Weekly Report Week 1(01/08 – 05/08)

Project: Development of security plugins for IDEs

Tasks done:

- I completed the environment set-up for LineVul repository on my google collab.
- I have read through(briefly) the paper (LineVul: A Transformer-based Line-Level Vulnerability Prediction)
- I have successfully ran the model and generated the RQ1(inference) results for:
  - LineVul

```
INFO - __main__ -   ***** Running Test *****
INFO - __main__ -     Num examples = 18864
INFO - __main__ -     Batch size = 256
INFO - __main__ -   ***** Test results *****
INFO - __main__ -     test_accuracy = 0.9909
INFO - __main__ -     test_f1 = 0.9142
INFO - __main__ -     test_precision = 0.9712
INFO - __main__ -     test_recall = 0.8635
INFO - __main__ -     test_threshold = 0.5
```

  - BPE+No Pretraining+BERT

```
INFO - __main__ -   ***** Running Test *****
INFO - __main__ -     Num examples = 18864
INFO - __main__ -     Batch size = 256
INFO - __main__ -   ***** Test results *****
INFO - __main__ -     test_accuracy = 0.8087
INFO - __main__ -     test_f1 = 0.0758
INFO - __main__ -     test_precision = 0.0519
INFO - __main__ -     test_recall = 0.1403
INFO - __main__ -     test_threshold = 0.5
```

  - Word-Level+Pretraining(Codesearchnet)+BERT

```
INFO - __main__ -   ***** Running Test *****
INFO - __main__ -     Num examples = 18864
INFO - __main__ -     Batch size = 256
INFO - __main__ -   ***** Test results *****
INFO - __main__ -     test_accuracy = 0.9002
INFO - __main__ -     test_f1 = 0.0646
INFO - __main__ -     test_precision = 0.0678
INFO - __main__ -     test_recall = 0.0616
INFO - __main__ -     test_threshold = 0.5
```

  - Word-Level+No Pretraining+BERT

```
INFO - __main__ -   ***** Running Test *****
INFO - __main__ -     Num examples = 18864
INFO - __main__ -     Batch size = 256
INFO - __main__ -   ***** Test results *****
INFO - __main__ -     test_accuracy = 0.9351
INFO - __main__ -     test_f1 = 0.0208
INFO - __main__ -     test_precision = 0.0667
INFO - __main__ -     test_recall = 0.0123
INFO - __main__ -     test_threshold = 0.5
```

- The pretrained model was run for RQ2 and I have captured the following results:

```
'codebert_attention_top10_accuracy': [66.0], 'codebert_attention_ifa': 4.54
```

- The pretrained model was run for RQ3 and I have captured the following results:

```
INFO - __main__ -   total functions predicted as vulnerable: 938
INFO - __main__ -   total predicted vulnerable lines: 25978
INFO - __main__ -   total lines: 504886
INFO - __main__ -   total flaw lines: 3966
INFO - __main__ -   total flaw lines in predicted as vulnerable: 2664
INFO - __main__ -   top0.2-Effort: 0.0072
INFO - __main__ -   total inspected line to find out 0.2 of flaw lines: 3643
INFO - __main__ -   top0.01-Recall: 0.2509
```

- I replicated the cppcheck results.

```
Top-1 Accuracy: 0.0732484076433121 (46 / 628)
Top-10 Accuracy: 0.15445859872611464 (97 / 628)
IFA: 21.598591549295776
Effort@20Recall: 0.12920353982300886
Recall@1LOC: 0.03919372900335946
```

- My code to run the model is in google collab and the link to access it is: https://colab.research.google.com/drive/1r8swy3jugh2dmlXC9cCl8Qte6jPI8rYa?usp=sharing. After final clean up, I'll be uploading it in the GitHub repo.

Challenges/Lessons done:

- Got a brief idea about the LineVul project.
- Got to learn new google Collab commands.
- I faced google drive quota exceed issue, but found a workaround for that as well. By uploading it on dropbox and another gdrive account.

- I tried running retrain model, but its computational costs are high and couldn't run it completely. We might have to run it on high-speed computer. I have requested for phoenix computers access, once I get that I can try running the model.
- The same issue was seen in RQ2 & RQ3. I got the initial results for the model but couldn't run it successfully and got the below error:

```
RuntimeError: CUDA out of memory. Tried to allocate 384.00 MiB (GPU 0; 14.76 GiB total capacity; 13.59 GiB already allocated;
```

  I will try finding the workaround to resolve the error next week. But having access to phoenix computer might help speed up the process.
- One of the challenges I'm facing is, after sometime google Collab keeps loosing connection with kernel and I have to run all the code parts to restore the project data. I will check if I can use other platforms to run the model.

Tasks for next week:

- I will be continuing the task of reading the paper (LineVul: A Transformer-based Line-Level Vulnerability Prediction).
- Debug and fix *CUDA out of memory* error and successfully run model for RQ2 & RQ3.
- I will analyse furthermore the results obtained from running the model for
  - RQ1
  - RQ2
  - RQ3
- Understand how to obtain function level vulnerabilities.