# Project 14 - Comprehensive AI Architecture

## Report Week: 8/30 to 9/10

## Progress summary

* Most of the raw data has been parsed and reviewed
* Programming the pull of geographical data like city, state, and zip code information has been written
* Programs have been created to edit the information far faster
* Discussions have been had about team meetings and the frequency of them
* Team meetings have been set up and organized
* Sever access has been granted and team members are setting up logins
* Server setup, logins, and server updates are being assessed and performed
* Created website

## Meeting summary

* Our sponsor gave us an overview of the project with a copy of their PowerPoint presentation full of useful links and an overview of what is needed to complete Milestone 1.
* Everyone needs to setup the VPN on their local machines. Everyone needs to establish a connection to the server provided by the sponsor. Everyone needs to work on getting Jupyter Notebook connected to the server. Look over the raw data to get a better understanding of what we are working on. Use the captured geographical data to work on programming the retrieval of weather data.
* We meet frequently via Teams and GroupMe. These are meetings where we discuss everything from what is left to be done on documentation and share updates as we make progress working with the heart rate monitor data. We use our time to share ideas and collaborate to come up with solutions as problems arise. Some issues are going to be addressed as we meet with our sponsor. If an issue, like getting the server setup for example, arises that is slowing down progress, Talia reaches out to the sponsor for guidance.
* Anytime documentation is due, we will spend time meeting to discuss what has been done in the documentation and what is left to be done. If a member can’t attend, we reach out to them via email and ask that they complete key pieces.

## Key events

* City, state, and zip code have been programmed in and we have found limitation with gathering that data in bulk. That will be addressed in our next sponsor meeting
* Weather data is being tested and programmed in new server

## Member activities

|  |  |  |
| --- | --- | --- |
| **Team member name** | **Major tasks and contributions** | **Workload (hours)** |
| Talia Brooks | Communicated with sponsor about setting up meeting times and questions  Set up Gantt chart and edited charts per team's requests and accommodations  Organized and submitted first assignment for team  Edited and created the weekly report | 15 |
| Aaron Bemis | Parsing out the raw data. Initially this was done on my PC working with a few new tools, so there was some learning curve there as well.  Helping with documentation for the project plan, weekly activities etc.  Our sponsor has a server setup so that we can share our code and data.  Once that was established, I set up a VPN on my local machine to allow me to connect to the server. Connected Jupyter Notebook to the server to allow me to move data and code to the server easily.  Updated a bunch of Python packages that were outdated on the server so that Python code could run correctly  Everything has been migrated to the server.  Working on a more efficient way to capture geographical data. There is a limitation on APIs that we will address when we meet with our sponsor next week. The initial code for that is built and working, it is just challenging to pull that data in bulk. | 38 |
| Jack Morris | Created Website.  Becoming familiar with Pandas Python Library and reviewing Python.  Filled out some of the project plan. | 20 |
| Alex Boyett | Worked to pull geographical data from raw data parsed by Aaron  Researched how to pull weather data using coordinates from Meteostat (https://github.com/meteostat/meteostat-python)    Connected to the server via PuTTy and then Jupyter notebook, which I will need to become more familiar with to work around in it. Attempted to update packages but received an admin access error. Will need to ask the sponsor about our level of access in the next meeting. | 15 |
| Lauren Bailey | Worked on creating the actual database that we will use. Aaron created two tables for it already, and I worked on adding to it.  Did my part for the project plan.  Researched databases that similar AI use. Trying to make sure we are doing exactly what would work best for our specific AI. | 15 |
|  | ***Team Total*** | 100 |

## \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

## Report Week: 9/11 to 9/17

## Progress summary

* All the raw data is now fully parsed into a single dataset on the server where the modeling will take place.
* Meeting with our sponsor produced solid results from the team this week.
* Jack took a deep dive into a Pandas which is a library used by Python that helps read, query, export and in our case combine datasets.
* Weather data has been acquired for all the workout data. There are some exercises that did not produce information and Professor Le let us know we can address that in our models.
* Pulling geographic data has been challenging and was fully collected this week.

