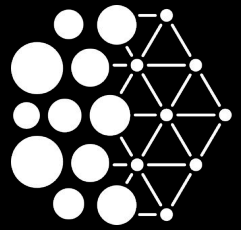
Institut
québécois
d'intelligence
artificielle

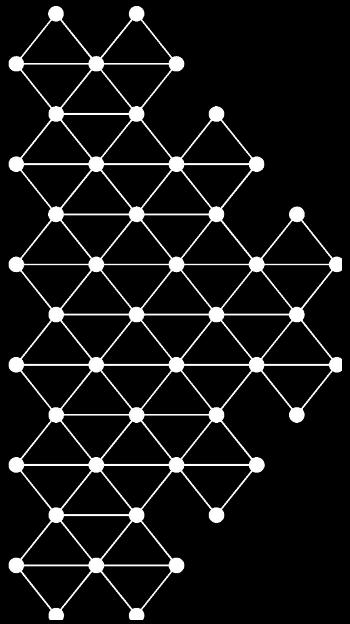


Mila

IFT6759
Advanced Projects In Machine Learning
Winter 2020

Institut
québécois d'intelligence
artificielle

Mila



Syllabus

Objectives

- Apply ML theory to complex projects
- This involves:
  - Understanding a problem along with the available data
  - Defining and implementing appropriate machine learning solutions
  - Writing high-quality code and readable documentation in Python
  - Using collaborative tools and machine learning libraries
  - Preparing reports and oral presentations

Course Structure

- Two (2) projects have been selected this year
  - Each project will have a total duration of 6 weeks
  - Teams of 4 students will be formed to work on each project

Class Schedule

- Start Date: January 8th, 2020
- End Date: April 17th, 2020
- Class Hours:
  - Wednesday: 15:30-17:30 (Mila Agora)
  - Friday: 9:30-11:30 (Mila Auditorium #2)
- Office Hours:
  - Ankit Vani: Tuesday 14:00-15:00, room J09
  - Breandan Considine: Friday 11:30-12:30, Cube 08
  - Krishna Murthy: Friday 11:30-12:30, room D07
- Contact email: ift6759@mila.quebec.

Instruction Style

There will mainly be two types of class sessions:

- Tutorial sessions (first two weeks of the semester)
  - Presentation of material useful for tackling the assigned projects (Unix, Helios cluster, Git / Github, TF, …)
- Working sessions
  - Students work on their assigned projects; TAs are available for support
  - Teams meet with the TAs to brainstorm and provide status updates

Note: students are expected to **continue working** on the projects outside of class hours.

Computing Resources

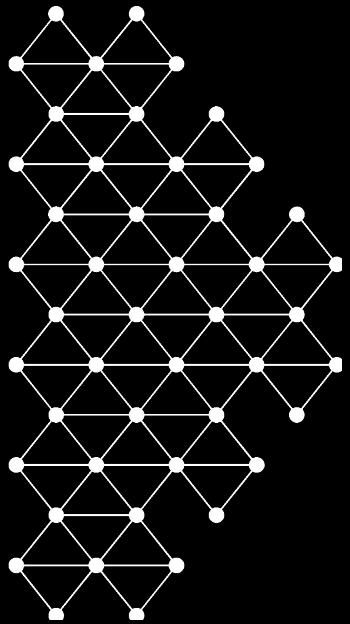
- Access to an external GPU cluster is provided to work on the projects
  - Tutorials on how to access the cluster will be provided this Friday
- The cluster is shared with other users outside of the course.
- Reserved time slots with high priority to students in the class:
  - Wednesdays from 11:30 to 19:30
  - Fridays from 8:30 to 16:30
- Students can still use the GPU cluster outside these time slots
  - Accessibility on a **"first-come, first-served"** basis

Policy on Contribution

- It is expected that each team member **contributes** to the team work
- If you suspect someone in your team is not contributing:
  - Have a team meeting to discuss with the person on how she/he can contribute:
    - Use a positive and respectful attitude
    - Try to help the person find a way to contribute
- If the situation persists:
  - Send an email to the TAs:
    - A meeting with all team members will be scheduled
    - Everyone should be able to justify their position with some relevant material
    - TAs and the professor will take actions to resolve the problem

