



Fundamentals of Machine Learning

Fondements de l'Apprentissage Machine

Lab 0 - Basics to know BEFORE you start

Bases à connaître AVANT de commencer

Ioannis Mitliagkas, Guillaume Rabusseau

Teaching Assistants:

Reyhane Askari, Thomas George, Meraj Hashemizadeh, Philippe Laferrière, Viet Nguyen, Michael Noukhovitch, Mandana Samiei, Alexis Tremblay, Vikram Voleti

Slides by Vikram Voleti, Michael Noukhovitch, Rémi Le Priol and Salem Lahlou, 2019

Contents

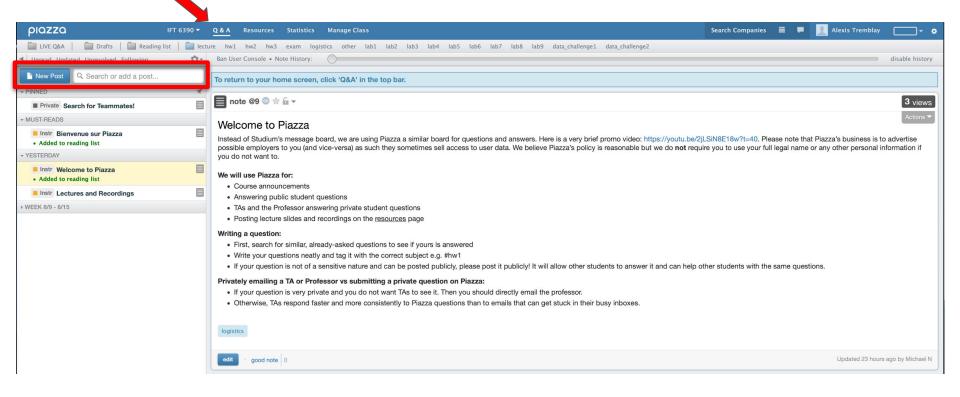


- 1. Piazza
- 2. Gradescope
- 3. Overleaf (Latex)
- 4. Miniconda (Package manager)
- 5. python hello_world.py
- 6. Jupyter
- 7. Google Colab

