

California State University of San Marcos

Jordan Fisher, Juan Gonzalez, Keith Gross, Miguel Morales , Noah Miera

CIS490 : Dr. Shaun-inn Wu

### Table of Contents

1. **Application development**

1.1 Statement of Business Context 4

1.2 Statement of Customer’s Business Problem 4

1.3 Statement of Project Proposal 4

1.4 Statement of Deliverables 4-5

1.5 Measures of Success 5-6

1.6 System Overview 6-7

1. **Requirements Matrix** 8
2. **Project Management**

3.1 Product Breakdown Structure 9

3.2 Work Breakdown Structure 10

3.3 Initial Schedule for Tasks and Deliverables 10-13

3.4 Statement of Total Price 13-14

3.5 Gantt Chart 14-18

3.6 Initial Project Cost Tracking Chart 19

3.7 Statement of Deliverables 20

3.8 Outline of Resources Needed 20-21

1. **Team - Appendix**

4.1 Access to Project 21

4.2 Attendance Records 21

4.3 Timesheets 22-26

4.4 Sprint Meeting Minutes 27-31

To: Jared Macshane, Machine Learning Supervisor

CC: Dr. Shaun-inn Wu, Director of Projects

The Sushi Team is grateful to be able to be part of this for Keep America Beautiful. Thank you Jared, as well as Dr. Wesley Schultz, for not only this opportunity to work in the Machine Learning field for this project, but also to work on something that helps build and maintain clean, green, and beautiful spaces in America. We look forward to working and learning from you this Spring semester of 2022, in order to assist in the making of this project.

In this third phase, the team began the construction of a system to receive, process, and detect litter in Google Street images from a user, for our Machine Learning project for Keep America Beautiful. Such progress so far has led us to creating the boneworks of our system through the utilization of a Google Cloud server, the integration of the the YOLOR object detection model - with training and testing to detect litter in Google Street images, and the development of a Python script to formulate the process of our system - pertaining to inputs, outputs, and it’s ability to do frequent operations of processing Google Street images. In the coming phase, we will strive to delve deeper into the development of our Machine Learning algorithm’s results, refining its ability to detect litter in Google Street images, and possibly categorize litter, as well as finally establish a hook by which the Web App team can use the Machine Learning algorithm for their Google Street images.

Regarding the projected cost of this project, the Sushi Team has determined that we will “charge” a rate of $23 per hour, per member of our team. In this third phase, we have accumulated a cost of $5,100.71 through 221.77 labor hours. Furthermore, due to provided discounts & promotions from Google Cloud, we have accrued zero costs from the platform within this phase, during our usage of it. In the next phase, we estimate a cost of $4,979.50, through 216.5 labor hours. With this rate and the currently accrued actual costs of Phases 1, 2, & 3, we estimate the total cost of the project to come to a total of $17,233.21, through a total of 403.77 completed hours and an estimated 345.5 future hours. Should alternative or additional costs be required, the costs will be updated accordingly.

By signing below, you hereby approve Sushi Team to continue working on the following project: Keeping America Beautiful: Litter Detective and agree to the aforementioned estimated costs.

From, Team/Scrum Leader: Noah Miera

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*Jared Macshane, Machine Learning Supervisor*

### 1. Application Development

**1.1 Statement of Business Context**

Keep America Beautiful is a leading national nonprofit organization that inspires and educates people to take action every day to improve and beautify their community environment. They envision a country in which every community is a clean, green, and beautiful place to live.

**1.2 Statement of Customer’s Business Problem**

* Need a new Machine Learning algorithm to produce data on Google Street images containing litter.
* Needs the data to not only detect if there is litter, but also categorize them.
* Make the results available and usable for the Web App team.

**1.3 Statement of Project Proposal**

* 1. Process Google Street images through a pre-trained model.
* 2. Adapt model to our specific needs of litter detection.
* 3. Collect output data/results.
* 4. Sync up input/output format with the Web App team.
* 5. Make results accessible to the Web App team.

**1.4 Statement of Deliverables**

* Machine Learning code/algorithm that will be compatible and usable with the Web App team.
* Algorithm will produce data that will include:
  + Identification of litter in an image.
  + The total amount of litter.
  + Show the detection accuracy in testing (e.g. 90% sure this is a Plastic bottle-Litter).
  + *Striving for/Conditional features*:
    - Categorize the litter (e.g. Plastic bottle, Paper bag).

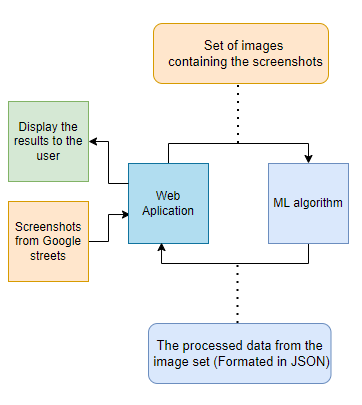
We will strive to deliver a new Machine Learning algorithm for Keep America Beautiful, developed through Python in a Anaconda/Miniconda environment, that will output usable and available results for the Web App team.

**1.5 Measures of Success**

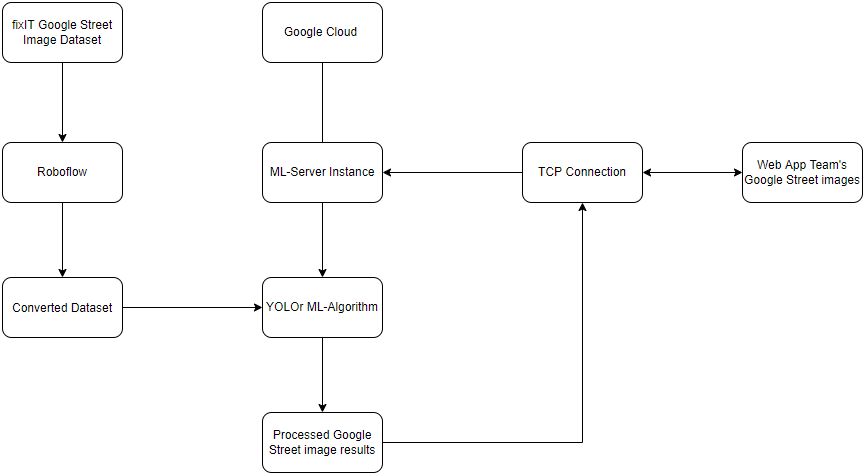
The Sushi Team will perform the following tasks in order to accomplish the requirements listed in the proposal.

* JAD #1 - Identify initial requirements and technical specifications for the project.
  + Connect to a Google Cloud Server.
  + Establish Anaconda/Miniconda environment.
  + Read documentation/code of TACO’s pre-trained model.
  + Collect/Annotate Google Street images for future initial testing.
* JAD #2 - Identify the final requirements and technical specifications for the project.
  + Further collect Google Street images for future initial testing.
  + Establish extra student labor to assist in Google Street image annotations, and the access to software and Google Street images they will need.
  + Research object-detecting pre-trained model options for our detecting litter process.
* Prototype #1 - Process images through the Machine Learning algorithm to produce litter metrics.
  + Fully utilized and integrated the YOLOR object detection model into a Google Cloud server.
  + Trained the YOLOR model with a previous group's Google Street image dataset for litter detection (converted through Roboflow).
  + Wrote a Python script to establish an input/output format, with a TCP connection, for the utilization of the Machine Learning algorithm.
  + Began to look into possible different sources of litter related datasets to better the results of our Machine Learning algorithm for Google Street images.
  + Modified previous group's Google Street image dataset annotations to try to get better results for the algorithms litter detection.
* Prototype #2 - Able to sync up the input/output format of the Machine Learning algorithm with the Web App team, as well as refine the algorithm's litter detection.
* Final Product - Collation and transfer of litter metrics from images, received by the Web App team, that have been processed through a Machine Learning algorithm, for use by the Web App team.

**1.6 System Overview**



* Closer View of the System - Prototype #1

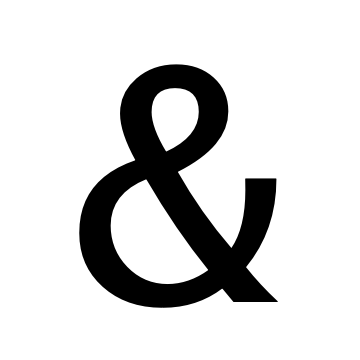










*(Disclaimer: Cropped images to fit in document.)*

* Currently, we have created a system that can receive and process Google Street images from a user, in which a Machine Learning algorithm will be able to detect instances of litter in the received Google Street images. Once the litter instances have been identified, the system will then output the results into an image with bounding-boxes, identifying the litter, and a JSON file, textualizing the results, for the user.

### 2. Requirements Matrix

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| REQ # | Requirement name | Description | Critical | Implemented | Task ID |
| 1 | Process Google Street images | Google Street image datasets can process through a pre-trained model. | Y | Y | 2.7.1 / 3.3.1  3.3.4 / 3.3.9  3.3.10 / 3.5.2  3.8.1 / 4.3.1 |
| 2 | Litter Detection | The algorithm will identify the presence of litter in Google Street images. Conditional: As well as classify them. | Y | N\* | 4.3.7 / 4.5.4  4.5.5 / 4.8.2  4.8.3 / 4.8.18  5.2.2 / 5.2.3 5.4.4 / 5.4.5 |
| 3 | Prepare Google Street Image Metrics | The Machine Learning program will process the algorithm’s litter results into JSON. | Y | Y | 4.8.9 / 4.8.16  4.8.17 |
| 4 | Send Prepared Metrics | Once the results have been prepared, the program will deliver the metrics to the Web App team. | Y | N | 5.2.9 / 5.2.10  5.2.11 |

### 

### \* (For REQ # 2 - Litter Detection: Functional in the system, but not fully implemented to our liking, in terms of results.)