## Meeting summary

## Our sponsor, Professor Le, met with us on Monday, September 12th. This meeting covered technical aspects of Milestone 1 and provided clarity to a few components. Professor Le gave us details on how he would like the first batch of data. He needs sample workout data for around 1,000 exercises.

* The team met several times to determine the best way to pull the geographical and weather data. Through these discussions some optimizations have been made.
* There was another team meeting to help decide the best way to combine the data, produce sufficient results for Professor Le’s first dataset and how to output that data.
* We had a team meeting to discuss the requirements for Milestone 1 documentation. We are going to compile all the data for that presentation and start getting the PowerPoint ready.

## Key events

* Clarification on the exact deliverable for Milestone 1.
* All the weather data has been gathered and combined back with the raw heart rate monitor data.
* The geographical data has been split into 10 jobs that are running against separate datasets to continue pulling the geographical data.

## Member activities

|  |  |  |
| --- | --- | --- |
| **Team member name** | **Major tasks and contributions** | **Workload (hours)** |
| Talia Brooks | Met with our sponsor, Professor Le. Worked on gathering data for the Milestone 1 PowerPoint, as well as creating the rough draft and editing of the milestone 1 report. | 25 |
| Aaron Bemis | Met with our sponsor, Professor Le. Spent almost the entire week continuing to run and monitor the jobs that pull geographical data. Professor Le, offered a few helpful ideas to see if I could speed that process up. I have implemented those and pulling the data seems much more stable. Helped organize the data in a way that Jack could start working on tying all the pieces of raw data, weather data and geographical data together. Updated the Gantt chart. Most of my time this week was spent monitoring and keeping the geographical jobs running. The collection of the geographical data was finalized on Friday the 16th. | 55 |
| Jack Morris | Becoming more familiar with Pandas documentation. Compile all the raw data, weather data and geographical data. Working on pulling 1,000 exercises as sample data for Professor Le. Working on the PowerPoint for Milestone 1. | 25 |
| Alex Boyett | Created program using meteostat to pull weather data from coordinates, which was then used by Aaron to pull weather data using existing data.  Began troubleshooting points where data was not returning from program.  Began work on ways to combine data to one CSV file for sponsor. | 25 |
| Lauren Bailey | Continued to research ways to make our database more efficient. Once I discovered a few different ways that could work best for all of our data, I began to code. I created a few different codes that may create the best database. I then decided on one and added to that. | 25 |
|  | ***Team Total*** | 155 |

## \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

## Report Week: 9/26 to 10/02

## Progress summary

Working on processing our combined data to get into a usable format for the models. This includes tasks like normalizing exercises. For instance, we take all the different types of bike rides and combine them into their own biking category. We have been doing research on the model types since we have been presented with four. Learning about the parameters, their settings and how changes will affect accuracy.

## Meeting summary

We met with our sponsor, Professor Le, on Monday to discuss the next steps to begin working on Milestone 2. He had a PowerPoint presenting, processing code and some explanation of the four models we will use to begin our machine learning. It was a short meeting, but he provided us will all the materials we need to begin the next phase.

## Key events

Established and moved processing code to our server so that we can share those resources. Reviewed and hosted our modeling code on the server as well. We are all busy doing research on the models, run time and desired results.

## Member activities

|  |  |  |
| --- | --- | --- |
| **Team member name** | **Major tasks and contributions** | **Workload (hours)** |
| Aaron Bemis | Met with our project sponsor to get an understanding on the requirements for Milestone 2. He has provided us with some code to help process our collected data and setup 4 models for us to test out. I have moved all that code to the server we have setup so that we can share our progress. Worked with Jack to make sure TensorFlow has been updated on the server. | 25 |
| Talia Brook | Meet with project sponsor as well as did research on required models and ran mini tests to see how to models were supposed to form and lay out | 15 |
| Alex Boyett | Researched the 4 model types for the start of this milestone. Reviewed material provided by sponsor. Researched how to add the data gathered from milestone one for modeling. | 10 |
| Lauren Bailey | Met with project sponsor to discuss how to go about this next Milestone. Researched the 4 different model types to get a better understanding of them. | 10 |
| Jack Morris | Communicated with Sponsor to evaluate our data for modeling and made sure we had the correct modeling code to be able to run the modeling tests. Uploaded the corrected modeling code that should work with TensorFlow. | 10 |
|  | ***Team Total*** | 70 |

## \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

## Report Week: 10/03 to 10/09

## Progress summary

A Python job was provided by our sponsor to get our combined data ready for our four learning models. This job was ran against a smaller set of data containing 1,000 rows to ensure things are working. From there, we tested out the four models to make sure those are working as well. We spent some time this week learning about the results the models produce and their accuracy based on those results.

