# Discussion 2018-02-01

A kick-off presentation was given for both sides of the graduation project with Peter and Joaquin. Unfortunately Dragan could not be reached by skype, but will be caught up when he is back. The discussion was based around what I have done for the project, what I was working on at the moment and what I planned on doing in the future.

The main points that became clear:

* Create one framework for use. This framework could be seen in two ways: a flow chart (figuratively) and as a combination of tools.
  + The flow chart should result in a GUI that makes everything possible for use.
  + The combination of tools should have tackle data within different boundaries. The aspects discussed should show these boundaries how well it would work.
  + Right now a start will be made with Python, using TPOT and scikit-learn.
* Testing and trying things out is smart to continue the project with. From these testing different parts should be continued to find better results. This will be done by also meeting with Felix Garza to have her expertise on the skin disease data set.
* During the project also other existing tools will be tested for comparison
* Discussion with the responsible people will be held to find out if a separate report is needed for both masters.

**Timetable specific**

Due to the discussion about the time-table it has been changed a little bit. Some points

* A meeting with report and presentation will be held every time at the end of the period I have been working in.
* More frequent meeting will be held with Dragan and Joaquin to discuss advancements in the project.
* Discussion with the responsible people will be held to find out if a separate report is needed

**Updated time table**

|  |  |
| --- | --- |
| **Month (36 weeks, 8 ½ month)** | **Focus** |
| February 5st – March 2nd (4 weeks) | * Finish literature research:   Data Mining side (TPOT/Other data mining tools)  Last fixes   * Testing Skin disease data sets problems for the framework (statistical analysis) |
| March 5nd – April 20th (6 weeks) | * Start focus on analysis high dimensional data (Psoriasis Data sets) * More data set gathering (open source?) * Trying out tools * Possible tool extension |
| April 24th – June 1st (6 weeks) | * Focus on analysis low quality data (error handling, missing value handling) * Relevant data set gathering (open source?) * Trying out tools * Possible tool extension |
| June 4st – July 13th (6 weeks) | * Focus on low sample number data (Psoriasis data sets, sample generation) * More data set gathering * Trying out tools * Possible tool extension |
| July 16th – August 31th (7 weeks)) | * Holidays (half a month) * “Free time” Extension other aspects or starting new aspect |
| September 3rd – September 28th (4 weeks) | * Write and hand in report (possibly joint report with separate sections) |
| October 1st – October 19th (3 weeks) | * Hand in report (together/separate?) * Give presentations (most likely separate) |