1. Topic & Research Question:
2. TOPIC: Text Vs Trees Vs Graphs: Deep Learning Techniques for Program Understanding.
3. RESEARCH QUESTION: What techniques are best for properly processing and understand source code in the way that NLP models process and understand text and natural language.
4. Scope of Review and Background of the Project (Sources and motivations):
5. Papers explaining how NLP models are built, how they work and how they process and understand text.
6. Explanations about why common Deep Learning and NLP models are not suitable for processing source code (context, control flow, paradigms, etc.).
7. Papers explaining what Graph Neural Networks are, how they are built and how they work.
8. Papers explaining what Gated Graph Neural Networks are, how they are built and how they work.
9. Papers explaining what Tree-Based Convolutional Neural Networks are, how they are built and how they work.
10. Data Sources for Training and Testing:
11. All data will be of a specific type: Sorting algorithms written in Python.
12. Due to the nature of the NN models, each python file will only contain one function.
13. Each file will have a label associated with it indicating what type of sorting algorithm it is (bubble, merge, etc.) for classification purposes.
14. To increase the accuracy of the results, the testing data will contain the same number of files as the training data.
15. Conducting the search for appropriate literature.
16. Reviewing the literature and making notes.