### 

### 

### 

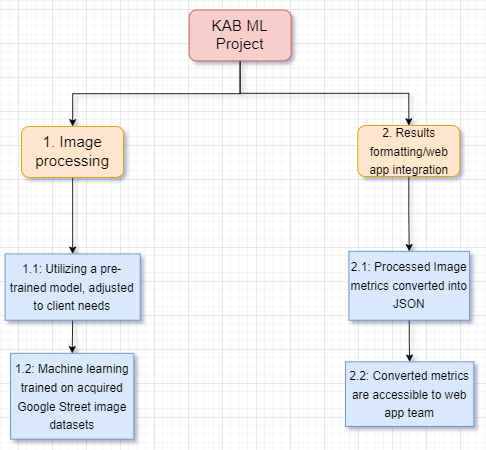
### 

### 

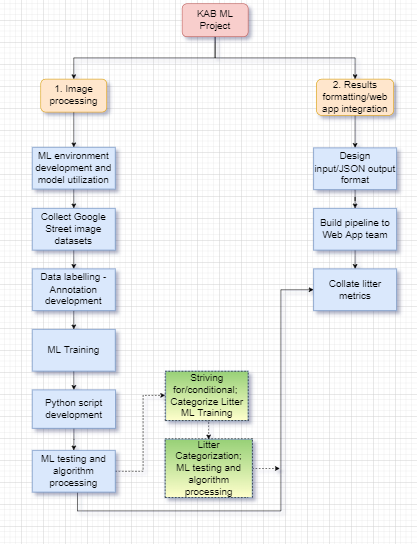
### 

### 3. Project Management

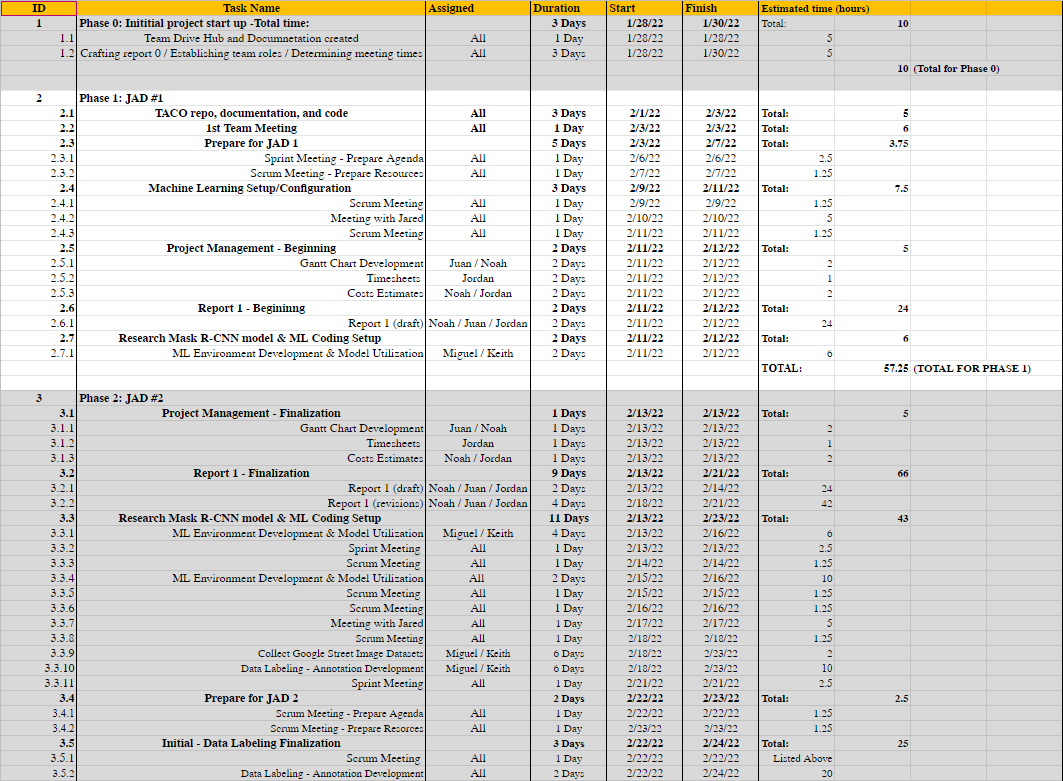
**3.1 Product Breakdown Structure**

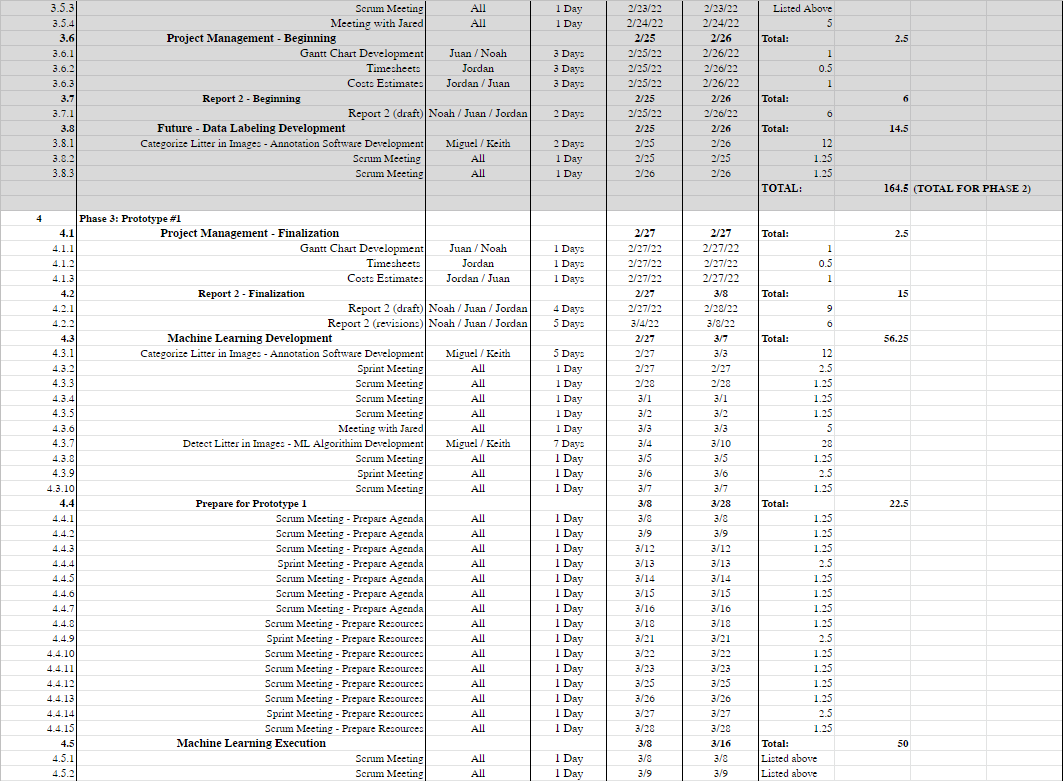


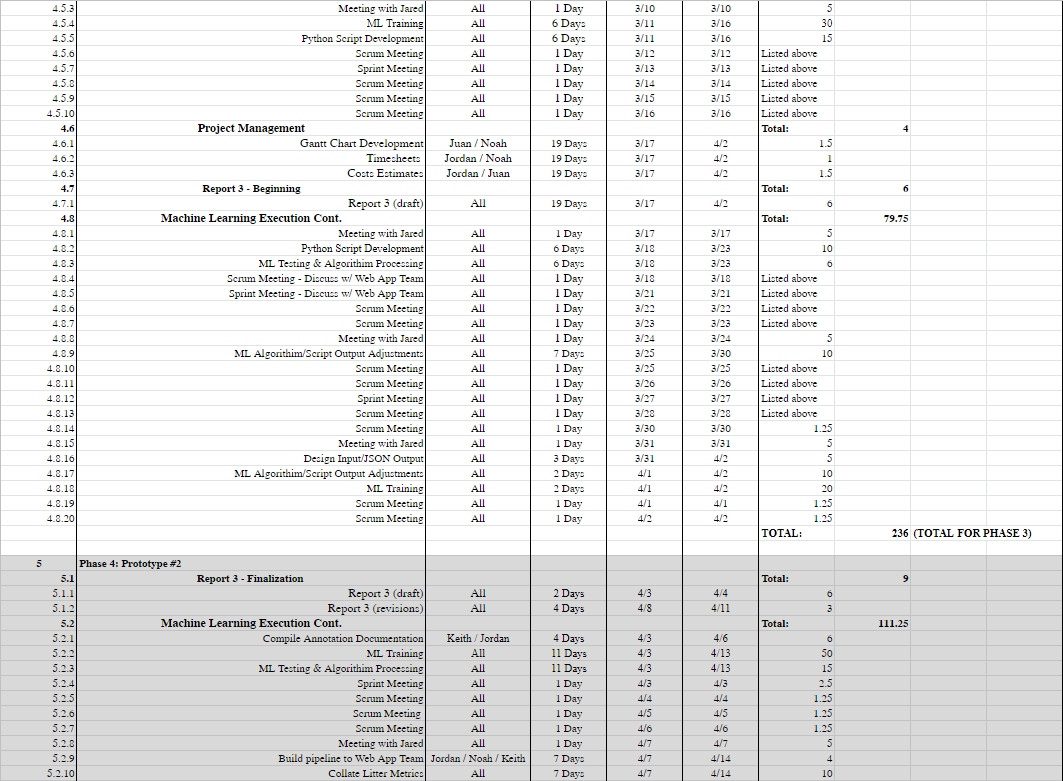
**3.2 Work Breakdown Structure**

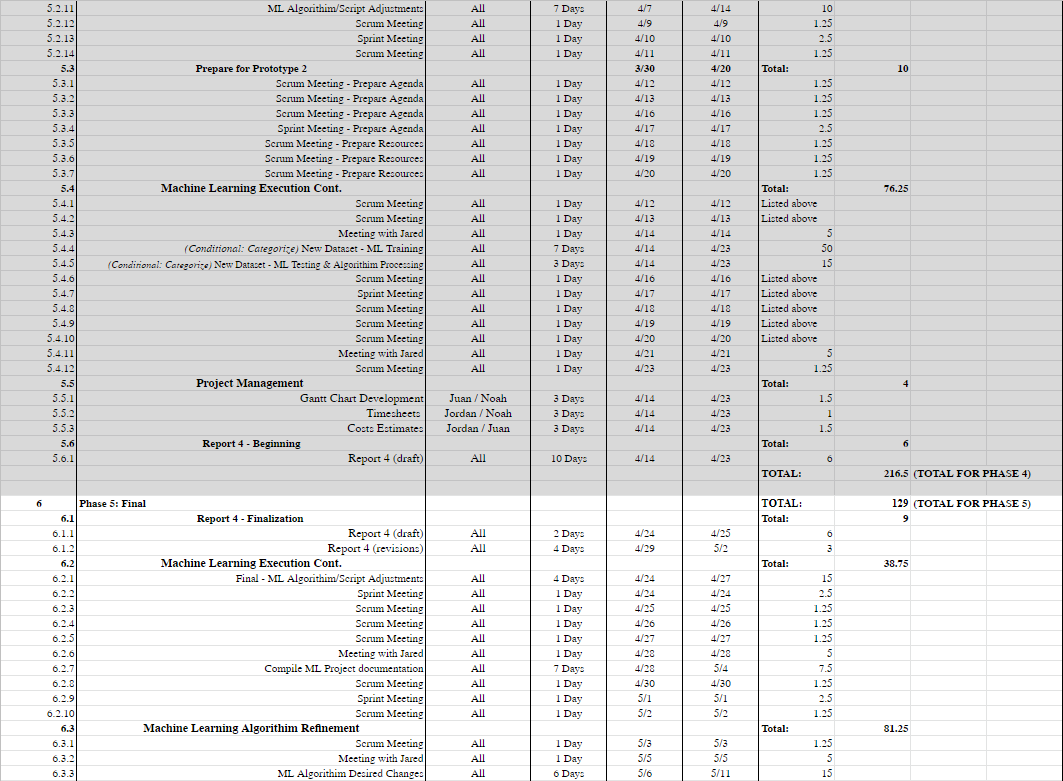
****

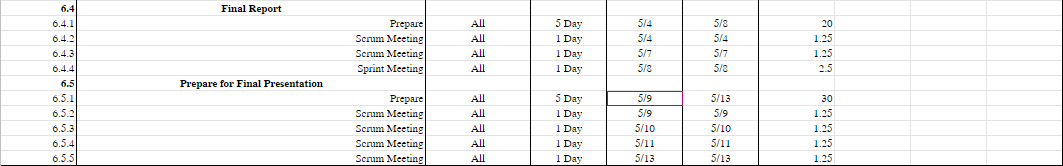
**3.3 Initial Schedule for Tasks and Deliverables**











**3.4 Statement of Total Price**

We will be charging a flat rate of $23.00 per hour for each member of the Sushi Team. The Google Cloud server will cost $0.47 cents per hour to run.

Through phase 1 of the project, it was estimated that the cost will be $1,316.75. This is based on an estimated 57.25 hours of work done by the team in this phase. In actuality, the cost was $1,138.50, having been completed in 49.5 hours of work.

During phase 2, all the members were estimated to work a combined total of 164.5 working hours. As such, the estimated cost for the phase was $3,783.50. In actuality, however, a total cost of $3,047.50 was accrued with a total of 132.5 hours.

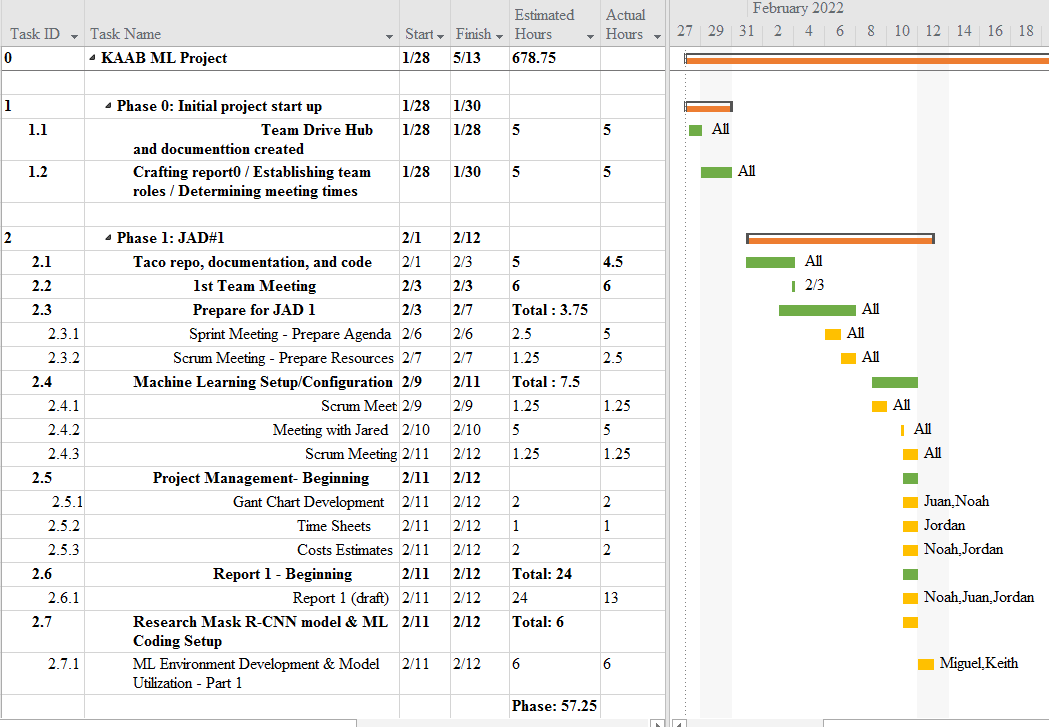
During phase 3 of the project, the Sushi Team was projected to work for 236 hours, with a projected cost of $5,428.00 for the phase. In actuality, the team worked a total of 221.77 hours, with a total labor cost of $5,100.71. Furthermore, due to provided discounts & promotions from Google Cloud, we have accrued zero costs from the platform within this phase, during our usage of it.

