

EE/CprE/SE 491 WEEKLY REPORT 2

9/19/2024 – 9/26/2024

Group number: sdmay25-proj031

Project title: Explainable AI for source code applications

Client &/Advisor: Arushi Sharma

Team Members/Role:

Manjul Balayar

Kellan Bouwman

Sam Frost

Akhilesh Nevatia

Ethan Rogers

Weekly Rotation
Manjul
Kellan
Sam
Ethan
Akhilesh

- **Weekly Summary**
- **Past week accomplishments**

Manjul Balayar: Set up and reviewed NeuroX and CodeConceptNet repositories, explored the NeuroX tutorial notebook and went through alignment markdowns to understand the assigning labels to cluster methods.

Kellan Bouwman: Tested repos, ran example code, tested new installation method, edited readme for new instructions, and began to refactor metrics. +9+

Sam Frost: Acquired pronto access for group, experimented/learned CodeConceptsNet and Neurox repos

Akhilesh Nevatia: Tested out running jobs on pronto, went through getclusters.sh on CodeConceptNet and transferred part of it to the get_clusters.py file
Ethan Rogers:

Individual Time Contributions

Name	Hours This Week	Total Hours
Manjul Balayar	6	12
Kellan Bouwman	6	15
Sam Frost	6	12
Akhilesh Nevatia	6	12
Ethan Rogers	5	11

Plans for the upcoming week

Manjul Balayar:

- Set up pronto
- Fix the alignment issue and refactor to NeuroX
- Get better understanding of the pipeline

Kellan Bouwman:

- Pronto set up testing
- V2 code refactoring
- Starting scaling research

Sam Frost:

- Clean up scripts/wikis/READMEs
- unit tests where able
- CI/CD implementation where able

Akhilesh Nevatia:

- Go through temp folder and scripts on Neurox
- Go through Readme better and understand pipeline workflow
- Test out get_clusters.py (NeuroX) vs getclusters.sh on CodeConceptNet
- Go through hyperbolic clustering and attachments

Ethan Rogers:

- Integrate drafted activations-related functions & refactoring via pull requests
- Learn and implement bash script
- Expand to more activations and pipeline related pieces for refactoring

Summary of weekly advisor meeting

This week, we reviewed the progress on tutorials, alignment, clustering, and activations for

the project. Moving forward, our priority for next week includes labeling, alignment, configuration, adding clusters, hyperbolic clustering, and activations, with the aim to run the full pipeline. Refactoring code, cleaning up the README, and achieving a unified starting point for the new package/library remain key goals.

Additionally, the following tasks are on the TODO list:

1. Create a directory structure suitable for scaling the entire pipeline. While extracting activations is already implemented temporarily, we need to add processing for activations, clustering (with results from all layers to visualize in the web app), and evaluation. Proactive questions are encouraged.
2. Complete refactoring of process activations, clustering, labeling (Gemini), and alignment.
3. Investigate alignment issues in more detail, particularly focusing on why results are subpar and why obvious clusters are not being labeled.
4. Clean up the repository, add unit tests, CI/CD, documentation, and create a fully functional Python package for importing into CodeConceptNet. Test the entire (scalable) pipeline in a temporary setup with a clear README and directory structure details. Once the pipeline is ready, models and datasets will be provided to scale and evaluate. Ensure web app compatibility.
5. Test hyperbolic clustering with our dataset on Pronto, ensuring it's formatted in the standard clustering format.

In terms of next steps, we need to:

- Add cluster purity and other standard clustering evaluation metrics to Metrics.py using sklearn.
- Update the Colab notebook with the new working imports.

Finally, diagrams like UML or flowcharts will be helpful to visualize and clarify the steps, and these should be added to the project wiki. Also, please update the Google Doc by removing outdated information and adding relevant updates.