Groups + Project Proposal

Due Mar 17 by 8am **Points** 2 **Available** Feb 2 at 12am - Mar 19 at 8am

Topics: You can choose any topic you'd like for the project, but here are some suggestions

- **New Facebook Projects (https://gatech.instructure.com/courses/294044/files/36258463?wrap=1)**
  - **(https://gatech.instructure.com/courses/294044/files/36258463?wrap=1) (Note you can still take a look at the original Facebook Projects (https://gatech.instructure.com/courses/294044/pages/project-ideas?module_item_id=2825318) but they may be outdated)**
- **Additional ideas (https://www.cc.gatech.edu/classes/AY2020/cs7643_spring/other_projects.html)**

Project description and rubric: Group_Project_Description.pdf (https://gatech.instructure.com/courses/294044/files/38921405/download)

Instructions for Project Proposal:

The final project provides an opportunity for you to (1) gain experience implementing deep NN models and (2) try non-trivial Deep Learning on problems that interest your team.

The project consists of two deliverables

- Project Proposal
- Project Report (**report template (https://www.overleaf.com/read/fdjpfsdhztfp)**)

Group size:

Recommended group size is 2-4. Typically smaller or larger groups than this are not allowed and will be severely penalized.

Project Ideas

You can come up with any topic you'd like! Below are some suggestions if you are having trouble coming up with one (again you are not limited to them). We highly recommend watching interesting talks, reading papers (feel free to ask during office hours!), etc. to get ideas.

- Facebook project ideas are posted in canvas.
- FB is working on updating these.

- Other topic ideas can be found **here** (https://www.cc.gatech.edu/classes/AY2020/cs7643_spring/other_projects.html)
- Ethical Implications: Are there any negative societal or ethical implications of your project topic (as an example, can be directly used for harassment)? If so, please list them and consider changing the topic to reduce this risk.

Rubrics

The final report rubrics will be on **Canvas** (https://gatech.instructure.com/courses/294044/assignments/1225504), along with the **LaTex template** (https://www.overleaf.com/read/fdjpfsdhztfp) for the final report (you can use Word or something else, but please follow the format and structure and font/margins to make sure you don't miss something).

The proposal is worth 2 points and used as a sanity check to assess the risk of the project.

Due Dates

Please refer to the **course syllabus** (https://gatech.instructure.com/courses/294044/assignments/syllabus) for due dates.

Instructions for Project Proposal

What goes in a project proposal?

- Team Name
- Project Title
- Project summary (4-5+ sentences). Fill in your problem and background/motivation (why do you want to solve it? Why is it interesting?). This should provide some detail (don't just say "I'll be working on object detection")
- What you will do (Approach, 4-5+ sentences) - Be specific about what you will implement and what existing code you will use. Describe what you actually plan to implement or the experiments you might try, etc. Again, provide sufficient information describing exactly what you'll do. One of the key things to note is that just downloading code and running it on a dataset is not sufficient for a description or a project! Some thorough implementation, analysis, theory, etc. have to be done for the project.
- Resources / Related Work & Papers (4-5+ sentences). What is the state of art for this problem? Note that it is perfectly fine for this project to implement approaches that already exist. This part should show you've done some research about what approaches exist.
- Datasets (Provide a link to the dataset). This is crucial! Deep learning is data-driven, so what datasets you use is crucial. One of the key things is to make sure you don't try to create and especially annotate your own data! Otherwise, the project will be taken over by this.
- List your Group members.

Where are project proposals submitted/graded?

- Groups will submit the project proposal on Gradescope. We also encourage groups that are looking for more members to post their proposals on Ed (some have been doing this already)
- project proposals are not a “one-shot” submission, you are free to refine it until the deadline and resubmit.

What you need to do for the proposal:

In summary:

- Create groups
- draft a project proposal with group-mates
- post a draft of the proposal on Ed if you are looking for teammates (do this sooner rather than later)
- Finalize proposal and submit by the due date.
- IMPORTANT: Include everyone in your group during the Gradescope submission (resubmit if new team members join), if someone is missing they will not get a grade. Don’t be responsible for your teammate getting a 0.

FAQ

- I don’t have a group, can I do my project Solo?
  - No, you will be penalized for not gaining approval for a solo group.
- It’s the week/day/hour before the proposal deadline and I couldn’t find a group. Can I do my project solo?
  - This is insufficient justification. You will be penalized for not gaining approval for a solo group.
- I don’t have a group, when should I start panicking?
  - We are releasing this info now but note the proposal is not due until basically the end of Oct. Expect churn in groups until early Oct. (students will drop, find other topics of interest, etc.). Ideally, you want to start locking down groups after assignment 2 is due and start experimenting with ideas before locking down a project topic.
- Can we use existing implementations?
  - Yes, you can use existing implementations (do not forget to cite). But again, remember that your project has to be comprehensive per the project guidelines; don’t provide just the results and shallow analysis. If you used existing models/code this is even more important as we can’t judge how much you have learned from that. The key is to have iteration between running, making hypotheses/claims about what’s happening (with evidence, which you should show), and documenting what decisions you made in response to improving things. Also, note that you have to be careful about unofficial implementations. Often they have bugs; typically we try to avoid using implementations that don’t reproduce the results from the paper. You can often tell as the good ones have tables comparing their performance to the paper in the README.
- Changing topic after the proposal deadline?
  - It depends, it is a higher risk further in the term. Feel free to attend office hours to ask or make a private post if you encounter this. Ultimately it's up to you and your group, the proposal doesn’t

lock you into anything however the expectations for the final report do not change even if you changed topics the week before.

- Can we use pre-trained models?
  - You are allowed to use pre-trained models however just running open-source code is not sufficient. What we're looking for in the project is the ability to perform and analyze non-trivial DL. In that vein, using a pre-trained model and fine-tuning on a dataset is insufficient by itself, however, using a pre-trained model and doing some non-trivial DL as part of the pipeline is acceptable. Do not forget to Cite.
- Can we use other libraries other than Pytorch?
  - Yes.
- What makes a good project?
  - Usually projects where the dataset is readily available and existing benchmarks are solid options however you are free to experiment.
  - Projects, where you collect your own data, take longer, require more infrastructure, and generally are not good candidates. We don't discourage passion projects, so if you had something in mind go for it! Just be forewarned that you should likely start looking into the infrastructure now to make sure you have both the right data and enough data amenable to DL. An additional challenge here is that you will have to create baseline models (naive or otherwise) to show DL shows improvement over something.
  - Deep RL projects are tough for the same reason, in DRL you'd be using a simulator that generates the data you train on. Oftentimes these simulators require significant debugging to get working so be forewarned. Again if you're passionate don't let this stop you from pursuing projects in this space.
- What type of resources are required to train DNN's for the project?
  - Depending on the dataset it can take hours/days for models to train on non-trivial data sets. If you haven't used GPUs or cloud computing before and plan to use it (often necessary for the FB projects or non-trivial data sets), plan on accounting for time to learn how to do this.

An Example Project Proposal:

- A sample will be posted on ed
- We will make a post with a sample final project report after the proposal deadline.