well as communicating clearly on what the tasks are and who is assigned to what task. Another comment I would like to add is that I am not sure doing all the assignments individually was a smart decision. I would like to work with my team more to create a more cohesive solution to our problem because in my experience with this assignment, everyone had different ideas on how the frontend was going to look like, the backend, as well as the model topology which made it difficult to make sure everyone contributes to a similar product. If there are three versions of the app that all look different that people did individually, how can we report on which part we worked on in the final submission, if that final submission was only one person’s version? It does not make sense to me. In the future, we are going to have more meetings to discuss ideas and make sure we come to a consensus on how the app is supposed to look like instead of having these inconsistencies.

Overall, though, I think I have learned a lot on the technical side of things. I contributed to the model topology and training and fine-tuning. It was a long and dueling process of gathering several datasets, preprocessing them, and trying different topologies to see which one would be the best to classify different types of diabetes. Originally, what I was going to do was take my assignment2\_2’s submission, combine all my datasets into one, and then train the model on that. However, I realized there were a lot of inconsistencies with data including missing data and not having the same amount/types of features between the datasets. So, my next solution is to use the concept of transfer learning to further improve and expose my model to different types of data. I used datasets from Kaggle, Github, and other sources like from the CDC to train my model on different features such as gender, A1C, glucose, bmi, etc. I also performed feature importance to see which features are contributing more to each prediction. There were features such as smoking and alcohol history, but those were quickly eliminated as possible features to train because they did not contribute much. The overall accuracy after all the training and testing was around 74%.

When it came to time series data, I initially trained the model fine-tuned after the processing I mentioned above, but that did not result in reasonable convergence or accuracy. I had to do a bit more research as well to find that sort of data but also on a new solution to solve this problem. I came across LSTM(s) or Long Short Term Memory layers that actually have a memory cell attached to them. What this does is allow the neural network to remember past inputs, which in time series data, is needed to calculate trends. So I decided to implement that and I got a much better result with around 95% accuracy without the model overfitting. I quickly realized in order to use both LSTM(s) and RNNs, I had to create an ensemble model that handles both structured data and time series data simultaneously whose outputs concatenate to get the final prediction. In the end, I got about 83% accuracy overall with the RNN attached to these LSTM layers for time series prediction.

Sreevardhan Reddy Soma: Review

When it comes to the front-end development, Sreevardhan did a good job of developing a simple web interface for our project, which is essential for user compliance. He suggested the frontend I did like the question/survey-based system that we have with our app. He did a moderate job of communicating clearly on when he will get tasks done and what he contributed to. However, like I have said before, since these previous assignments were individual, I wasn’t sure exactly what his model looked like until we met in person. He also didn’t incorporate time series data into his model which isn’t exactly what we wanted. But going forward, he will improve on his work.

Venkata Sai Veeramalla: Review

During our research phase of the project, I appreciated his contributions to finding specific datasets to train our models, which I used some of them in my training process. He found datasets outside of Kaggle and Github which is going above and beyond my expectations on research. He and Sai worked together on the assignments of 2\_1 and 2\_2 and created a model with a combined dataset, which they managed to get working properly without issues. I remember experiencing some difficulty in doing that, so I am glad they were able to do it. He did a good job on communicating how far along he was at accomplishing the tasks and was able to ask questions when needed.

Sai Kiran Basetty: Review

Sai and Venkata worked together on their version of the app and model. Their contributions were basically the same as each other. I would like to add that Sai did a better job at asking questions throughout to make sure he was doing the process of building, training, and testing the model correctly. However, I do feel like he might have benefitted from his own research of finding solutions to that problem.