WEEKLY INTERNSHIP REPORT

3rd July – 10th July

HUYNH THANH QUAN

# Plan of this week

* Setting up PC and creating accounts
* Self-introducing to division
* Writing a Python script to collect records on GitHub repository
* Storing records into MongoDB
* Assessing the commit is bug or non-bug

# Achievement of this week

## Tuesday

* + Attended the first Sprint Planning of the team
  + Given the Proof of Concept project, this project is in prototype stage

## Wednesday

* + Created Outlook, Taiga, and Mattermost account
  + Set up proxy account
  + Changed the language interface of Outlook, Windows and Microsoft Office to English
  + Designed self-introduction presentation and self-introduced to division
  + Installed Anaconda 5.2 for Python 3.6 and necessary libraries
  + Paired programming with Kubota-san

## Thursday

* + Continued to pair-programming with Kubota-san
  + Searched bug-fix commits that contain “FIX” in commit message by using GitPython library. For each bug-fix, searched its parent commit and store them to evaluate as bug commit or non-bug commit later
  + Attended the Japanese training class

## Friday

* + Wrote Python script to fetch commit data. Collect those information from each commit such as
    - number of files
    - number of lines
    - number of modified lines
    - committer
    - date
  + After that, all those information are converted in JSON which is called “record”. And store them in RxJavaRecord collection of CommitInfo database in MongoDB server (Kubota-san’s virtual machine)
  + Ran the virtual machine to fetch more than 5000 commits in the weekend.

## Monday

* + Divided the total of records into 4 fractions to label the commit as bug or non-bug (150 commits for each person)
  + Evaluated those commits and labelled them in MongoDB server

## Tuesday

* + Attended the Sprint Review and Retrospective
  + Filtered the unduplicated records in MongoDB

# What I have learned

In summary, I learned the following things in the first week:

* How to use GitHub, GitLab, Taiga, Mattermost in Web Browser
* How to organize the task in Agile and Scrum process
* How to use Git API on Python
  + Fetch commits data to get information (committer, date, number of files…)
  + Trace the difference lines between the current commits and its parents
* How to use GitBash
  + Create new branches
  + Add and checkout new changes to commit
  + Commit the change
  + Push commit to GitHub
* How to manipulate data on MongoDB on Virtual Machine
  + Insert one or many records into database
  + Update columns value into existing data
  + Add new columns into existing data
  + Backup and remove duplicated data

# Impression

My first impression is that our team celebrated the welcome party for me and many invitations for me to explore and visit the Japan country.

Secondly, the working environment in FUJITSU is extremely friendly and fun. Akasofu-san, Sakai-san, Kubota-san, Oyu-san, Bunta-san Morishima-san and many others are willing to help me when I have problems. Sometime, they can’t express their words in English but they have tried to express in many different ways to help me.

I really appreciate very much about this. This makes me easier to work together efficiently.

Thank you very much

# Plan of next week

* To collect more features such as “insertion lines” and “deletion lines”
* To normalize the range of the value of each feature
* To divide the training data into training set and test set (possible validation set)
* To learn and test the machine learning model
* To prepare the presentation on 18th July and the business trip to Numazu on 20th July (next Friday)
* To prepare the interview with HR on 12th July (this Thursday)

HUYNH THANH QUAN