

## Sprint 5 Plan

Understanding Healthcare Data  
Sprint Completion Date: 6/7/2020

Goal: Introduce an RNN into the Docker container for Rob to run on actual patient data. In addition, zero out the first 12 months in the latest Synthea training/testing set to compare the effectiveness of our RNN and our CNN.

### User Stories

**User Story 1:** *As a programmer, I want more predictive models at my disposal.*

**Task 1:** Add the RNN to the Docker container. (5 points)

- Add RNN model to w2vtrain.py and push the new docker container to docker hub.

Total for user story: (5 story points)

**User Story 2:** *As a programmer, I want to compare the effectiveness of my RNN model vs. my CNN model on Synthea generated data.*

**Task 2:** Generate and train on a dataset with the oldest 12 months zeroed out to compare the effectiveness of our RNN vs. our CNN. (10 points)

- Modify our one-hot encoding and word embedding dataset by zeroing out the first 12 months in our 24 month window.
- Test and train on the modified embeddings.

Total for user story: (10 story points)

**User Story 3:** *As a group, we want to be able to present our project in the best light to the audience*

- **Task 3:** Work on our poster and submit the final draft. (10 points)
- **Task 4:** Assign slides in the powerpoint and rehearse them before the presentation. (10 points)

Total for user story: (20 story points)

### Initial Task Assignment:

Cassidy Norfleet - task 2, 3, 4

Brendan Reilly-Langer task 1, 3, 4

Aman Prasad - task 1, 3, 4

Harshitha Arul Murugan - task 1, 3, 4