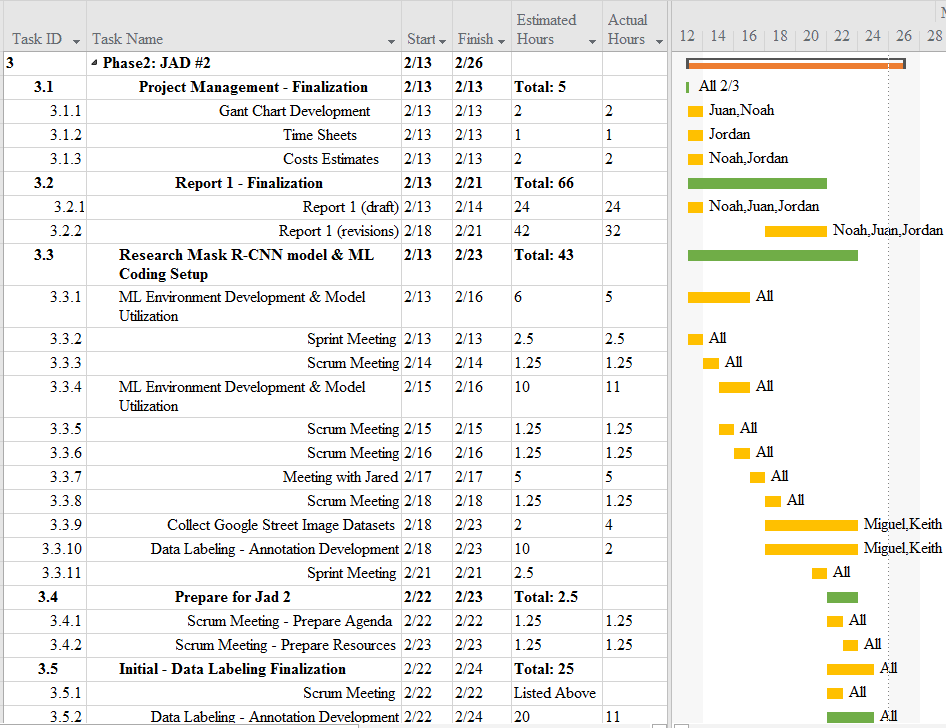
During phase 4 of the project, the total number of estimated working hours spent towards development amounts to 216.5 hours, with a total cost of $4,979.50.

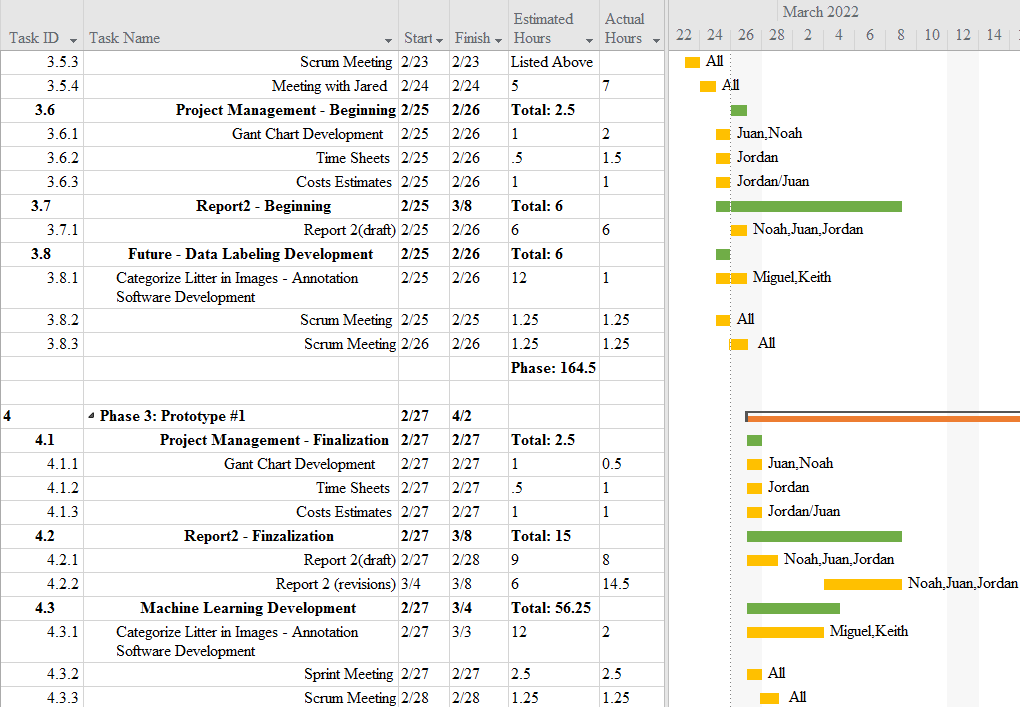
During phase 5 of the project, the total number of estimated working hours spent towards development amounted to 129 hours. Thus, the total cost for this phase will be $2,967.00.

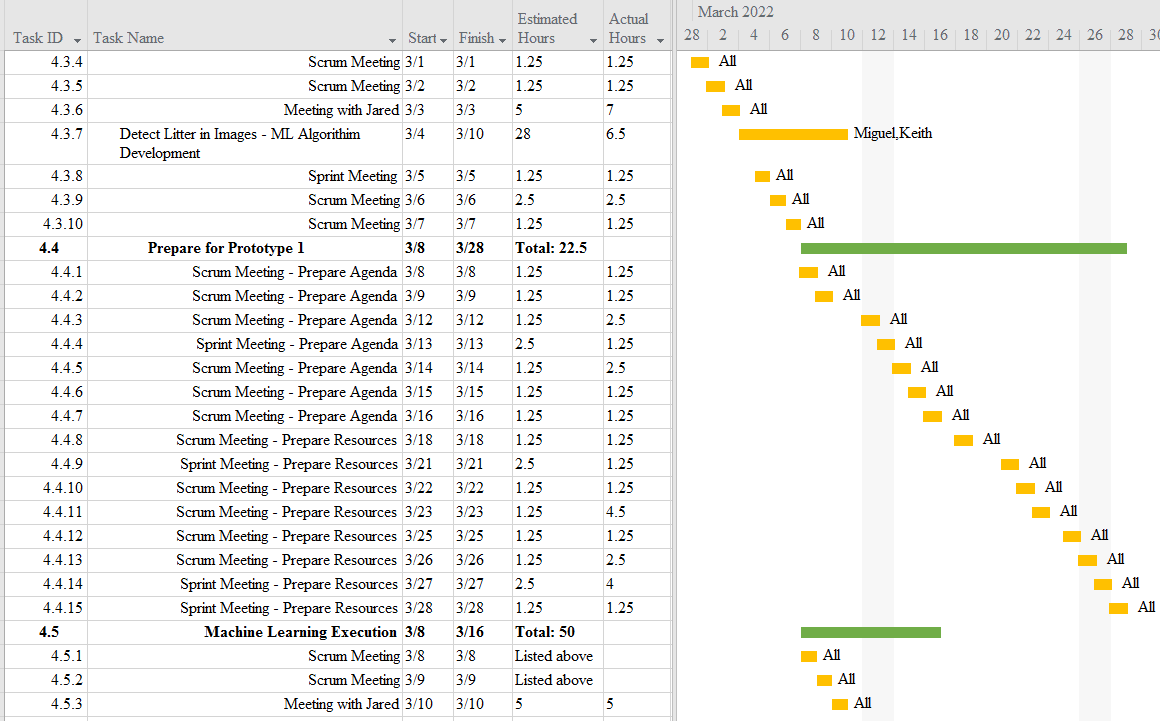
Based on our original estimates,the final cost of this project was projected to be $18,474.75, with 803.25 hours spent by the team towards the development of the Machine Learning algorithm. However, when adjusting for the actual hours of the Phases 1, 2, & 3, we expect that the total cost of the project will come to $17,233.21, accrued over a total of 749.27 labor hours. Should alternative or additional costs become required, the above estimates will be updated accordingly.

