EE/CprE/SE 492 WEEKLY REPORT 3

Feb 11 - Feb 24

Team 48

Allergy Prediction Using Artificial Intelligence

Client/Advisor: Ashraf Gaffar

Client Lead: Joseph Trembley

Team Lead: Noah Ross

Minute Taker: Ella Godfrey

Research Lead: Xerxes Tarman

Quality Assurance Lead: Alex Ong

- Weekly Summary
  We created the model and used various techniques to validate and evaluate the predictions such as k-fold validation, accuracy, precision and recall. We also met multiple times to ensure everyone could touch base on what they had been working on, and get a baseline for future work.
- Past week accomplishments
  - Joseph Trembley - wrote a base for the model, experimented with different model hyperparameters and metrics to get a better method for creating predictions and evaluating the quality of them.
  - Noah Ross - Using the base model Joseph implemented, I experimented with different layers and activation functions to see how they impact the area under the curve, precision, and recall.
  - Ella Godfrey - With the base model I worked on addressing the issue we were faced with where the area under the curve is not meeting the expectations we are wanting to see. This entailed experimenting with class weights to see if changing that could significantly help the auc.
  - Xerxes Tarman -These past weeks, I got up to speed on the Keras model that Joseph has been working on. I implemented K-fold analysis and conducted additional testing on the python notebook. I also researched different metrics that we can apply to our project. Additionally, I tested using dropout weights, Support Vector Classification (SVC), and random forest models on our dataset.
  - Alex Ong - Did research on understanding the models from keras and the effect of each hyper parameter. Did more research on data pre-processing and one hot encoding method to double check if we were handling that correctly before porting it to the model. Small experimentations and changes on the existing model.

- Pending issues
  - Joseph Trembley - need to determine how to further refine the model to get better predictions consistently.
  - Noah Ross - Recall of the model is low and seems to be unrelated to layers. Need to look at how data is processed
  - Ella Godfrey - need to continue to experiment with the model to try and get the model to have better predictions
  - Xerxes Tarman - All tested models show very low performance, even when trimming the data to its most popular ingredients. This leads me to believe that the pre-processing of the data needs to be corrected or there is no pattern to be learned from the data.
  - Alex Ong - Continue to get a better understanding of each hyper parameter effect and explore other processes to create a more sophisticated model
- Individual contributions

Name	Hours this week	Hours Cumulative
Joseph Trembley	13	54
Noah Ross	6	41
Ella Godfrey	7	35
Xerxes Tarman	8	40
Alex Ong	6	38

- Plans for the upcoming week
  - Joseph Trembley - refine the model further to get better predictions.
  - Noah Ross - Need to look at how data is processed to try and increase recall
  - Ella Godfrey - continue working with the model to obtain better predictions
  - Xerxes Tarman - Verify that data pre-processing is correct. If model predictions do not significantly improve, possibly reach out to the advisor about the quality of the data.
  - Alex Ong - continue working on the model, begin more refinement of the front end if significant progress is made on the model
- Summary of weekly advisor meeting
  - We discussed our current progress that we have made in the project, which was that we have created our model and are attempting to refine it further to make it better at predicting allergies.
- Broader Context
  - In our design from last semester, we had included a database as a possibility for the data to go in as it enters the model. However, after discussion with our advisor, we realized that

this would affect our global design context by violating HIPAA, so we removed the component from the design.