Institut
québécois d'intelligence
artificielle

Mila



Evaluation

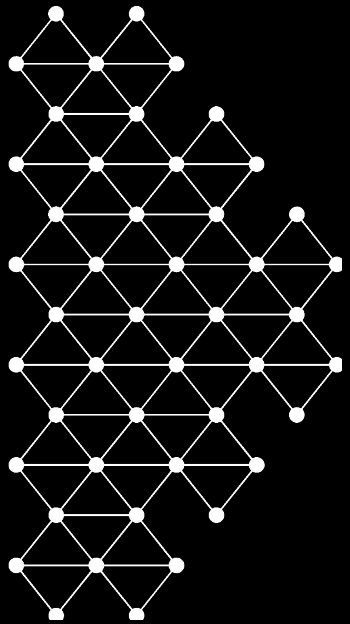
Project Deliverables

For each project, each team will need to hand in the following deliverables:

- A **report** summarizing the experiments and the results of the project
- The **code** written as part of the project notably to:
  - Generate, analyze, visualize and pre-process the data
  - Train and finetune models
  - Run a model on a data set to generate predictions
  - Analyse the results
- The **best model** obtained (used by the TAs for blind test evaluation)
- A **presentation** in front of their peers

Institut
québécois d'intelligence
artificielle

Mila



Project Evaluation Grid

Evaluation Grid (50% Per Project)

- Code review: 10%
- Report Evaluation: 20%
- Model Performance on blind test set: 10%
- Project Presentation (oral): 10%

Code Review

Code quality	/8
Code is well organized : modular code design (different files for data processing, model definition, model training, model inference, …), no/minimal code duplication Evaluation : -0.1 per error	/1
Code respects the PEP8 standard Evaluation : -0.2 per error	/1
Code is easy to read : complex code blocks are accompanied by comments, comments are relevant (see article), functions and variables have meaningful names that reflect their function. Evaluation : -0.1 per error	/1
Proper management of input arguments in the training script (see argparse, python fire, configparser) Evaluation : 1/1 for overall adequate input management, 0/1 otherwise	/1
Proper utilization of GitHub (e.g. branching, relevant commits and messages, usage of pull request) Evaluation : 1/1 if git branches and Github pull request are used for most contributions to the repository. 0/1 otherwise.	/1
Executable scripts have a “main” function (see article) Evaluation : the points given will be equal to the ratio of executable scripts with main function over the number of executable scripts in the project.	/1
Reproducible experiments (e.g. seed) Evaluation : -1 per error	/2

Report Evaluation

Introduction (~1 page)	**/1**
Purpose : introduce the project (present and motivate the problem, summarize your contributions, outline the following sections of the report) Evaluation criteria : • Task is introduced and described • Model evaluation metric is explained	 /0.5 /0.5
Data analysis (~1 page)	**/3**
Purpose : in-depth analysis of the provided data using statistics, visualizations, etc. Evaluation criterion : • Relevant data attributes are analysed : number of samples, available features, missing/corrupted values, underlying structure, and any other data property relevant to the project.	 /3
Literature review (~1.5 page)	**/3**
Purpose : identify and summarize any prior work (research articles, datasets, models, …) that are relevant to the project Evaluation criteria : • Literature review is relevant to the project and demonstrates a reasonable level of effort • Links are drawn between articles when relevant	 /2 /1
Methodology (~2 pages)	**/4**
Purpose : describe your experimental pipeline Evaluation criteria : • Experimental pipeline is described in sufficient amount of detail to allow other people to run equivalent experiments : description of models, external data sources (if needed), data splits, data processing, loss function, learning procedure, hyperparameter tuning, … • The selected methodology makes sense (coherent with the results of the data analysis and the literature review).	 /2 /2

Report Evaluation (2)

Results and discussion (~2 pages)	/6
Purpose : present your experimental results and analyse them. What can be learned from your results? What works and what doesn't work? How could the results be improved? Evaluation criteria : • Results are presented clearly using the appropriate media (tables, figures, etc.). • Analysis demonstrates a good understanding of the results. The presented conclusions are supported by the results.	 /1 /5
Conclusion (~0.5 page)	/2
Purpose : summary of project state Evaluation criteria : • This section summarizes the work that was done and its limitations • Recommendations are made regarding future work on the project (what would be the next step, if more time was available?)	 /0.5 /1.5
Overall quality of the report	/1
Report format (7-9 pages including references, font size 11, title with team member names, clear sections, flow between sections, figures and tables titled, axes titled, etc.)	/1

Model performance on blind test set

The best model provided by a team will be compared to a baseline model built by the TAs:

- If the best model provided by a team crashes or provides results that are statistically worse than those of the baseline model, the team gets 0%.
- Otherwise, if the best model provided by a team is statistically equivalent to the baseline model, the team gets 6%.
- Otherwise, if the best model provided by a team is statistically better than the baseline model:
  - The team gets 10% if the model is the best performing one or is statistically equivalent to the best performing model provided by another team.
  - Otherwise, the team gets 8%.