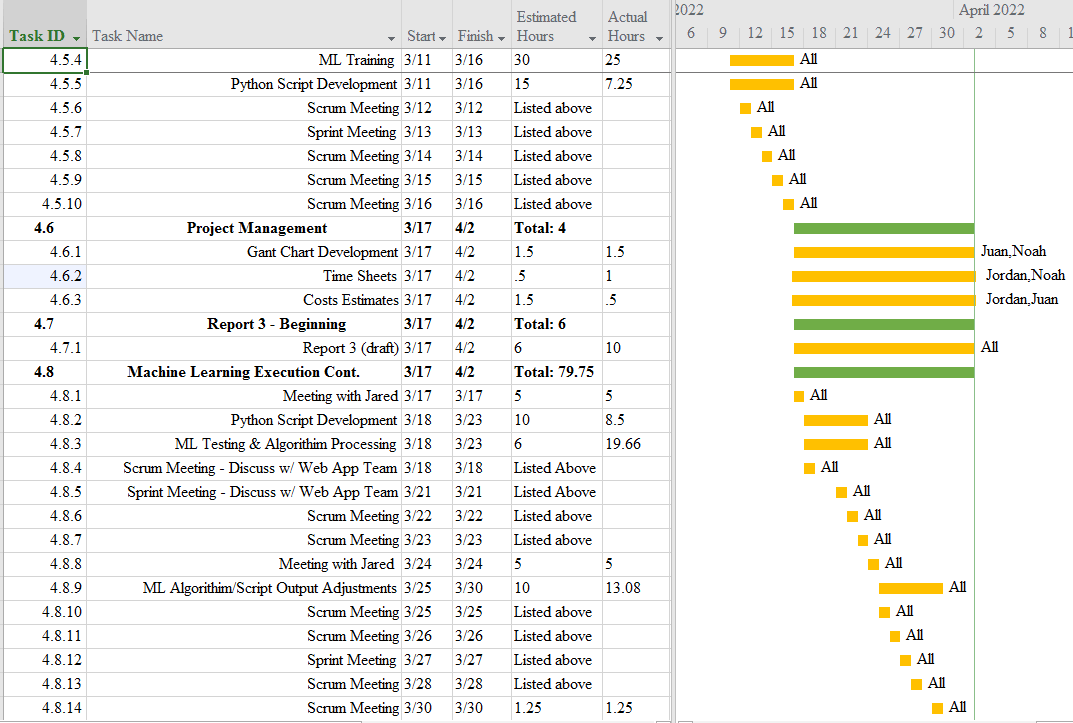
**3.5 Gantt Chart**

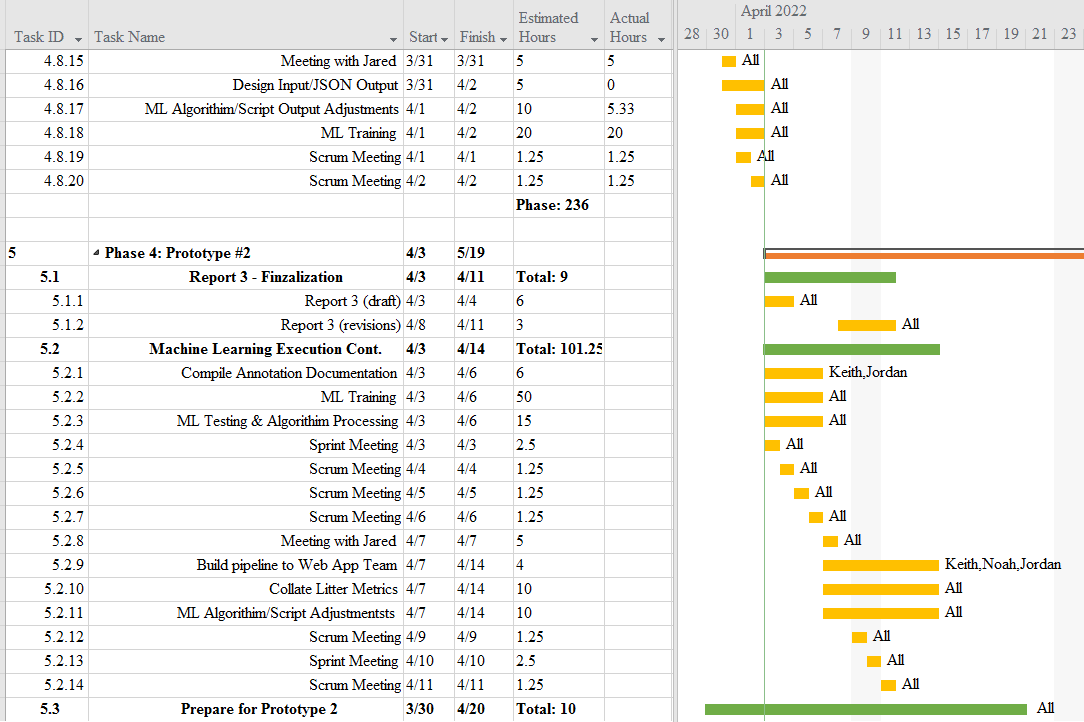


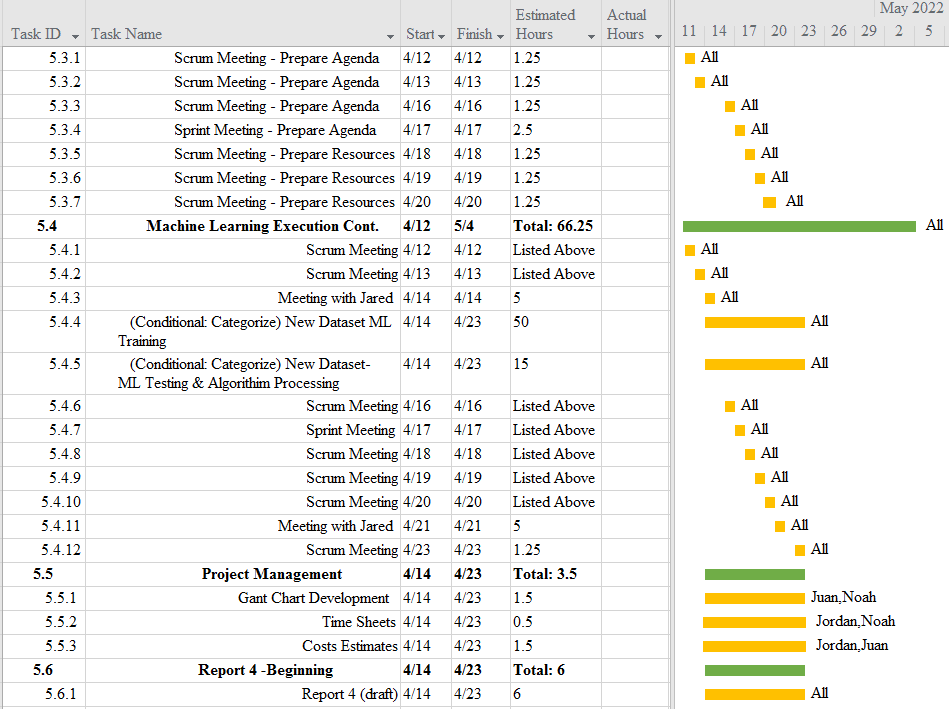


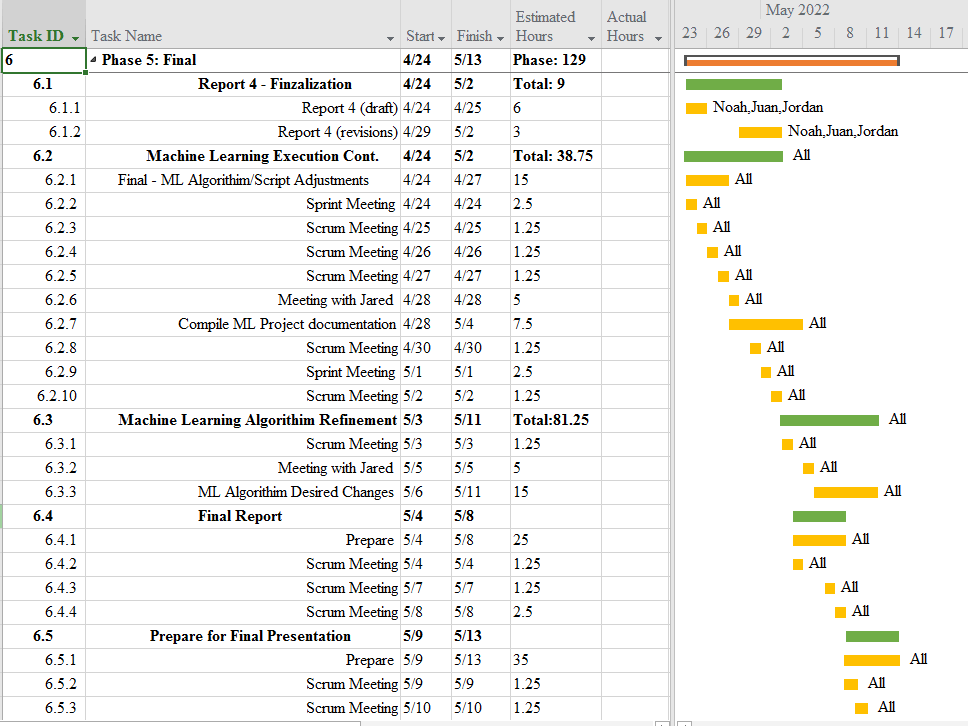








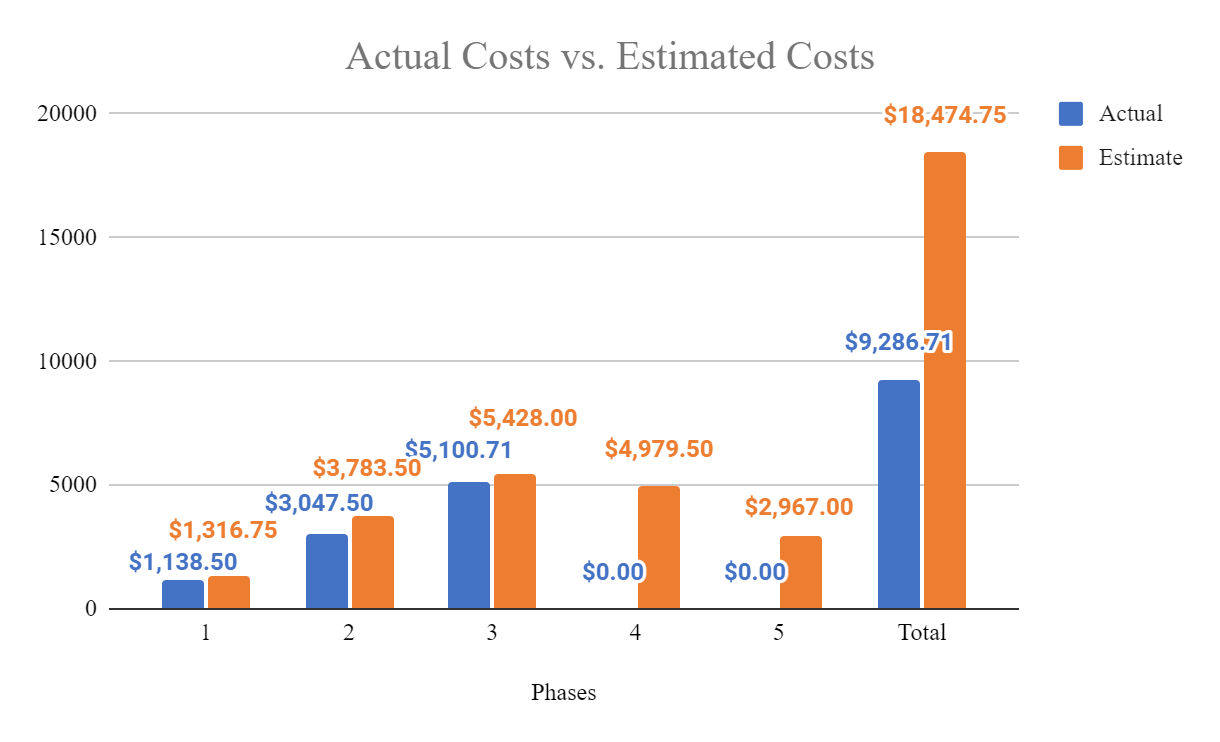


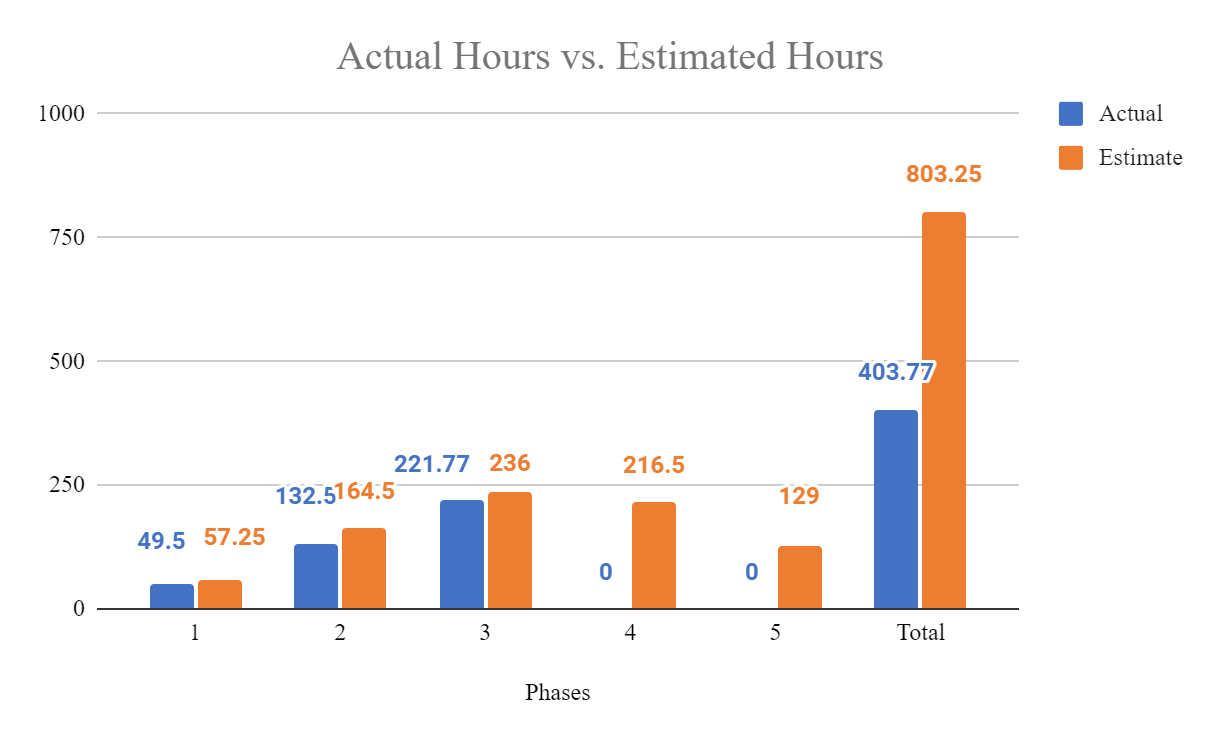




**3.6 Initial Project Cost Tracking Chart**

(As of now, by Phase 3)

****

****

The above chart tracks the estimated and actual costs/hours for each phase of the project as well as our total accumulated costs compared to our estimated costs. We will assume an hourly rate of $23.00 per hour for this project, as we estimate to work a total of 803.25 hours.

