Project Plan

VIP Intestinal Worms Team, Fall 2023

Nathaly Jose-Maria

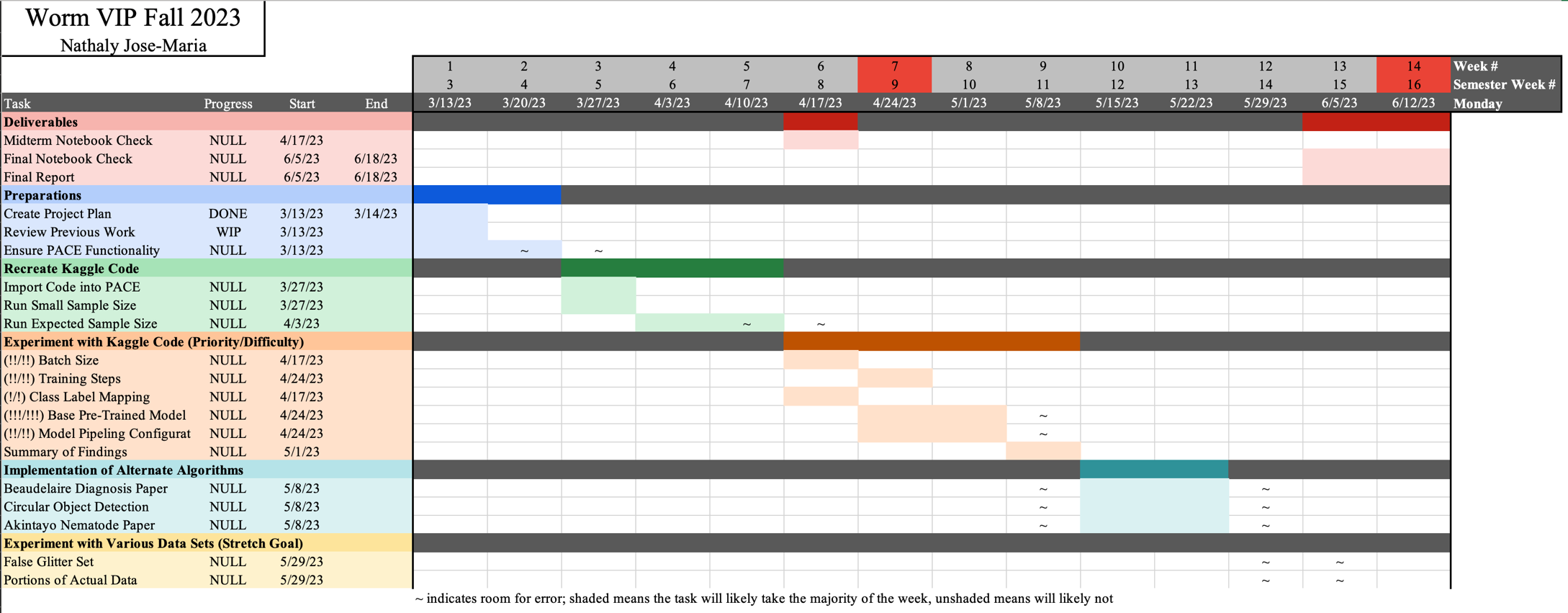
The overall efforts of the Intestinal Worms VIP team look to improve the diagnosis of soil transmitted helminths for public health professionals in areas where the issue is prevalent and necessitating quick, efficient work. We seek to improve diagnosis by creating an identification system that reliably counts the presence of roundworms, hookworms, threadworms, and whipworms in stool samples, which is indicative of the quantity and severity of infection in an area. The current system in place is the Kato-Katz system, and it utilizes pathologists that manually quantify these worm eggs in stool samples. The labor is timely, intensive, and uncomfortable, and as such, limits what can be done for the communities impacted by these worms, including the provisioning of deworming medication. The creation of a low-cost, easy to use automatic quantifier conforming to World Health Organization sensitivity and specificity criteria will enable outreach in need-based communities even without the physical presence of pathologists in these areas (Children Without Worms, 2021a; Children Without Worms, 2021b).

The general problem of this semester is to have a manipulable, executable version of the code and data presented in the paper “Affordable artificial intelligence-based digital pathology for neglected tropical diseases: A proof-of-concept for the detection of soil-transmitted helminths and *Schistosoma mansoni* eggs in Kato-Katz stool thick smears" (Ward 2022). The code has already been analyzed and adapted in theory to fit the needs of the Children Without Worms team, evident in our previous work “GTVIPFall22 STH ML Notebook” (Jose-Maria et al., 2022). Therefore, the theoretical adaptations need to be brought into reality, and this will happen with the help of high-performance computers. These computers, particularly Georgia Tech’s PACE ICE clusters, provide better environments for data processing and code runtime as opposed to the limitations presented in the original online server Kaggle, as previously detailed in our most recent team report (Hart et al., 2022).

