MISI Interns Git Team

Philip Roberts, Walker Buckle, Anastasia Lindquist, Kieran Garner,
Jamell Battle, Sam Thoguru, Cooper L, Cameron S

07/27/20

Contents

Abstract

This paper discusses the work done by the interns at Dreamport MISI on Machine Learning. We were given the goal of improving a machine learning algorithm that could test a patient for COVID-19 based on a computed tomography scan of the patient's lungs.The paper focuses on the Git team, or team three, who had the responsibility of managing the GitHub repository for all the teams. This involved creating and merging branches, writing Git hooks to format our python code, writing documentation for our repository in our README, and assisting other teams in their usage of Git.

1 Introduction

As part of our internship at Dreamport MISI[1], we worked on a Machine Learning project that can test if a patient is COVID-19 positive based on a CT scan. Our repository is forked from a machine learning project done by UCSD[2]. We managed the Git repository for all of the interns in addition to writing the README for the repository[3], a text file that gives information about a repository, and how to use it.

2 Git Collaboration

As part of our Machine Learning project, we used the Git version control system to enable simultaneous development between teams. We used GitHub to host our git server. We created several branches of the master branch for each of our development teams to allow them to work separately. Our team assisted other teams with interacting with the repository, in addition to merging branches with the master branch. In addition to managing the versions of our software, we created the README file with information about who we are, what the repository is for, and how to use it.

3 Git Hooks

Git Hooks are what allowed us to keep formatting pretty and keeping our code up to the standard we needed when collaborating. We use a pre-commit hook and autopep8 in order to help keep our code within the standardized PEP 8 format. This kept our code easy to read and helped find issues along the way, especially when merging branches into the master.

[1] https://www.misiacademy.tech/
[2] https://github.com/UCSD-AI4H/COVID-CT
[3] https://github.com/walkerjbuckle/COVID-CT-Starlight-Saviors/blob/master/README.md

4 README.md

In addition to managing the versions of our software, we updated the README file with information about who we are, what the repository is for, and how to use it, and citations from the dataset we used for the CT scans.

5 Conclusion

Through this project, we learned how to interact with Git to collaborate with multiple development teams. We also wrote a Git pre-commit hook to keep our code formatted consistently. We wrote the README file for our repository and an installation script to streamline the process of installing our git hooks and dependencies.

6 References

Soares, E., Angelov, P., Biaso, S., Higa Froes, M., & Kanda Abe, D. (2020). Sars-cov-2 ct-scan dataset: A large dataset of real patients ct scans for sars-cov-2 identification. medRxiv. Retrieved from https://www.medrxiv.org/content/early/2020/05/14/2020.04.24.20078584 doi: 10.1101/2020.04.24.20078584