1. Piazza



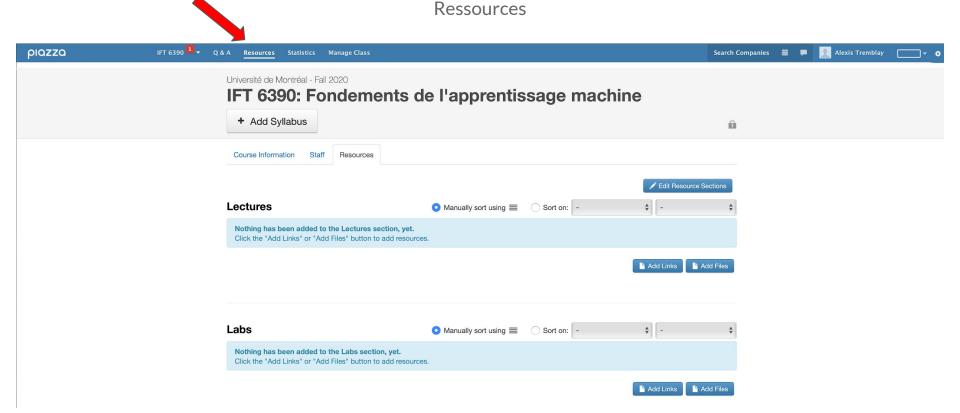
Announcements - Forum - Annonces



1. Piazza



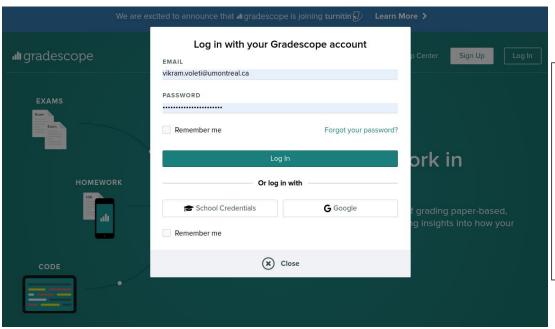
Resources





All assignment submissions will be on gradescope.com

Toutes les soumissions seront sur gradescope.com

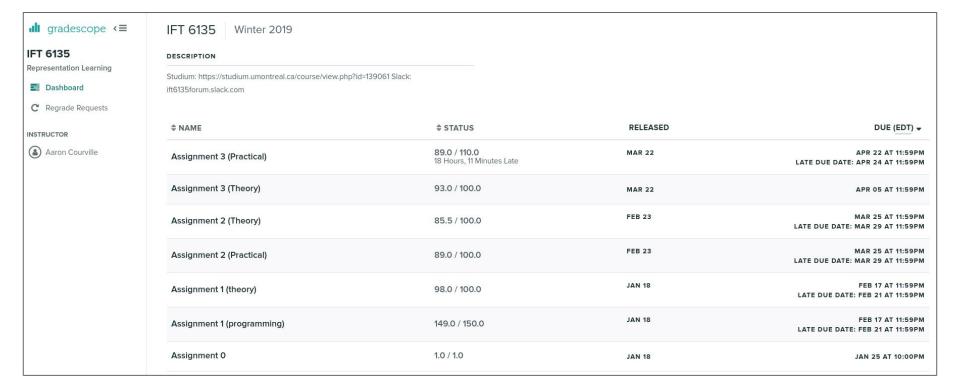


ıllı gradescope <≡	Your Courses	
Your Courses Welcome to Gradescope! Click on one of your courses to the right, or on the Account menu below.	Winter 2019	
	IFT 6135 Representation Learning	
	7 assignments	



Find all your assignments here! (with their due dates, status, etc.)

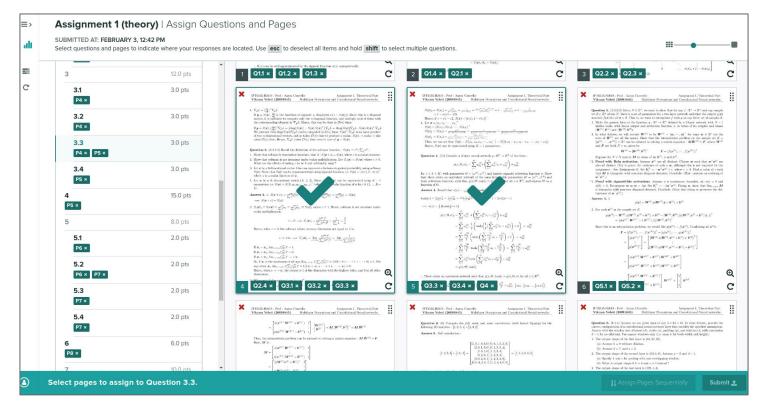
Trouvez tous vos devoirs ici! (avec leurs dates d'échéance, statut, etc...)





<u>Submission</u>: for <u>each</u> question, select <u>all pages</u> with the answer (hold shift)

Pour chaque question, sélectionnez toutes les pages correspondant à votre réponse (en appuyant sur shift)



- TAs can point out errors or comments Les TAs peuvent spécifier vos erreurs et vous faire des
- "Regrade requests" can be submitted within a 2 week deadline

retours dessus.

Sur Gradescope, vous pouvez demander qu'on recorrige certaines de vos réponses (justifier pourquoi) jusqu'à 2 semaines après publication des notes.

- 2. Let Γ be a diagonal matrix with $\Gamma_{ii} = (X^{\top}X)_{ii}^{1/2}$. Show that the expectation (over R) of the loss function can be rewritten as $L(\mathbf{w}) = ||\mathbf{v} - p\mathbf{X}\mathbf{w}||^2 + p(1-p)||\Gamma \mathbf{w}||^2$.
- Show that the solution w^{dropout} that minimizes the expected loss from question 2.2 satisfies

$$pw^{\text{dropout}} = (X^{T}X + \lambda^{\text{dropout}}\Gamma^{2})^{-1}X^{T}y$$

where λ^{dropout} is a regularization coefficient depending on p. How does the value of p affect the regularization coefficient, λ^{dropout} ?

- Express the solution w^{L2} for a linear regression problem without dropout and with L² regularization, with regularization coefficient λ^{L_2} in closed form.
- 5. Compare the results of 2.3 and 2.4: identify specific differences in the equations you arrive at, and discuss qualitatively what the equations tell you about the similarities and differences in the effects of weight decay and dropout (1-3 sentences).

Answer 2. 1. $L(w) = ||y - (R * X).w||^2$, where * is element-wise multiplication.

2. $\mathbb{E}_{R}[L(w)] = \mathbb{E}_{R}[||y - (R * X).w||^{2}] = \sum_{i=1}^{n} \mathbb{E}_{R}[(y_{i} - (R_{i} * X_{i}).w)^{2}]$ We know (from Assignment 0) that:

$$\operatorname{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \implies E[X^2] = E[X]^2 + \operatorname{Var}(X)$$

$$\therefore \mathbb{E}_R[L(w)] = \sum_{i=1}^n (\mathbb{E}_R[y_i - (R_i * X_i).w])^2 + \operatorname{Var}_R(y_i - (R_i * X_i).w)$$

$$= \sum_{i=1}^n (y_i - (\mathbb{E}_R[R_i] * X_i).w)^2 + \operatorname{Var}_R(y_i - (R_i * X_i).w)$$

