**Which classes to select**

You can work with all 6 classes in the data, select some of them, or group some of them (e.g. healthy/not) - of course with motivation. This is part of the data exploration in task 0 that was at the start of the project. You can also choose to filter out some images which are too low quality for example.

**Feature extraction**

Your feature extraction algorithm does not need to return binary values, you can use continuous values in your classification step.

You don't have to have a single value for each feature, for example if you want to look at multiple aspects of color, you can have features color1 and color2. You might end up choosing one or both them, based on their correlations with other features, how much they contribute to the performance, etc.

**Evaluation with external data**

Please have a look at the Github repository <https://github.com/vcheplygina/fyp2023> and the file 3\_evaluate\_classifier.py to get an idea of how the TAs will evaluate your methods on external data. In short you need to provide an extract\_features function, and a saved classifier that outputs probabilities, for the classes you have chosen (6 classes, healthy vs not, etc).

**Using meta-data**

If your classifier uses some of the metadata, please ALSO submit a classifier that uses the mask- and image- only, because not all the external test files will have all the metadata available. So for example your submission could have the files group7\_classifier\_allfeatures.sav and group7\_classifier\_imageonly.sav

**Should we do more for XYZ?**

Try to have something for each task first so you have a complete project. Then if you have time, work to extend a more specific part.

**Report structure**

Please revisit the slides of the last lecture on writing, you can use either the Introduction/Methods/Results/Discussion format, or Introduction/Data exploration/Feature extraction/Classification/Discussion format. Your open question can be an additional section (for example after Classification), or a subsection (for example a subsection under Feature Extraction if you investigated different features).

**Figures, tables**

Your plots and tables should be produced via code and not by copy pasting things into Excel for example. It is OK for me if you use R code for creating this, instead of Python, just add it to your Github repository.

**Questions before hand-in**

On Thursday is the last session with the TAs, use it! If you have questions afterwards **please post them on the LearnIT forum.**

**Exam**

The exam is similar to the exam of project 1 but we have a bit more time. The goal is to figure out if you have achieved the learning objectives (see course description). I want to see if you understand the methods we studied, why/why not to apply them, what do you expect to happen if you instead used another method, etc. It is not sufficient to only know the definition.

**Exam presentation slides**

You can still update plots etc in your presentation after the hand-in. It is also fine if you notice you made a mistake in the report, you can say this in the presentation and show the correct result instead. You cannot submit an empty PDF and then present everything at the exam :)

**Exam time**

As Louise said in another post, once we know how many hand-ins there are, she will post the exam schedule.