**3.7 Statement of Deliverables**

The Sushi Team strives to deliver the following final products:

* A fully working machine learning algorithm that will detect instances of litter in an image. When executed, it will display the Street image with outlines of the litter that is visible, along with percentages based on the accuracy. We will also strive to, though conditional, identify the type of each litter instance.
* A scheduling hook by which the web app team can use this algorithm by sending a set of images.
  + A system to format the data collected from the set of the set of images sent using the web application.
  + A system to send the formatted data back to the web application for further use as determined by the web app team.
* A Github repository for the machine learning project/algorithm.
  + Source code for the algorithm.
  + Documentation, manuals, and reports for users and programmers.

Phase 4 - Prototype #2 will result in the refinement of Prototype #1 to detect litter , and possibly categorize, as well as finally establish a hook by which the Web App team can use the Machine Learning algorithm for their Google Street images. (Expected 4/21/22).

**3.8 Outline of Resources Needed**

The following are what the Sushi Team will be supplying:

* Knowledge of Python code, documentation, and implementation.
* Updated documentation.
* Research and testing.

The following resources are what the Sushi Team will need supplied from Jared Macshane:

* References and guidance of Machine Learning through existing models & open-source data/source code/software.
* Google Cloud Server Instance funding.
* Signage and approval for the Sushi Team to continue working on the project.
* Availability for future meetings.
* Access to Keep America Beautiful’s private server.
* Possible extra student labor to assist in Google Street image annotations for Google Street image datasets.

### 

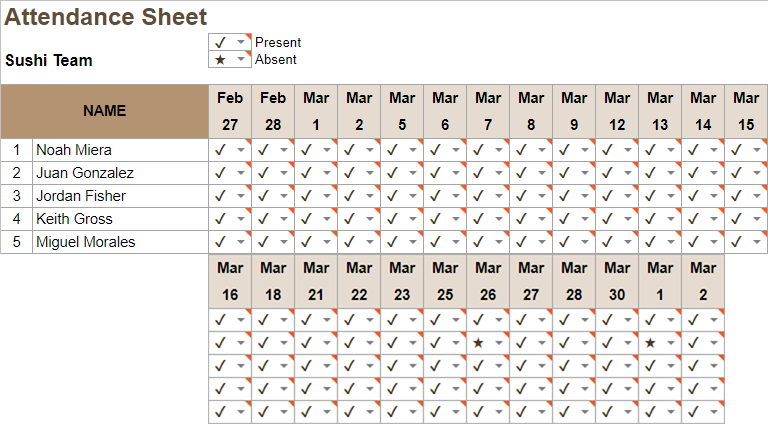
### 4. Team - Appendix

**4.1 Access to Project**

The following will be how the Sushi Team will be collarbaring during this project so-far:

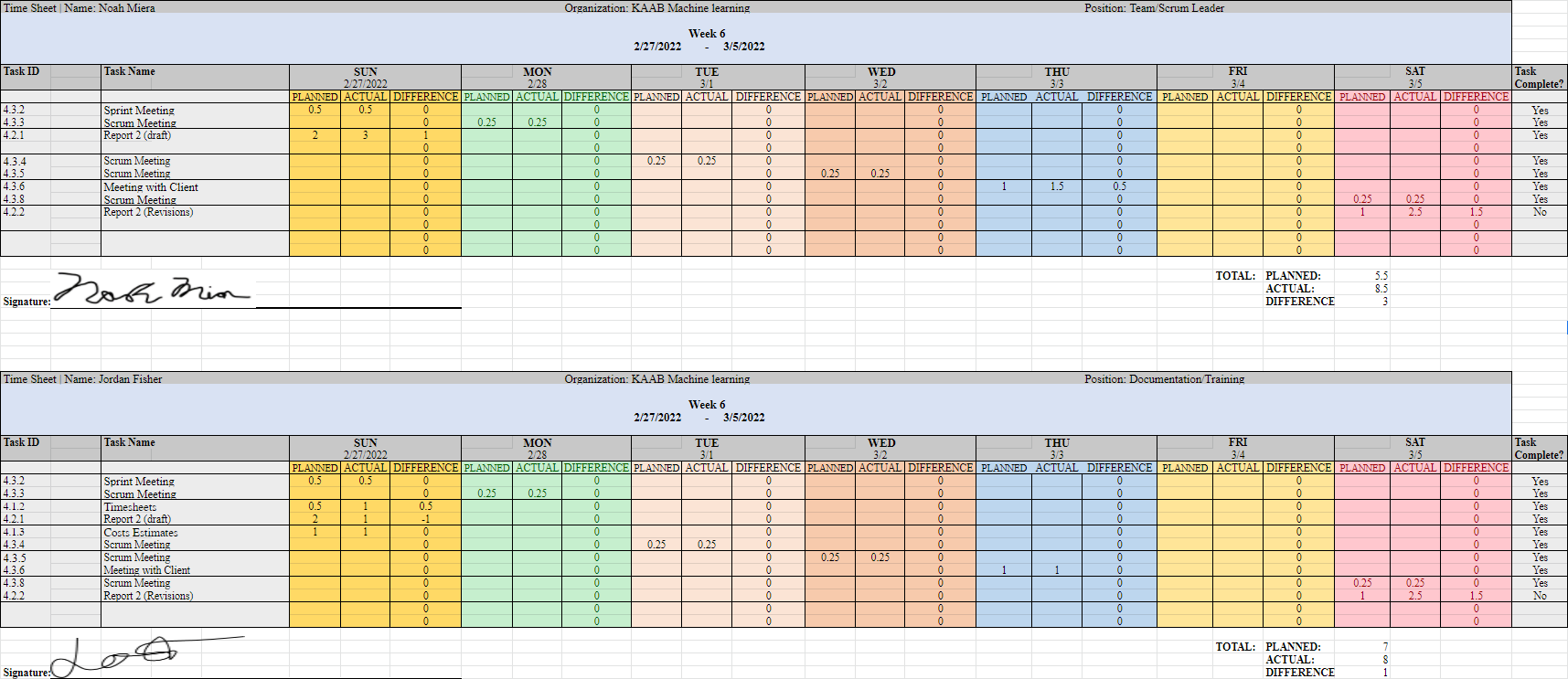
* Discord.
  + https://discord.gg/dr5785e9
* GitHUB
  + https://github.com/JuanGonzalez2020/KAAB-ML
* Google Cloud Server Instance.
  + Access is granted by Keith by giving him your SSH public key.
  + ssh username@34.152.63.61