Model performance on blind test set

Do not attempt to obtain the blind test set in any way (abusing a security loophole, attempting to recover the original data from the internet, etc.). We reserve the right to reject a model if we have valid reasons to believe that any part of the test set was used to build/refine the model.

Project Presentation

Content of the presentation	/4
Description of the project	/1
Description of the solutions adopted	/1
Presentation of the achievements	/1
Synthesis of findings, obstacles, and recommendations	/1
Format of the presentation	/4
The presentation is clear and structured	/1
The presentation is engaging and speakers have good delivery (looking at the audience, adequate voice level and speaking rate, speaker enthusiasm, ...)	/2
Respect of time	/1
Questions period	/2
The answers to the questions are precise and clear	/2

Deadlines

- Deliverable (report + code + model) deadlines:
  - **11:59 PM on the day before** the project presentation
  - For **every late day**, the project will see its **grade reduced by 10%** (out of 100)
  - **After 4 late days**, the project will automatically get **0%**.

- Peer evaluation deadlines:
  - “Informal” peer evaluations: **Friday 11:59 PM** on the **third week** of the project
  - “Official” peer evaluations: **Friday 11:59 PM** on the **sixth week** of the project

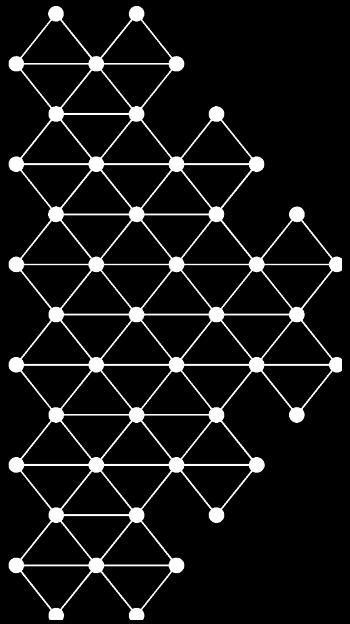
Deadlines - Example

Example for a project with the presentations on a Friday:

- Project handed in at the latest on Thursday 11:59 PM : No penalty
- Project handed in before Friday 11:59 PM : **10%** penalty
- Project handed in before Saturday 11:59 PM : **20%** penalty
- Project handed in before Sunday 11:59 PM : **30%** penalty
- Project handed in before Monday 11:59 PM : **40%** penalty
- Project handed in after Monday 11:59 PM : **100%** penalty

Institut
québécois d'intelligence
artificielle

Mila



Grading Policy

Grading Policy

- Each project will contribute 50% of the final grade
- At the end of a project handled by a team:
  - The **project's group grade** is provided based on the evaluation of their deliverables
  - Each student's individual grade will be based on **the group grade with a small (factor) adjustment based on peer evaluations**.

Grading Policy

- Peer evaluations are provided by team members at two moments:
  - Halfway during the project: **“informal”** peer evaluations (**no effect** on final grades)
    - Students will be provided with the average peer evaluation grade
    - An opportunity to adjust one’s contribution during the remainder of the project
  - At the end of the project: **“official”** peer evaluations
    - They will have **an effect** on the students’ final project grades.

Individual Grades

- Obtained by multiplying the overall grade for the group project **(the same for every student in a team)** with an **individual adjustment factor** which will reflect each student's personal contribution to the project as a whole

- As part of the peer evaluation, each team member will have 100 "contribution" points to split between him/herself and the others. Following is an example assuming a team of 4 students:

Eval Of		Eval by		
	Student A	Student B	Student C	Student D
Student A	25	27	30	25
Student B	23	25	30	25
Student C	47	46	36	25
Student D	5	2	4	25
SUM	**100**	**100**	**100**	**100**

Individual Grades

- Every student's contribution score is obtained by taking **the average of the evaluations given to him/her by the other students** (a student's own evaluation doesn't count).

