Senior Design Bi-Weekly Status Report

Team: sdmay21-35
Team Email: sdmay21-35@iastate.edu
Team Members:
      Ahmad Alramahi - Lead Developer
      Austin Boling - Meeting Facilitator
      Joseph Naberhaus - Project Lead
      Ekene Okeke - Report Coordinator
      Ethan Ruchotzke - Documentation Manager
      James Taylor - Linguistics SME

Project: POS Tagger for Software Documentation
Report Period: Feb 8th - Feb 22nd

Summary of Progress in this Period

The objective of this week was to return to working condition and plan ahead for the coming weeks. The major steps remaining in the data pineline are to tag data (automatically and manually), and test & train the model.

During the week, the group planned ahead and scheduled the bi-weekly module calendar we will be using for the tokenizer and data pipeline. In addition, more refinement work was done on the automatic tagging system. Finally, the group divided into groups for working on tokenization, manual training, and tag system finalization.

Meeting Notes

- SSH Login Issues
  - Might be down for maintenance
  - Check daily to try and log in - reach out if not done by wednesday
- Autotagging
  - Expect two weeks of manual tagging work
  - Follows work on tokenizer
- Tokenizer
  - Direct focus on a 2 week workflow for the tokenizer
  - Everyone interested can help, main "leaders" will be James & Ethan

Pending Issues

- Refinement of the automatic tagging pipeline
  - Inaccuracies / room for extra automatic tagging
  - Minimization of manual tagging is a must
- Tokenizer
  - 2 week schedule
  - Ethan & James will be working on this
  - Implement the "middle" of the data pipeline (turning data into tokens)
- Finalization of Tags
  - Complete the set of tags to use in the pipeline
    - List of used english tags
    - List of new software tags

Plans for upcoming Reporting Period

Our plans for this upcoming reporting period **(Feb 8th - Feb 22nd)** are broken down as follows:

Who	What	Due When
All (Led by Austin)	Finalize Tag Set	Two weeks from 8th February (meeting)
Ethan & James	Implement the Tokenizer	Two weeks from 8th February (meeting)
All	Manual Data Tagging	Four weeks from 8th February
All	Setup Training Pipeline (Schedule out)	Unknown
Ethan	SSH into GPU VMs	Wednesday, February 10th
	.	