Our main goals detailed by indented intermediate goals for this semester are as follows:

* Preparations
  + Reviewing previous work
  + Preparing PACE-ICE on local computers.
* Recreate execution of Kaggle notebook code in PACE-ICE clusters.
  + Successful import of code into PACE-ICE.
  + Successful execution of a small data batch in PACE-ICE.
  + Successful execution of an expected data batch in PACE-ICE.
* Experiment with variable executions of Kaggle notebook code.
  + Manipulate various parameters within the original notebook code and observe their output: Batch Size, Training Steps, Class Label Mapping, Base Pre-Trained Model, Model Pipeline Configuration
  + Present a summary of findings comparing the results of various parameter changes and optimal choices among those observe for our team’s needs.
* Implement alternate parallel-running algorithms of use.
  + Circular Object Detection
  + Ideas presented in “A deep learning framework to discern and count microscopic nematode eggs” (Akintayo et al., 2018).
  + Ideas presented in “Towards an automated medical diagnosis system for intestinal parasitosis” (Beaudelaire et al., 2019).
* Experiment with various data sets.
  + False Glitter Set created in previous semesters.
  + Smaller Subset of Real Data from Kaggle notebook
  + Any new images that can be collected

A detailed product timeline is attached (as of Mar 15th, 2023), but can also be viewed at the following link: <https://gtvault-my.sharepoint.com/:x:/g/personal/njosemaria3_gatech_edu/ES2T6abj-w9GrK6Ss4pbAQYBilGP5-SjPtZZaV_FIqaQ7g?e=QJjnb2>.



Works Cited

Akintayo, A., Tylka, G.L., Singh, A.K., Ganapathysubramian, B., Singh, A., & Sarkar, S. (2018). A deep learning framework to discern and count microscopic nematode eggs. *Sci Rep* 8, 9145. <https://doi.org/10.1038/s41598-018-27272-w>

Beaudelaire Saha Tchinda, Michel Noubom, Daniel Tchiotsop, Valerie Louis-Dorr, Didier Wolf (2019). Towards an automated medical diagnosis system for intestinal parasitosis. Informatics in Medicine Unlocked, Volume 16, 2019, 100238, ISSN 2352-9148, <https://doi.org/10.1016/j.imu.2019.100238>.

Children Without Worms (27 May 2021). "Innovative Technology to Address the Elimination of Morbidity of Soil Transmitted Hemelinth."<https://gatech.instructure.com/courses/214292/files/24538503?module_item_id=1970154>

Children Without Worms (2021). "Welcome Georgia Tech BME Students!" <https://gatech.instructure.com/courses/214292/files/24797641?module_item_id=1982086>

Hart, A., Jose-Maria, N., Tanguturi, S., Jayanthi, R., & Shalom, E (December 2022). Intestinal Worms Final Technical Report. Google Doc. <https://docs.google.com/document/d/1RYWKkyzSzECGPsMKJpO_tUcqUoXJWaytBaS01V0QSjg/edit?usp=sharing>.

Jose-Maria, N., Hart, A., & Tanguturi, S. (2022). GTVIPFall22 STH ML Notebook. (Version 7). Kaggle. <https://www.kaggle.com/code/nathalyjosemaria/gtvipfall22-sth-ml-notebook?scriptVersionId=111027532> Adapted from Ward et. al, 2022b.

Ward P, Dahlberg P, Lagatie O, Larsson J, Tynong A, et al. (2022) Affordable artificial intelligence-based digital pathology for neglected tropical diseases: A proof-of-concept for the detection of soil-transmitted helminths and *Schistosoma mansoni* eggs in Kato-Katz stool thick smears. PLOS Neglected Tropical Diseases 16(6): e0010500. <https://doi.org/10.1371/journal.pntd.0010500>

Ward P, Dahlberg P, Lagatie O, Larsson J, Tynong A, Vlaminck J, et al. (2022). AI4NTD KK2.0 P1.5 STH & SCHm. (Version 6). Kaggle.<https://www.kaggle.com/code/peterkward/ai4ntd-kk2-0-p1-5-sth-schm#Explore-the-data-sets>