- These average contribution scores are then rescaled such that they sum to 100 for the team

Eval Of		Eval by				
	Student A	Student B	Student C	Student D	Average	Rescaled
Student A	25	27	30	25	27,33	28,37
Student B	23	25	30	25	26,00	26,99
Student C	47	46	36	25	39,33	40,83
Student D	5	2	4	25	3,67	3,81
SUM	**100**	**100**	**100**	**100**	**96,33**	**100,00**

Individual Grades

- Transformation of these scaled average scores into **bounded adjustment factors**
  - The bounds for the adjustment factors will be decided later (here, we use 85% - 115%)
  - The first value is the adjustment factor corresponding to average scores $\leqslant 12.5$ (half of the expected contribution score if everyone contributed equally in a team of 4)
  - The second value is the adjustment factor corresponding to average scores $\geqslant 50$ (twice the expected contribution score if everyone contributed equally in a team of 4)
  - To convert average scores to adjustment factors, these scores are constrained to lie in the interval [12.5, 50.0] and a quadratic polynomial function is fitted on these bounded scores to transform them into an adjustment factor.
    - $12.5 \rightarrow 85\%$
    - $25 \rightarrow 100\%$
    - $50 \rightarrow 115\%$

Individual Grades

Eval Of		Eval by							
	Student A	Student B	Student C	Student D	Average	Rescaled	Floor	Ceiling	Adjustment factor
Student A	25	27	30	25	27,33	28,37	28,37	28,37	103,19%
Student B	23	25	30	25	26,00	26,99	26,99	26,99	101,93%
Student C	47	46	36	25	39,33	40,83	40,83	40,83	111,82%
Student D	5	2	4	25	3,67	3,81	12,50	12,50	85,00%
SUM	**100**	**100**	**100**	**100**	**96,33**	**100,00**			

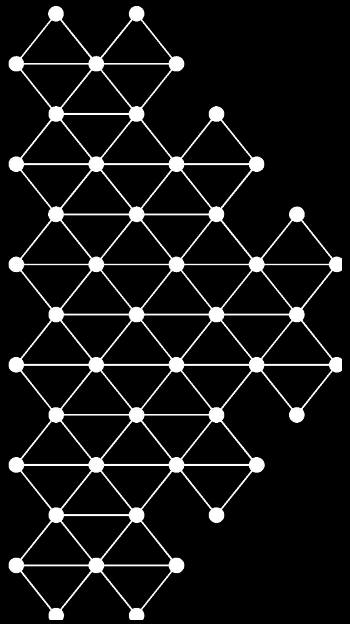
Notes:

- You are asked to base your peer evaluations not only on **their contribution of results**, but also on their **contribution of time** and **contribution of effort**.

- The contribution scores you assign to your teammates should reflect your honest opinion. We reserve the right, if we suspect foul play within a team, to step in and adjust the scores. If this happens, the team in question will be notified.

