lo*SE 491 WEEKLY REPORT x*
11/01/2024 - 11/07/2024

Group number: sdmay25-30

Project title: *Explainable AI for source code applications*

Client &/Advisor: *Arushi Sharma*

Team Members:
Manjul Balayar
Kellan Bouwman
Sam Frost
Akhilesh Nevatia
Ethan Rogers

Weekly Rotation
Manjul Balayar
Kellan Bouwman
<mark>Sam Frost</mark>
Ethan Rogers
Akhilesh Nevatia

Weekly Summary

This week was an implementation sprint, with a focus on progressing the tasks specified on the issues board. These issues being Unit Testing, Pipeline Integration, HPC Start Auto Labeling, Activation Expansion, Interactive Clustering Visualizations, Pipeline Automation Configuration, Clustering Analysis and Expansion. Each member was assigned to one or more tasks to work on over the sprint. At the conclusion of the sprint, we met with our advisor to ensure our work aligns with the needs of the user base.

o **Past week accomplishments**

Manjul Balayar:
- Worked on implementing parallel corpus into NeuroXCode
- Implemented awesomealign and fastalign to do manual test on them

Kellan Bouwman:
- Focused on adding visualization modules
- Updating pipelines and workflows
- Testing files
- Updating setup.py
- Updating Readme

Sam Frost:
- Added the main config (src/NeuroXCode/conf/config.yaml)
- Added more config files in accordance with directory structure
- Further research into Hydra

Akhilesh Nevatia:
- Further implemented the Birch and Clustering Algorithm
- Debugged error where clusters were not getting saved
- Included Birch and Clustering Algorithms in the Clustering NeuroXCode Test Script
- Started working on integrating extracted activations into clustering, switching away from randomly generated files function
- Updated Issues and made a draft Merge Request

Ethan Rogers:
- Gained a good understanding of what DSPY has to offer and how it related to various facets of our project
- Became well acquainted with data available in a reference HPC ontology repository
  - Looked into means of using this to expand our dataset list

o **Pending issues**

Manjul Balayar: N/A
Kellan Bouwman: N/A
Sam Frost: N/A
Akhilesh Nevatia: N/A
Ethan Rogers: N/A

Individual Time Contributions

Name	Hours This Week	Total Hours
Manjul Balayar	6	47

Kellan Bouwman	6	54
Sam Frost	6	47
Akhilesh Nevatia	8	49
Ethan Rogers	6	49

○ **Plans for the upcoming week**

Manjul Balayar:
- Work on publishing and initial version of NeuroXCode package
- Add documentation

Kellan Bouwman:
- Integrate everything into main
- Work on design
- Update based off of feedback from customer

Sam Frost:
- Finish config files for each subdirectory
- Take feedback from advisor meeting and adjust
- Automate unit tests

Akhilesh Nevatia:
- Work with Ethan to finish Clustering, Activation Integration
- Work on Clustering Purity Test
- Update Readme and analyze which Clustering is doing best based on purity and other mentioned test
- Help Kellan out with the visualization of the clustering algorithms

Ethan Rogers:
- Using example input/output data, use DSPY to improve auto-labeling capabilities
- Bring together the package into one whole piece that is cohesive and connected
- Document modules that I have worked on

Summary of weekly advisor meeting

The meeting began with each member sharing what they worked on during the sprint, and clearing up any questions that came up. Demos were done to show what progress has been made. The goal that was defined for next week is to integrate everything from the scripts folder in CodeConceptNet into a formal library. Before publishing, we need to make sure everything works. Arushi wants this done so that we web app team can use what we've developed. Clustering, and documentation should also be worked on. All the stuff is there, we just need to put things together.