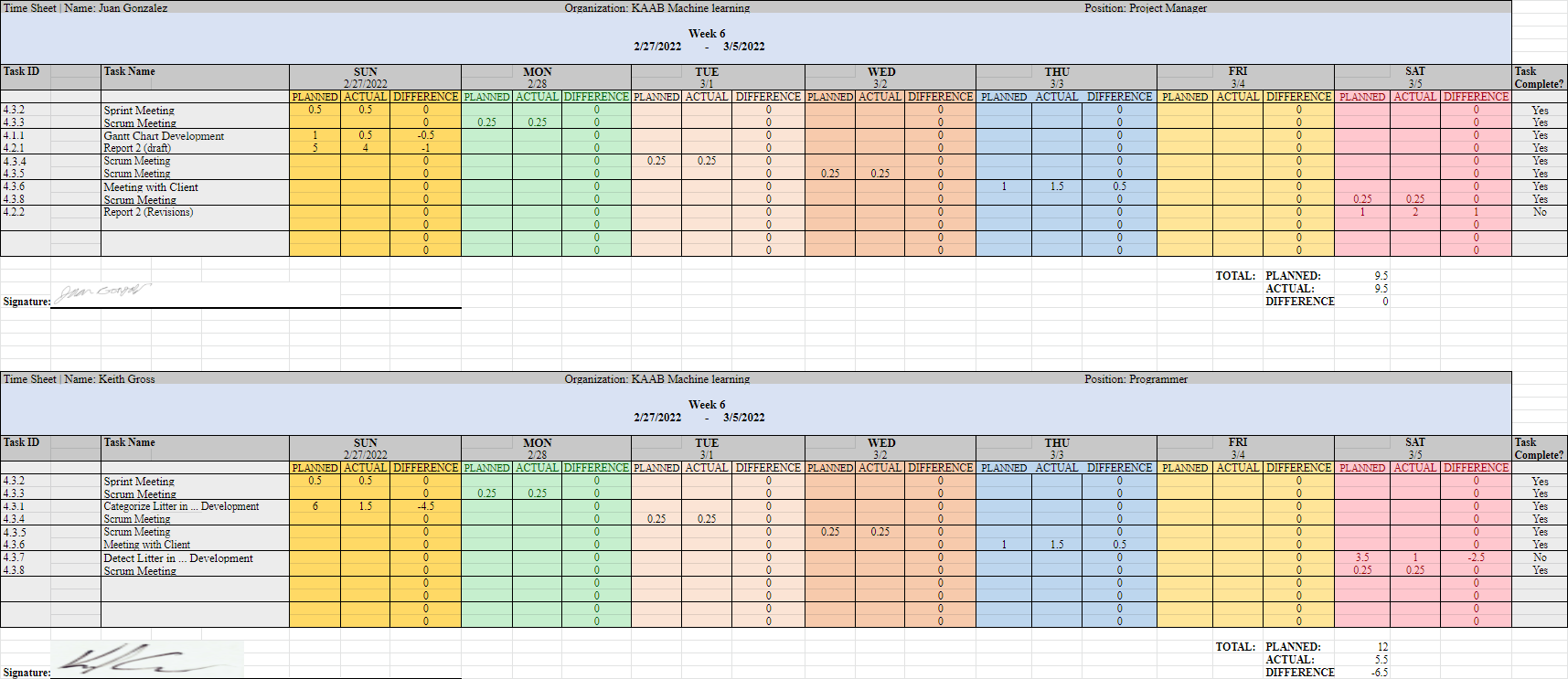
**4.2 Attendance Records**

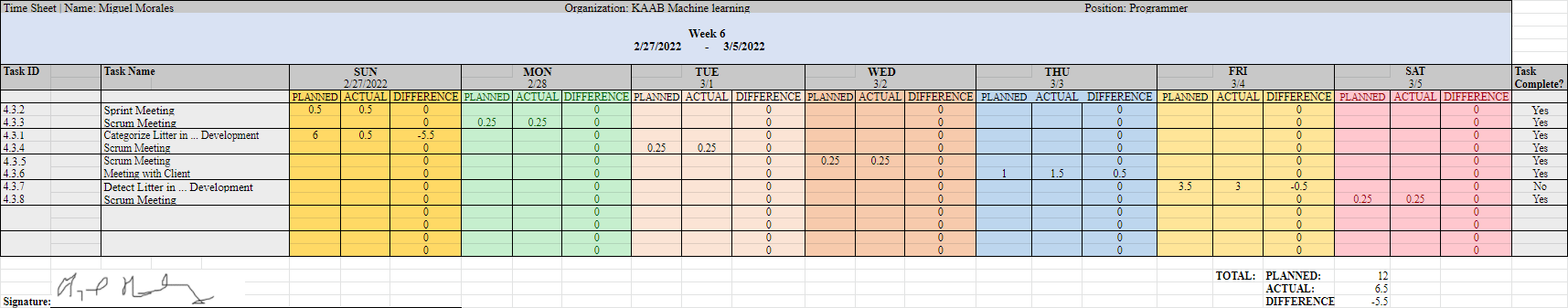
****

**4.3 Timesheets**

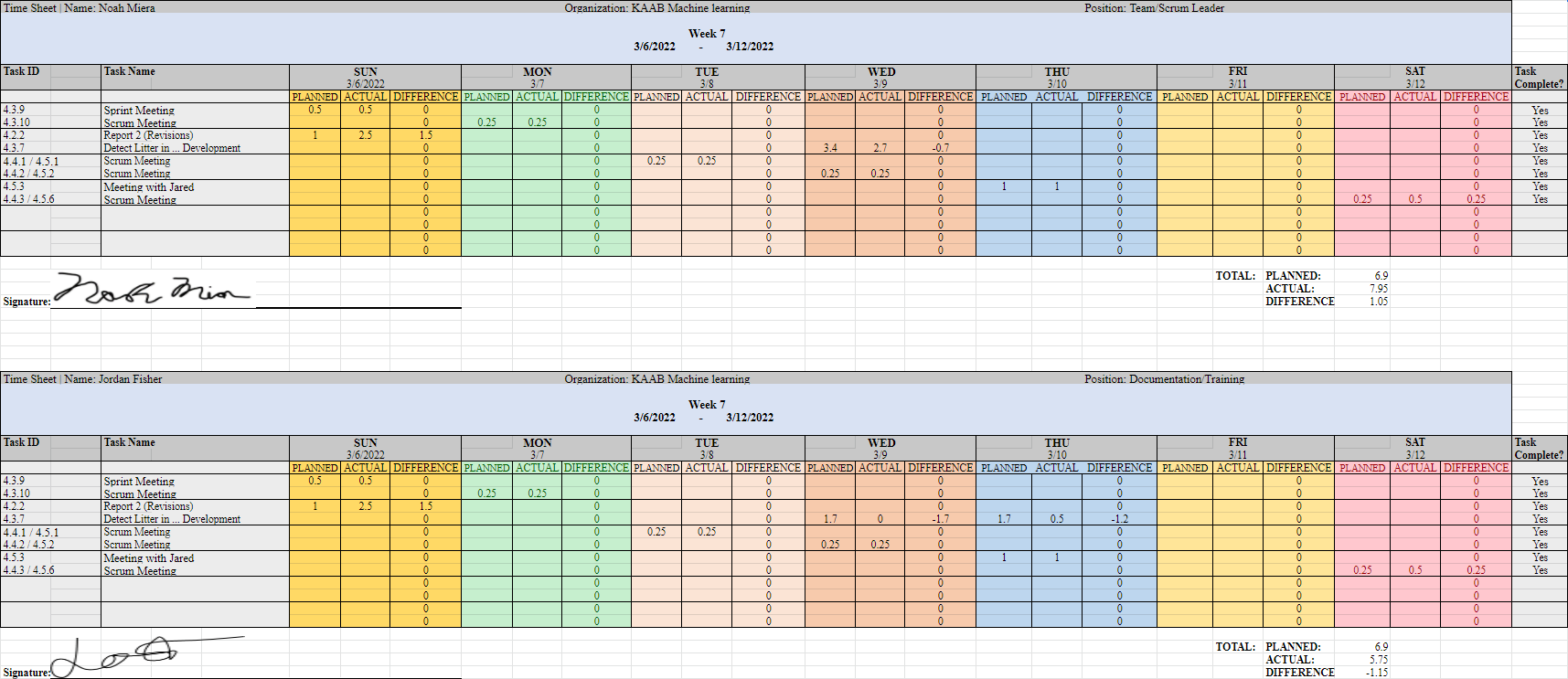
**Week 6**

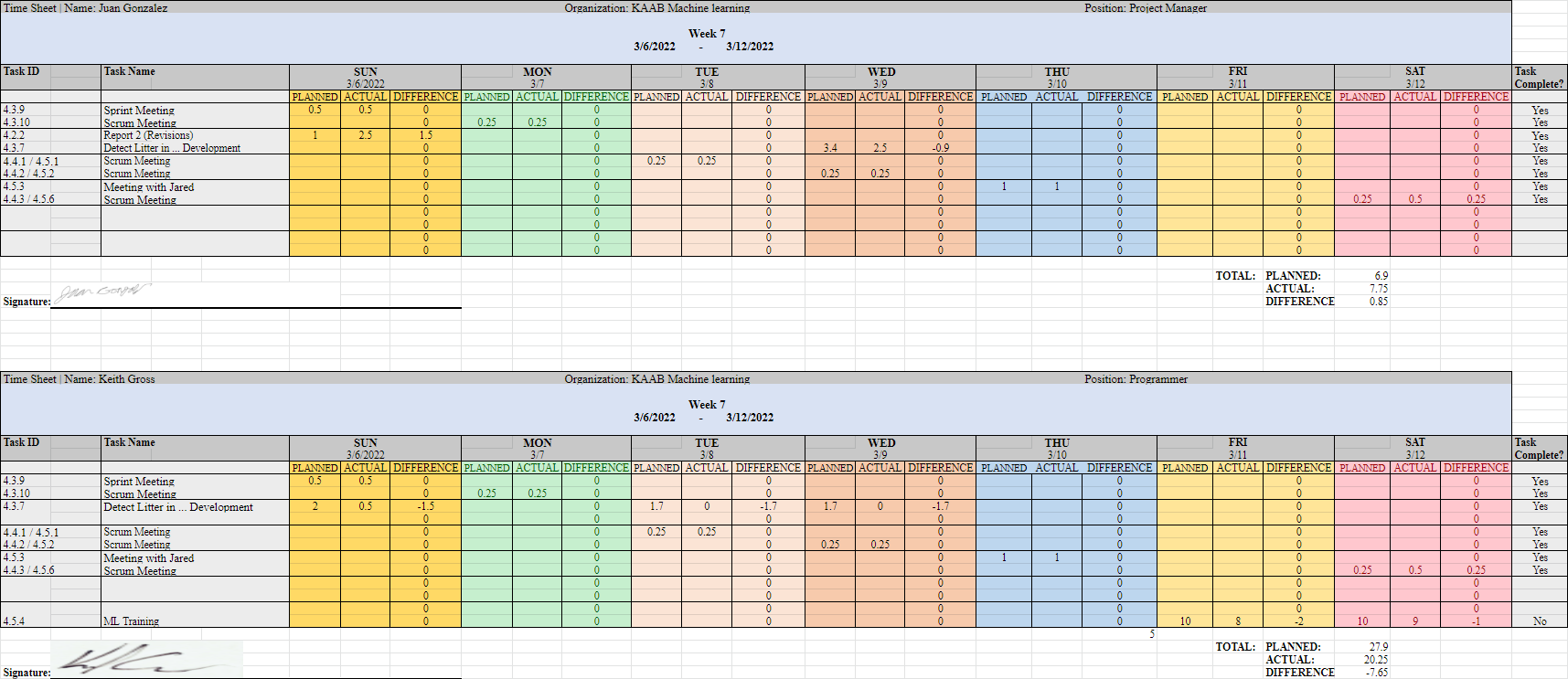


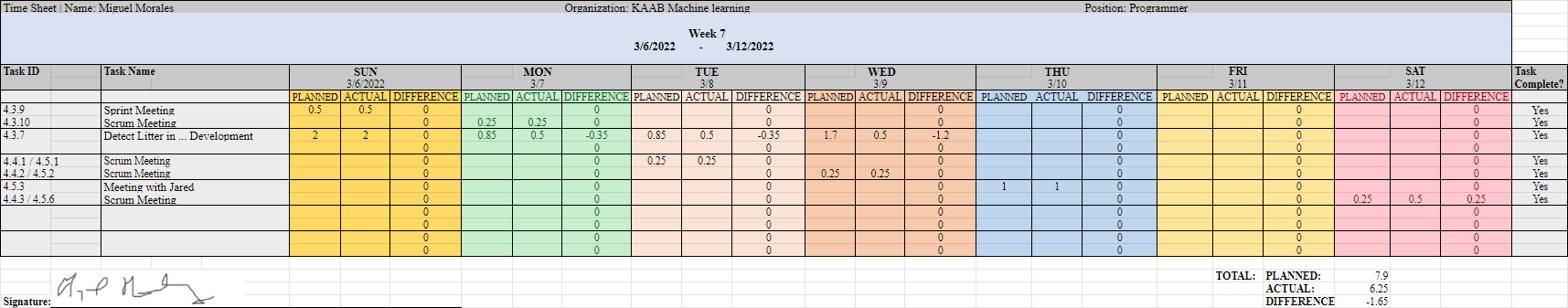




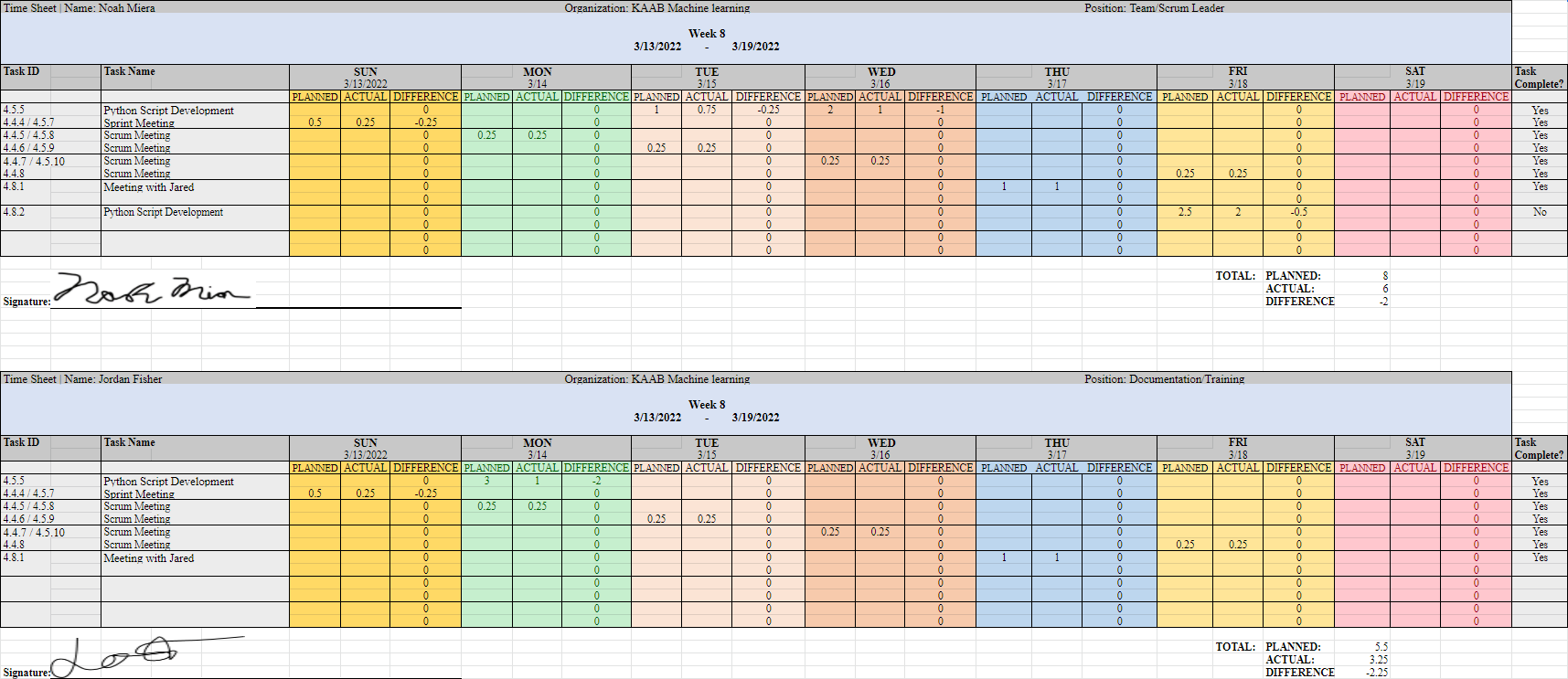
**Week 7**

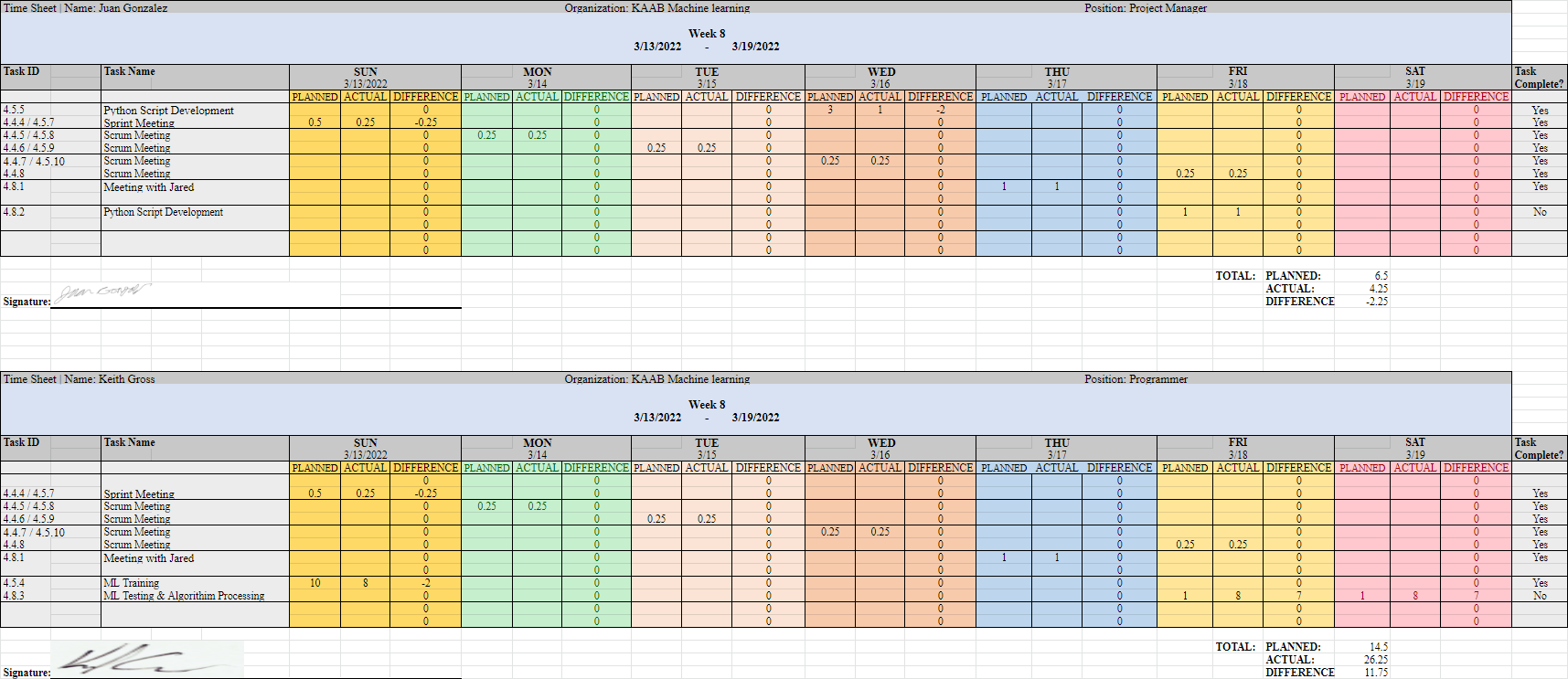


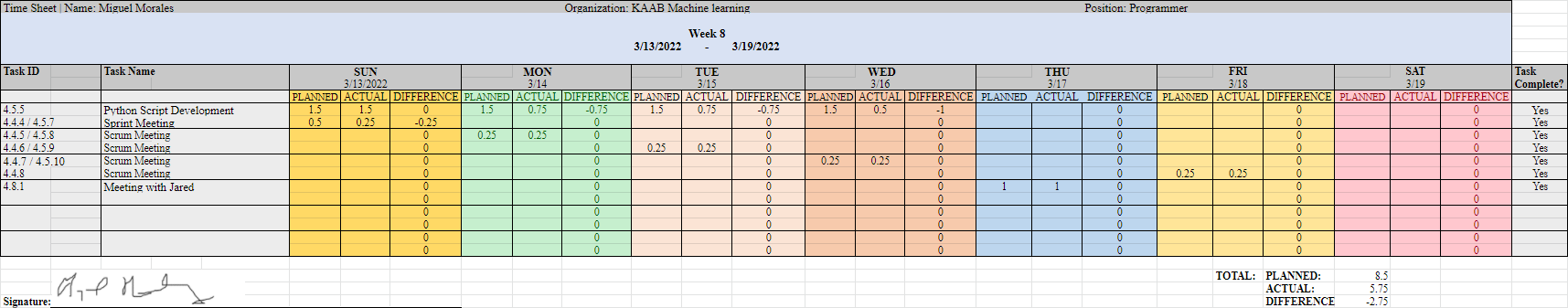




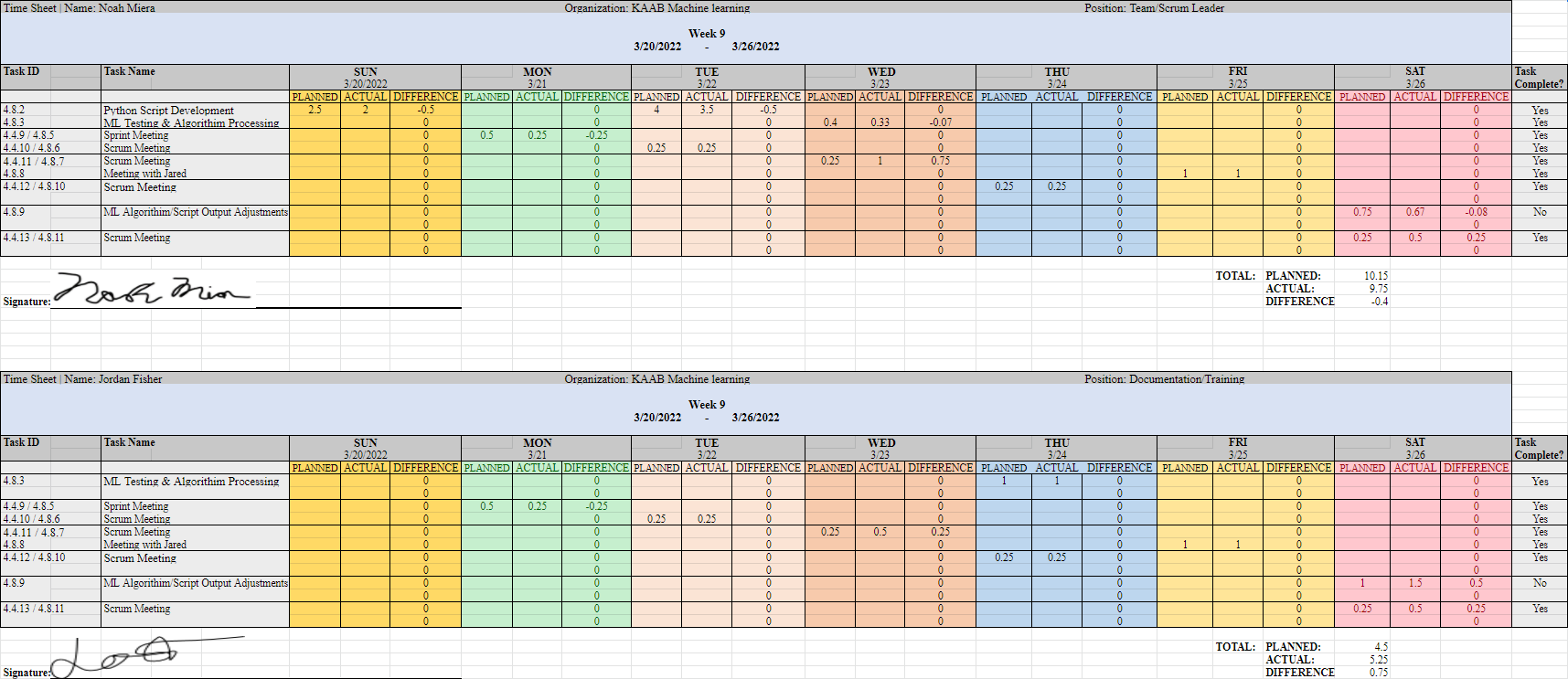
**Week 8**

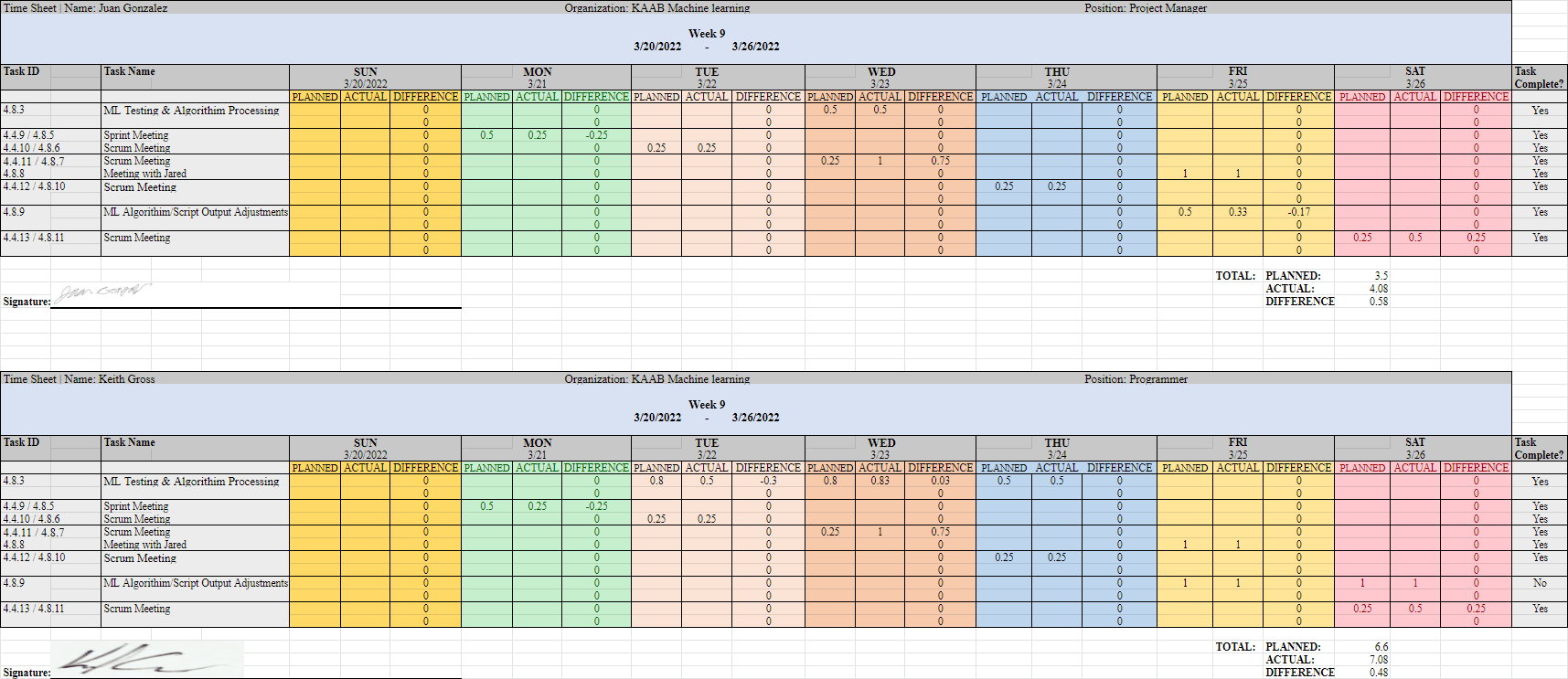


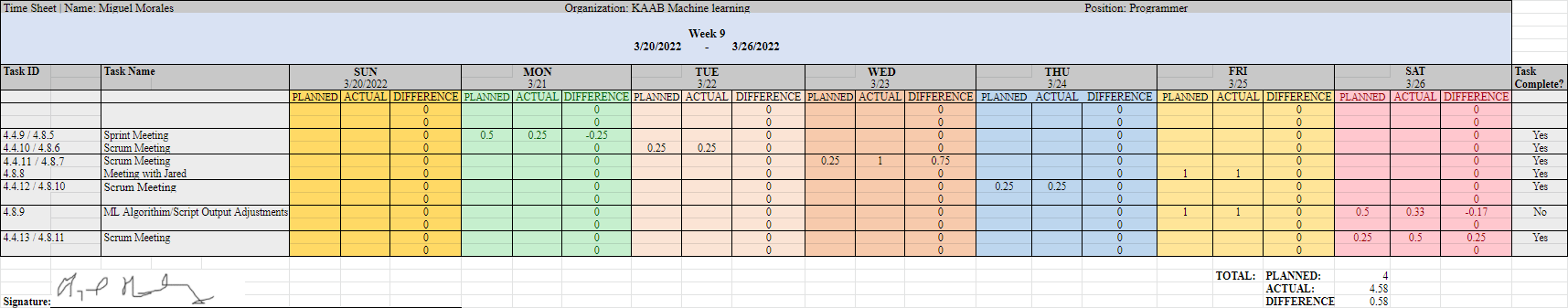




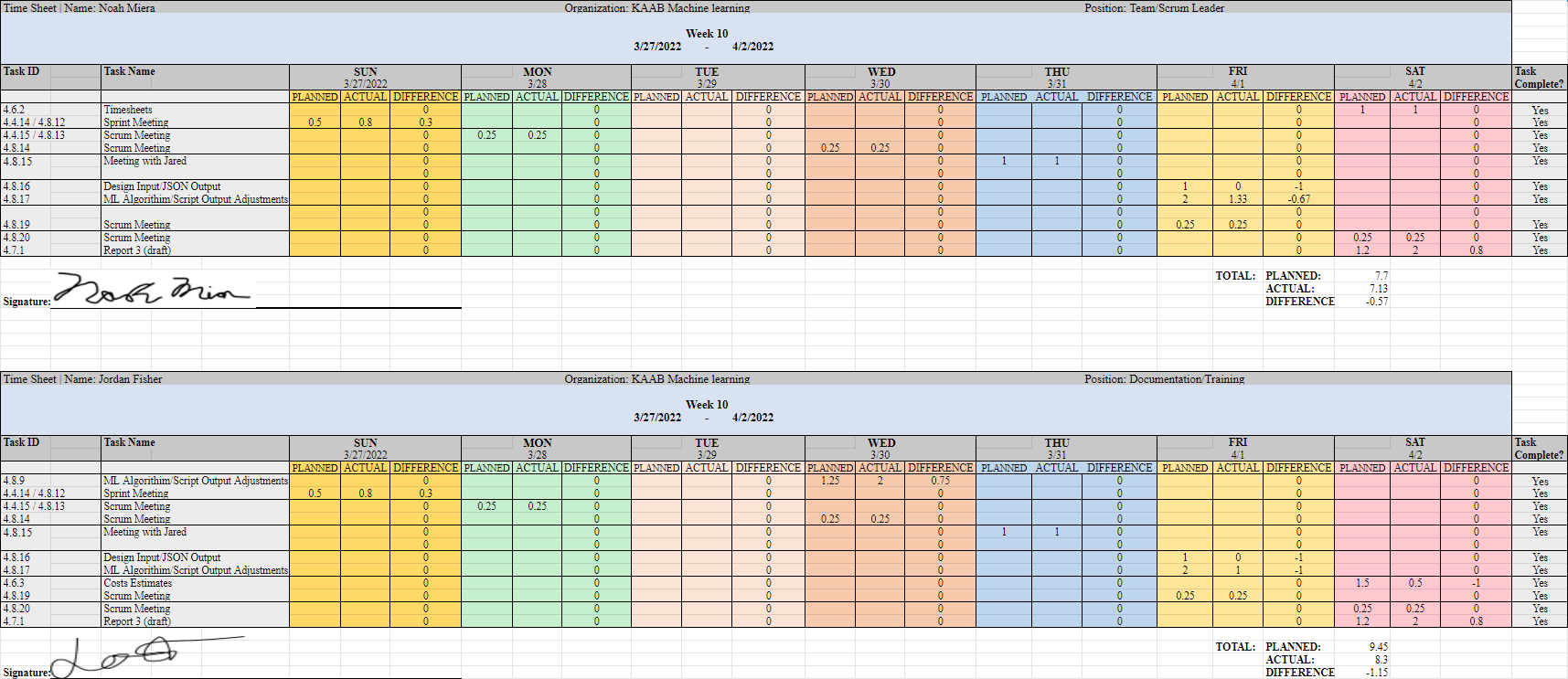
**Week 9**

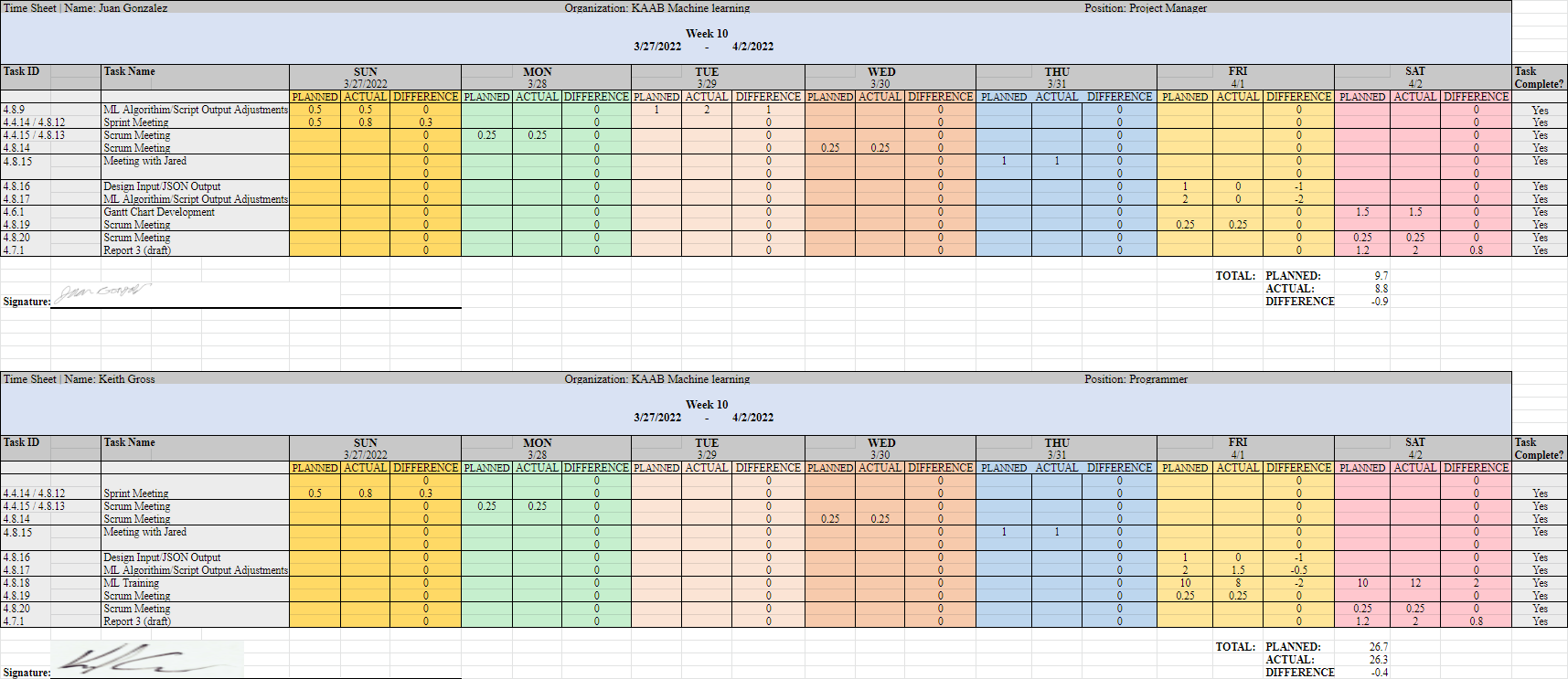


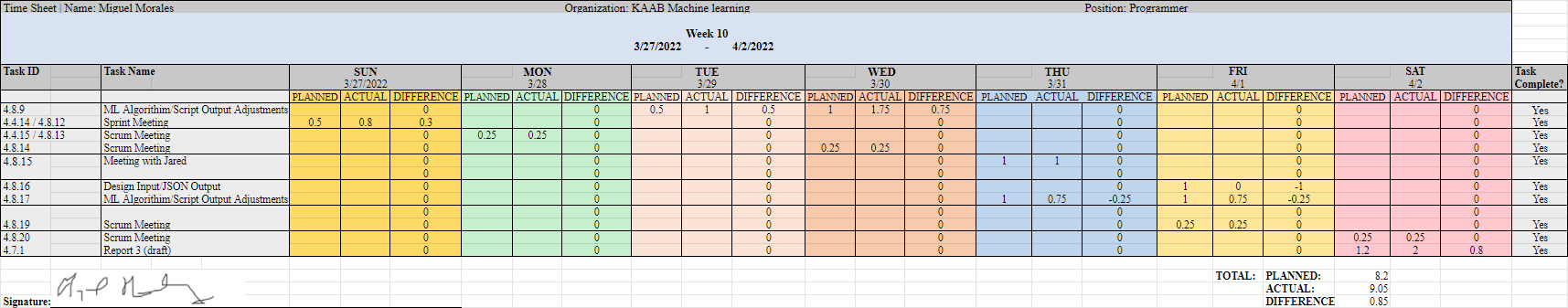




**Week 10**







**4.4 Sprint Minutes**

Sushi Team Sprint #4 Meeting Notes

Sunday, 02.27.2022

# **Attendees**

Noah Miera: Team/Sprint Leader

Juan Gonzales: Project Manager

Jordan Fisher: Documentation/Training

Keith Gross: Programer

Miguel Morales: Programer

# **Minutes:**

## 12:30PM: Review of Previous tasks & confirmation of actual times on timesheet

## 12:40PM: Explanation of VIA3 annotation software by Keith Gross

* Hosted online. All Dr. Schultz’s students would need to do is go onto the website and open the project with the correct project id
* Images can be added to the project
* Potential issue: importing images via URL is most effective, however they may need to be publicly hosted before they can be imported

## 12:51 PM: Discussion of Upcoming Tasks

* Keith and Miguel will be looking in to setting up VIA3 for Dr. Shutz’s students
* Noah, Juan, and Jordan will be working on report 3

## 1:00PM Meeting Adjourned!

Sushi Team Sprint #5 Meeting Notes