Institut
québécois d'intelligence
artificielle

Mila

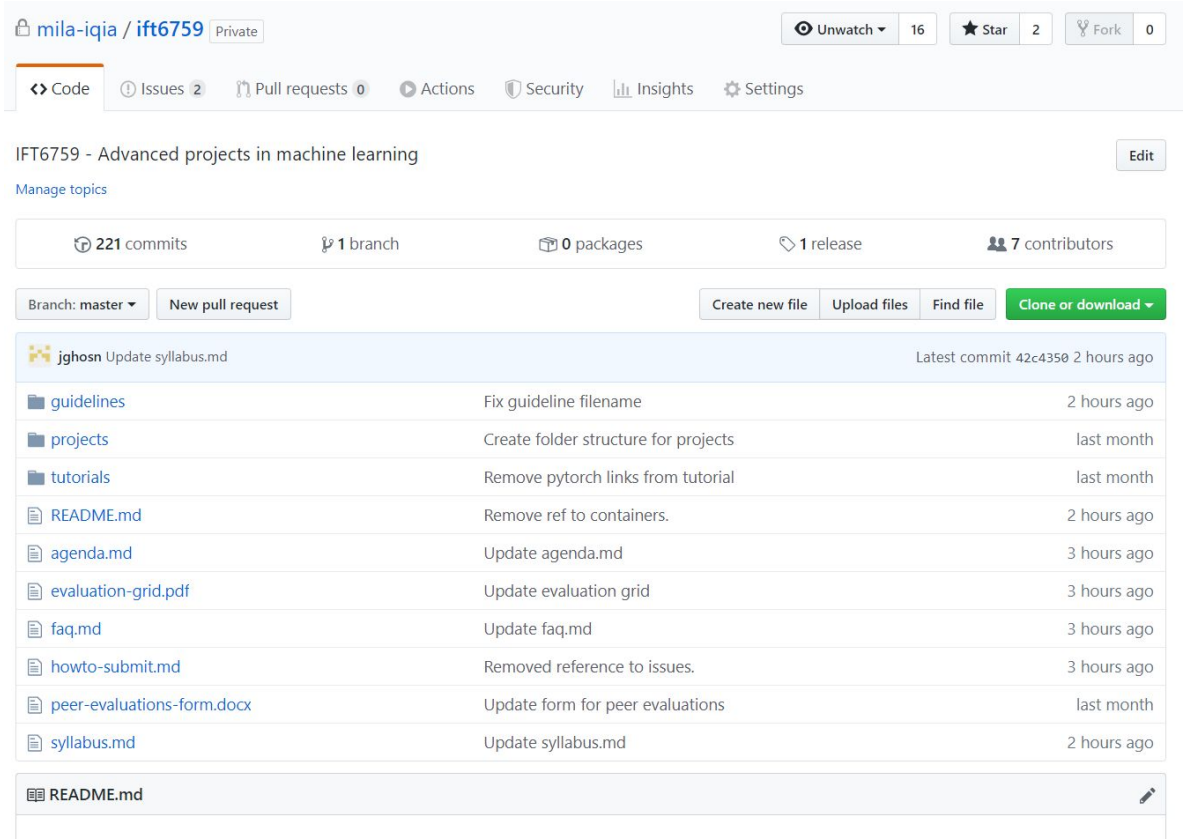


Resources for the course

Course Github repository

The github repository contains links to all the necessary course material including the syllabus, evaluation grid, weekly agenda and more.

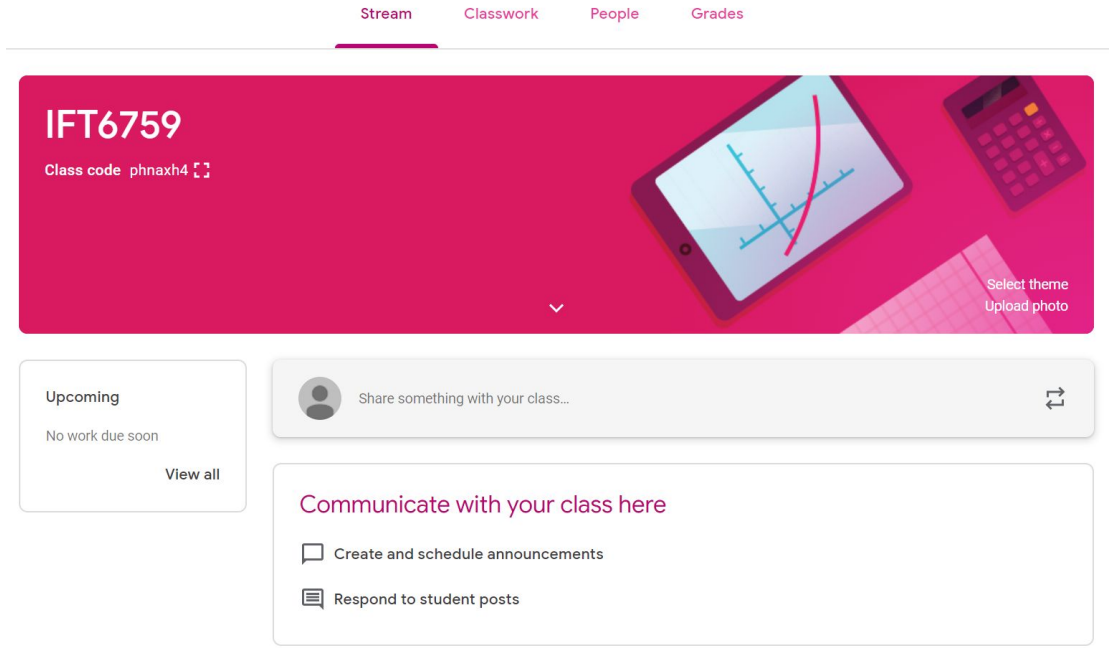
URL : https://github.com/mila-iqia/ift6759



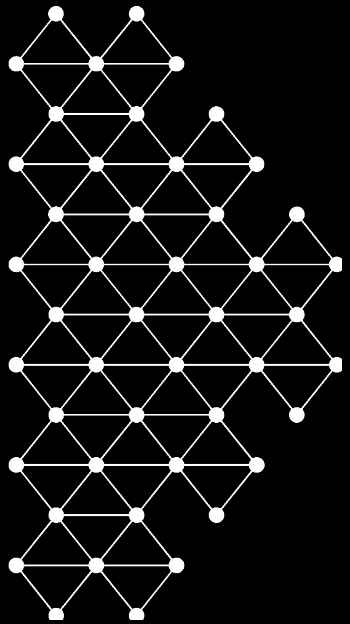
Course message board

The message board can be used to communicate with other students or to ask public questions to the TAs.

The link to the message board is available from the github repository.



Institut
québécois d'intelligence
artificielle



Thank you