EE / CprE / SE 492 Weekly Report 01: SDMay21-29
Intelligent Code Editor

Jan. 25 - Feb. 8
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 continued our attempts to create a working GNN to do our translation from pseudocode to code. Progress was made around the formatting of our input training data, in the sense that it can now take in a full set of node attributes, edge attributes and an adjacency matrix. Additionally, the model is almost completely set up, although issues remain in the connection between the ECC layer and the decoding LSTM layer. In the previous weeks client meeting, we also discussed the possibility of shifting our efforts from trying to create a true translation GNN to creating a text classification GNN due to the fact that there are not enough resources available to create a good translation GNN within a reasonable amount of time.

Past Week Accomplishments:

- Translation GNN - Cory
    - Developed parsing program to split text into attribute and adjacency matrices for input into the GNN
    - Connected input, embedding, and LSTM layers to the ECC layer
        - ECC layer output has shape (X, 5, 256), but the decoder LSTM needs an input of shape (X, 10, 256), adjusting the input shape might be able to fix this
- Classification GNN - Cory
    - Read over both papers that were provided by the client
        - The major difference between the 2 options appears to be that one uses Keras while the other uses PyTorch
        - The PyTorch GNN seems to require a specific GPU driver to run, which may make it less practical
        - Keras uses a higher level API, which is better for our group given our limited experience with neural networks
    - Began examining the Keras classification GNN
- Prototype on Unsupervised MT(unsuccessful) - Jordan
    - Researched Unsupervised Machine Translation Using Monolingual Corporal Only

        - Read the designated paper from the client and advisor (along with my own additional research)
        - Tried to create a working prototype but failed due to minimal knowledge on the subject
        - Do not believe going in this direction will help overall project and will switch my work efforts
- Research on Classification GNN vs. Translation GNN and datasets - Evan
    - Read over papers provided by client
        - New translation GNN will likely require use of external library other than original plan of Spektral if we choose that route
        - Looking to Cory for specifics, as GNNs are his area of expertise
    - Studied new dataset provided by client found on https://sumith1896.github.io/spoc/
        - Last semesters dataset was too small, new dataset is much larger and should offer a notable increase in translation/classification accuracy
        - New dataset is primarily in pseudo code to C++, original plan was to translate natural language to Java, client agrees pseudo code to C++ is agreeable
        - Formatting for insertion into GNN may be necessary, looking into now

Individual Contributions:

Name	Contributions	Hours This Week	Total Hours
Evan Christensen	Research on Classification GNN vs. Translation GNN and datasets	4	4
Ben Gonner	Reading and SPoC exploration (WIP)	3	3
Jacob Puetz	Communicated with the team on the formatting preferences for the data set.	4	4
Jordan Silvers	Unsupervised Machine Translation	6	6
Cory Smith	Translation GNN, Classification GNN	6	6

Upcoming Plans:

- Formatting SPoC dataset to fit into either the translation GNN or the classification GNN
    - This dataset is much more detailed than the one that the previous group used, but its formatted as a translation from pseudocode to C++, which may not work for either of our possible GNNs

- We can either find a way to make a graph out of the pseudocode for the translation GNN, or we can make a parser to determine the classifications for the classification GNN