Sunday, 03.6.2022

# **Attendees**

Noah Miera: Team/Sprint Leader

Juan Gonzales: Project Manager

Jordan Fisher: Documentation/Training

Keith Gross: Programer

Miguel Morales: Programer

# **Minutes:**

## 12:00PM: Confirmation of time for meeting with Group 2 and Jared

* 1PM with Jared as well as Ray and Angelica from group 2
* Develop questions for that

## 12:05PM: Review of complete and upcoming revisions for report 2 revisions

* Group will reconvene at 3:30 to finalize any changes

## 12:07 PM: Checking in with programmers about XML to YOLOR annotation conversions

* Roboflow seems to work well,
  + however some of the annotations are sized poorly
    - Could be a product of the original annotators rather than the program itself
  + also allows for augmentations which may increase the number of images we functionally have
* Discuss the changes to old annotations with Jared

## 12:15 PM: Discussion of upcoming tasks:

Programers will follow up on converting old dataset to YOLOR and the preparation for the training

* Miguel will focus on the isaychris method
* Keith will prepare for the training phase

## 12:25 PM: Confirmation of meeting time tomorrow at 10:30:

## 1:00PM Meeting Adjourned!

Sushi Team Sprint #5 Meeting Notes

Sunday, 03.13.2022

# **Attendees**

Noah Miera: Team/Sprint Leader

Juan Gonzales: Project Manager

Jordan Fisher: Documentation/Training

Keith Gross: Programer

Miguel Morales: Programer

# **Minutes:**

## 12:00PM: Housekeeping/Timesheet review

## 12:05PM: Overview of Upcoming tasks

* Research processing images using YOLOR model.
* Look into script options for YOLOR
  + TCP may be an option, check with Jared
