

# Big Project

## (50.039 Deep Learning, Y2021)

---

Matthieu De Mari, Liu Jun

### 1. Introduction

In this big project, two options are given to you: you may work on a topic of your own choice (fulfilling a minimal list of requirements), or you may work on the topic we have prepared for you.

In this project, you will have to design your custom model, and possibly create your own mock dataset or look for one online. We advise to have a look at the typical dataset repository websites such as Kaggle et Google DataSearch.

Please assume that this project is to be submitted to a serious client (played by us professors) and take this project as an opportunity to work on your project delivery skills. More specifically, pay attention to the way you code and document your code.

Finally, your submission will have to be uploaded for grading. Please only submit your code and notebooks on eDimension, not the dataset! (Otherwise, the submission will be too large). We also advise to upload your final project on Github, to enrich your personal portfolio of projects. Shall the dataset be required to run your project, let us know where to download it (Kaggle or even better, upload it to our own Google Drive/Dropbox).

### 2. TLDR

Groups of 3-5 people, will not allow people to work alone.

Mixing people from both classes is allowed.

Choice between a custom project and a plan B project if uninspired.

**Custom project proposal deadline: April 7<sup>th</sup>, 11.59pm.**

**Let us know about your groups before April 7<sup>th</sup>, 11.59pm at**

**<https://docs.google.com/spreadsheets/d/1EQ-85BTAK0Tn8yxeVHNp4dFdGLYBvauh7ku3x3VWlZA/edit?usp=sharing>**

**Submission deadline: 2<sup>nd</sup> May 2021, 11.59pm.**

### 3. Choice 1: custom proposal for a big project

#### Objectives and guidelines

**If you plan to go for the custom project, you will need to submit the project proposal via email to me by the 7th of April, 11.59pm for approval.** You may submit before this deadline, so you can get to start early.

**In addition, let us know about your group on the following link, before April 7<sup>th</sup> as well.**

<https://docs.google.com/spreadsheets/d/1EQ-85BTAK0Tn8yxeVHNp4dFdGLYBvauh7ku3x3VWlZA/edit?usp=sharing>

The proposal should be a PDF, containing a brief description of the following elements

- Topic, problem to be investigated,
- Expected inputs and outputs,
- What dataset you are going to use,
- Architecture draft (briefly hinting at the type of architecture you plan to use, does not have to be extremely descriptive),
- Team members,
- What you are going to deliver,
- Deliverable must be at least code for training, code for deployment, a small GUI runnable on user-selected data (at least a few examples), and a report.

**Important note:** no proposals using MNIST datasets or its derivatives, or Cats vs. Dogs datasets will be approved. No trivial binary classification problems in vision, either. It should at least include some aspects of CNNs/RNNs (image captioning, etc.) or some graph dataset concepts. Apart from these constraints, we are rather open to crazy topics!

### 4. Choice 2: Implement an image captioning

#### Objectives and guidelines

The objective in this plan B project is to implement an image captioning engine, that is, a model which could receive images and give them a caption in the form of a sentence briefly describing the content of said image (see [https://github.com/yunjeypytorch-tutorial/tree/master/tutorials/03-advanced/image\\_captioning](https://github.com/yunjeypytorch-tutorial/tree/master/tutorials/03-advanced/image_captioning) for some examples)

For inspiration on possible datasets to use and dataloaders, you may refer to [https://github.com/yunjeypytorch-tutorial/blob/master/tutorials/03-advanced/image\\_captioning/data\\_loader.py](https://github.com/yunjeypytorch-tutorial/blob/master/tutorials/03-advanced/image_captioning/data_loader.py).

For obvious reasons, transfer learning will only be allowed, to some limited extent. If you only use a pretrained model and do not iterate on it, or use one that is too close in terms of norms of weights, then you will get strongly penalized... You may for instance reuse a pre-trained ResNet architecture as a starting point, but it will have to be retrained to match your chosen dataset and problem.

You may however use a word embedding from the internet, without having to train your own one on your own vocabulary (we strongly suggest to use FastText!). Training your custom word embedding is not exactly interesting, therefore using a pre-trained one will not be penalized.

Regarding the dataset, we strongly advise to use the MSCOCO 2014/2015 captioning dataset (<https://cocodataset.org/#home>).

You are not required to use attention models. During the demo, your system should be able to run with custom images, not in the train/val/test set.

**If you plan to go for the plan B project, you will not need to submit the project proposal via email, but should simply let us know for your decision by the 7th of April, 11.59pm.**

**In addition, let us know about your group on the following link, before April 7<sup>th</sup> as well.**

<https://docs.google.com/spreadsheets/d/1EQ-85BTAK0Tn8yxeVHNp4dFdGLYBvauh7ku3x3VWlZA/edit?usp=sharing>

## 5. List of expectations for both projects

Report your performance on the validation set. Give a quick comparison of your model performance against some state of the art ones (you may for instance use the BLEU score for models comparison). You are not expected to beat the state-of-the-art models!

As with the small project, you are expected to describe the different steps you have taken to create your architecture from scratch/pre-trained models. A brief explanation should be given on how to recreate the model from scratch.

We also expect you to save your model weight for reproducibility. Visualization of your model performance (e.g. accuracy and loss curves, performance of your system on some validation set images, etc.) are also expected. For each figure used in your report, there should be a clear description on how to recreate said figure. Please show examples of your model malfunctioning if any (in the case of the plan B project, you may report the two weirdest/funniest wrong predictions (screenshot with example image)).

Your report should contain everything we need to know to run your code (including the package dependencies).

Put in your submission the group members and their contribution to the project.

## 6. Project Delivery

### Delivery details

Groups of 3-5 people, will not allow people to work alone.

Mixing people from both classes is allowed.

**Submission deadline: 2<sup>nd</sup> May 2021, 11.59pm.**

### Recapitulative report

Your recapitulative report shall be submitted in a PDF format, along with your code.

Your code may consist of .py files, ipynb jupyter notebooks, or Google Colab notebooks. Your PDF report should explicitly mention what needs to be done to run your code.

Properly documenting your notebooks/code files would be much appreciated, and you should practice it anyway (they are good practice!). Personally (Matt), I believe the most presentable format is a Jupyter Notebook, combining Markdown cells and code cells, along with .py files containing the largest parts of your code: you just have to import them later in your Jupyter Notebook, to minimize the amount of code in your Notebook!

### Project delivery

We strongly advise to upload your submission (code/notebooks + PDFs, but no dataset due to space restrictions) on a Github repository. You can then submit the link to your PUBLIC Github repository, during your submission on edimension.

Your Github repository for this project should contain your PDF report, your DOCUMENTED code/notebook files. It should also contain directions showing the required libraries and steps needed to re-train the model from scratch. And more importantly, it should also contain clear directions on how to recreate the exact trained model and its performance results you are presenting in the PDF, by loading some saved PyTorch weights for your model. This is essential, for reproducibility reasons and something that we, researchers, must do constantly for our papers. If the weights files are too heavy to be submitted on edimension or Github, you may upload them to a Google Drive/Dropbox instead.

**Important note:** you are nearing the end of your curriculum at SUTD, gathering your projects and uploading them on your personal Github is much appreciated by recruiters as it allows them to immediately identify what your coding capabilities are. If you have not done so yet, please consider starting your project portfolio on Github!

## Grading rubric

Project originality and project proposal: 10%

Report (organization, readability and content): 40%

Code (implementation, documentation, good/bad practices): 50%

**Important note: if you choose the plan B project, you will forfeit the points given to originality and project proposal (as we are giving you the idea of this topic and the dataset to use). This means, the maximal score you will be able to get for this project will be 90%.**