EE / CprE / SE 492 Weekly Report 02: SDMay21-29
Intelligent Code Editor

Feb. 9 - Feb. 22
Client: Hung Phan
Advisor: Dr. Ali Jannesari

Team Members & Roles:

Evan Christensen - Meeting Scribe
Ben Gonner - Report Manager
Jacob Puetz - Chief Engineer Software Systems
Jordan Silvers - Meeting Facilitator
Cory Smith - Test Engineer

Weekly Summary:

In the previous week, we officially switched our focus from a direct translation GNN to a classification GNN. We are working off of the TensorFlow-based GNN provided to us by our client, along with the SPoC dataset from our previous attempts. Research began on the development of plugins for Visual Studio. We are continuing to work on the preprocessing of the SPoC data because the GNN we are using has a very specific set of input requirements for its datasets. Finally, we have also begun working on exporting a reusable model from the GNN so that training does not have to be repeated for every classification.

Past Week Accomplishments:

- Model Exporting and Reuse - Cory
    - The GNN comes with a function to save a model to a file and restore it later
    - The GNN, as it was given to us, already creates two files from the training process: a doc_vectors file and a word_vectors file
        - Both files include data from the embedding process
        - These files don't come with any information about how to use them, so it's mostly trial and error
        - This is the difficult part of the process at the moment
    - The model object has a predict function (unused in the GNN itself), which should take in a vector of some form and return a vector which contains the probabilities of different classifications

- Visual Studio Plugin Research and Development - Jordan/Ben
    - Researched on how to make a plugin for Visual Studio
        - Tutorials such as "Hello World" and "Menu Command" were completed from given tutorials
    - Research on C# and XML syntax for plugin development
    - Research on on how to do certain tasks in the plugin:

- ■ Reading in a line of code at a time
- ■ Changing a line of code
- ■ Adding Command to Text Editor Context menu (Right Click menu)
- Training data pre-processing - Jacob/Evan
    - Meet with Cory to better understand the format for input data
        - ■ Talked about the various categories the data needs to be broken into
        - ■ Learned what info need to be in all of our training files
    - Started work on dividing the data up
    - Scripts to sort the data into their respective categories are in the works
    - May need to inquire about index files with advisor

Individual Contributions:

Name	Contributions	Hours This Week	Total Hours
Evan Christensen	Training data pre-processing	4	8
Ben Gonner	Visual Studio Plugin	4	6
Jacob Puetz	Training data pre-processing	4	8
Jordan Silvers	Visual Studio Plugin R&D	2	8
Cory Smith	Model Exporting and Reuse	4	14

Upcoming Plans:
- Determine how to insert data into a loaded model
    - This is mostly about figuring out how to use the word and document embedding files correctly.
- Continue work on the Visual Studio Plugin
    - Set up a command in a context menu for the text editor that can take a line of text and replace it with something else.
- Develop scripts to sort large dataset into categories
    - There are around 300,000 lines in the dataset, finding ways to sort them all into categories may take some time.