* Look into output options

## 12:20PM: Roboflow Reannotation Updates

* Get public keys to Keith so that he can add us to the workspace

## 12:25 PM: Meeting Adjourned!

Sushi Team Sprint #7 Meeting Notes

Wednesday, 03.23.2022

# **Attendees**

Noah Miera: Team/Sprint Leader

Juan Gonzales: Project Manager

Jordan Fisher: Documentation/Training

Keith Gross: Programer

Miguel Morales: Programer

# **Minutes:**

## 10:30 AM: Timesheets/housekeeping

* Time with for machine training will be counted towards time accrued
* Develop questions for that

## 10:40 AM: Presentation of text formatting -Noah

* Next goal is getting the code onto the google instance and see if it can be run there

## 10:45 AM: uploading the tcp transfer program to the google instance and initial testing

## 11:00 AM: Meeting Adjourned!

Sushi Team Sprint #8 Meeting Notes

Sunday, 03.27.2022

# **Attendees**

Noah Miera: Team/Sprint Leader

Juan Gonzales: Project Manager

Jordan Fisher: Documentation/Training

Keith Gross: Programer

Miguel Morales: Programer

# **Minutes:**

## 12:00 PM: Timesheet confirmation

## 12:05 PM: Updates to prototype model

* Prototype model can now run in a loop

## 12:10 AM: Discussion of Prototype 1 presentation

* Training process will be included on the network diagram
* Make sure that we are specifically saying google street images rather than simply “images”

## 12:40 AM: Discussion ofGroup 2’s progress with image collection

* It is unlikely that we will have the images ready from group 2. The backup dataset may see use.

## 12:47 AM: Review of Tasks

* Training process will be included on the network diagram

## 12:48 AM: Meeting Adjourned!