## Meeting summary

This week’s meeting was mostly used to explain what is working, what models need to be ran, how to deal with the results and we also discussed any questions we have for our upcoming sponsor meeting on 10/10.

## Key events

We have working processed data that is now in an object that our predictive models can use to learn about our data and make heart rate predictions based on its learning. The models have all been tested and we have a much better understanding of what the results are showing us. As the models make predictions, they are also check what the real results are to see how close or far off the predictions are. This will be useful as we work toward our final reports.

## Member activities

|  |  |  |
| --- | --- | --- |
| **Team member name** | **Major tasks and contributions** | **Workload (hours)** |
| Lauren Bailey | Ensured that what out sponsor gave us was enough information to get us where we need to be. Worked on researching types. Evaluating the models we were provided. | 20 |
| Talia BRooks | Editing our next presentation, setting up meetings with the sponsor, evaluating basic models and communicating with the others what parameters should be for the models | 25 |
| Aaron Bemis | Made sure the job that will process our combined data will work. Did a test batch of 1,000 records, produced an object that can then be used by our models. I ran each of the four models to make sure the object was created correctly. Worked with Jack to start running the models and getting an understanding of what the results mean and how they are useful. | 45 |
| Jack Morris | Reviewing the modeling code so it might conform to our requirements. Running GRU model with certain file data multiple times to get an average error rate. | 30 |
| Alex Boyett | Ran test model to confirm how to work modeling. Reviewed the current code to check current state. Verified current required parameters for sponsor with team | 15 |
|  | ***Team Total*** |  |

## \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

## Report Week: 10/10 to 10/16

## Progress summary

The models for milestone 2 have been ran, the data has been charted in an Excel document. We now need to do some final summaries of the data and prepare that as the deliverable for milestone 2.

## Meeting summary

Talia and Aaron met with our sponsor to get clarification on the format for the milestone 2 deliverables. He provided an example of how to display the run parameters and to average the output of 5 run per set of parameters.

## Key events

All of the models have been ran with 4 sets of parameters, 5 times per set. Work has already began on building out the next part of the report draft for milestone 2 and the PowerPoint is in progress for the presentation.

## Member activities

|  |  |  |
| --- | --- | --- |
| **Team member name** | **Major tasks and contributions** | **Workload (hours)** |
| Lauren Bailey | Contributed to the Milestone 2 PowerPoint presentation and essay. Ensured that the reports for this milestone have been created and worked on. Put in work for next week since I will be out of the country. | 20 |
| Aaron Bemis | Met with our sponsor on Monday for clarification on a few milestone 2 deliverables. Ran the model for 1DCNN 20 times with 4 sets of parameters, 5 times with each set to get an average of those 5 runs.  Ran the model for LSTM 20 times with 4 sets of parameters, 5 times with each set to get an average of those 5 runs. | 65 |
| Talia Brooks | Worked on the essay as well as the power point according to the new parameters set by professor, creating skeleton and analyzing models to ensure their fit. | 25 |
| Jack Morris | Ran GRU and Attention model types 5 times each to get average of each mode type. | 55 |
| Alex Boyett | Worked on getting lager data set from existing data. Ran test models to compare to other team member models | 15 |
|  | ***Team Total*** |  |

## Report Week: 10/24 to 10/30

## Progress summary

## Meeting summary

…

## Key events

…

## Member activities

|  |  |  |
| --- | --- | --- |
| **Team member name** | **Major tasks and contributions** | **Workload (hours)** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  | ***Team Total*** |  |

## \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

## Report Week:

(Keep adding more future weeks here …)