Weekly Report ---*Friday, 19 March 2021*

Student Name: Siqi Sun

Student ID: a1752383

- What I have done

1. Created a seperate branch called 'Siqi' from main branch on Github so that I can continue submitting my work afterwards.
2. Created four pages under Github wiki page (Home page, Agenda page, Diagrams page and Meeting Summaries page) and add a side bar to organise them.
3. Read the given paper from client and online researched the mentioned methods.
4. Attended group meetings before today's client meeting and discussed a series of questions as followed :

- What's the expected results of the project by the end of this semester?
- What's the suggested milestones of the project?
- Do you have any suggestions on where we get started?
- Where can we get the skills in C/C++ software programming cause we all know python has the natural advantage in machine learning and deep learning fields.
- How to determine the performance of the model?
- How to detect the vulnerability of the data input?
- Will there be 2 main modules for this project: Firstly, detection method algorithm Secondly, Evaluator (used to evaluate the the vulnerability detection method) ?
- Where can we find the sample sources?

5. Wrote a meeting summary for today's meeting.

- What I plan to do

1. Do the online research about data sources e.g. NVD (The National Vulnerability Database) and CVE (Common Vulnerabilities and Exposures).
2. Research on the existed models such as TextCNN and BiLSTM.
3. Hands on other open source relevant GitHub repositories and focus on deep learning models.
4. Learn how to convert the Python code to C/C++ code within the topic deep learning.
5. Get familiar with the approach for Cross-Project Defect Prediction (CPDP).
6. Learn to extract features from labelled data bases and source code within the domain of data science.

- the issue I've encountered

1. Don't know how to sift an appropriate cross-domain vulnerability detection method and evaluate their performances.

   Details: In this semester, we are expecting to research and develop a benchmark system that could visually rank the performances of different models. Also, explain why the benchmark system is convincible.
2. Lack of the experience of deep learning study.

   Details: haven't use training models before and unfamiliar with the procedures of data training.
3. Don't know the detailed procedures of knowledge transferring.

   Details: Knowledge transfer is all about convert and combine different classifiers from multiple projects.
4. Don't know how to define a series of labelled data is sufficient or not.

   Details: get the knowledge of 'over-fitting' and 'under-fitting' theory after group discussion.