

**SE 491 WEEKLY REPORT 9**  
**11/07/2024 - 11/14/2024**

**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*

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

**Weekly Summary**

This week we shifted our focus toward flushing out our previously created pipeline by making it much more cohesive with only requiring a few inputs from the user. With this, we will be updating our previously instantiated package. Additionally, we are working on flushing out the documentation for the project's various components on its github. There should be enough guides considering enough use cases that the expected user has few issues (if any).

○ **Past week accomplishments**

Manjul Balayar:

- Implemented alignment CLI
- Changed clustering output format to match with alignment
- Finishing clustering-alignment pipeline workflow

Kellan Bouwman:

- Pushed Visualization code for interactive looks at clusters
- Implemented the auto deploy pipeline for package registry
- Started looking at HPC integration
- Worked on implementing pipeline implementation
- Evaluation of Scripts from CodeConceptNet to be turned into work flow pipelines
- Worked on readme for both the NeuroXCode library and for sdmay25-30 repos.

Sam Frost:

- Finished all initial .yaml files for pipeline
- Began integrating Hydra functionality with clustering algos
- Updated senior design website, which had not been updated

Akhilesh Nevatia:

- Linked Hyperbolic Clustering and Birch Clustering to be command line callable as well
- Mainly worked on Integrating Clustering and Activations Workflow which is almost there, Clustering works on the codebert model's processed activations, where the user can run it by mentioning the layer he wants to form clusters for

Ethan Rogers:

- Worked on creating a seamless experience for the user by automating the directory system using a default case unless otherwise specified by the user
- Worked with Akhilesh to provide the information necessary to allow clustering scripts to take the output files from processing activations
- Moved activations scripts to being terminal-executable using standard argparse

○ **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	53
Kellan Bouwman	7	61
Sam Frost	7	54
Akhilesh Nevatia	8	57
Ethan Rogers	8	57

#### ○ **Plans for the upcoming week**

Manjul Balayar:

- Implementing newer alignment lexical from CodeConceptNet over to our new library
- Making sure our new integrated pipeline package works in CodeConceptNet

Kellan Bouwman:

- Pipeline Automation and Testing
- CLI run options
- Pip install pipeline testing for colab
- Verification of pipeline results

Sam Frost:

- Finish integrating Hydra with clustering
- Expand to other subdirectories in NeuroXCode as able
- Flesh out senior design website further, in preparation for panel

Akhilesh Nevatia:

- Work on Clustering Analysis, and comparing Evaluations and Heuristics of different Clustering Algorithms
- Assist on completing the Integration for the full pipeline from activations to labeling and ideally publish an updated package with the same
- Get a Colab notebook running with a live and easy example of the same
- Documentation on running clustering ( CLI & Import Wise Usage )

Ethan Rogers:

- Make sure the entire pipeline is easily runnable by a user without much knowledge of its internals
- Create documentation for all files I am responsible for
- Work back toward integrating DSPY into an LLM for automated dataset labelling to provide more (and better) datasets

### **Summary of weekly advisor meeting**

We have not had our weekly advisor meeting at the time of this report due to a meeting scheduling change to the coming weekend. We expect the meeting to consist of us presenting our

altered work extended into her other repository and receiving feedback from our client as to how we can better fit the integration needs into the other repository. We have kept in contact with our client, though, sharing packages as we iteratively develop packages and offering insight into the added functionality.