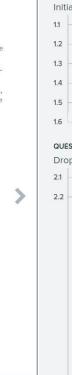
$$\sum_{i=1}^n (y_i - (\mathbb{E}_R(R_i) * X_i).w)^2 = \sum_{i=1}^n (y_i - pX_i.w)^2 = ||y - pXw||^2$$

$$\begin{array}{l} \sum_{i=1}^{n} \operatorname{Var}_{R} \left(\mathbf{y}_{i} - (R_{i} * \mathbf{X}_{i}).w \right) = \sum_{i=1}^{n} \mathbb{E}_{R} \left[\left(\mathbf{y}_{i} - (R_{i} * \mathbf{X}_{i}).w - \mathbb{E}_{R} | \mathbf{y}_{i} - (R_{i} * \mathbf{X}_{i}).w | \right)^{2} \right] \\ = \sum_{i=1}^{n} \mathbb{E}_{R} \left[\left(\mathbf{y}_{i} - (R_{i} * \mathbf{X}_{i}).w - \mathbf{y}_{i} + p\mathbf{X}_{i}.w | \right)^{2} \right] = \sum_{i=1}^{n} \mathbb{E}_{R} \left[\left((p\mathbf{X}_{i} - R_{i} * \mathbf{X}_{i}).w \right)^{2} \right] \\ = \sum_{i=1}^{n} \mathbb{E}_{R} \left[\mathbf{w}^{\top} . (p\mathbf{X}_{i}^{\top} - R_{i}^{\top} * \mathbf{X}_{i}^{\top}).(p\mathbf{X}_{i} - R_{i} * \mathbf{X}_{i}).w \right] \\ = \sum_{i=1}^{n} \left(\mathbf{w}^{\top} . (p^{2}\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \right) \\ = \sum_{i=1}^{n} \left(\mathbf{w}^{\top} . (p^{2}\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p^{2}\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \right) \\ = \sum_{i=1}^{n} \left(\mathbf{w}^{\top} . (p^{2}\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p^{2}\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \right) \\ = \sum_{i=1}^{n} \left(\mathbf{w}^{\top} . (p^{2}\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p^{2}\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \right) \\ = \sum_{i=1}^{n} \left(\mathbf{w}^{\top} . (p\mathbf{X}_{i}^{\top} \mathbf{X}_{i}) - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \right) \\ = \sum_{i=1}^{n} \left(\mathbf{w}^{\top} . (p\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p^{2}\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \right) \\ = \sum_{i=1}^{n} \left(\mathbf{w}^{\top} . (p\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \right) \\ = \sum_{i=1}^{n} \left(\mathbf{w}^{\top} . (p\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \right) \\ = \sum_{i=1}^{n} \left(\mathbf{w}^{\top} . (p\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \mathbf{X}_{i}^{\top} \right) \\ = \sum_{i=1}^{n} \left(\mathbf{w}^{\top} . (p\mathbf{X}_{i}^{\top} \mathbf{X}_{i} - p\mathbb{E}_{R} | \mathbf{X}_{i}^{\top} \mathbf{X}_{i}$$

 $=\sum_{i=1}^{n} (w^{T}.p(1-p)(X_{i}^{T}X_{i}).w)$ $= \overline{p(1-p)}(\boldsymbol{w}^{\top}.\operatorname{diag}(\boldsymbol{X}^{\top}\boldsymbol{X}).\boldsymbol{w})$

 $= p(1-p)||\Gamma w||^2$

$$\mathbb{E}_{R}[L(w)] = ||y - pXw||^{2} + p(1-p)||\Gamma w||^{2}$$





QUESTION 1

Initialization 12 / 16 pts 11a Gaussian w/activations R 2/3 pts R 0/3 pts 1.1b Uniform w/activations 12a Gaussian w/Rel U 3/3 pts 1.2b Uniform w/ReLU 3 / 3 pts 12c name 2 / 2 pts 1.2d why does it work? 2/2 pts QUESTION 2 Dropout 19 / 21 pts Write L(w) 4 / 4 pts Derive E[L(w)] R 4/6 pts √ - 2 pts incorrect or unjustified derivation of variance term, e.g. - replacing X^T X with Gamma^2 handling expectation of (RXw)^T (RXw) incorrectly C Regrade Request Submitted on: Apr 09 Hi, the definition of Gamma is mentioned in the question (in sub 2). what's the problem in "replacing X^TX with Gamma^2"? they are different, X^T X is not 0 off diagonal, do an example if you don't believe me:) Reviewed on: Apr 09







- If you're enrolled in this course, you should have received an invitation to sign up sent to your @umontreal.ca email (yes even you have one, McGill infiltrators)

 Si vous êtes inscrits à ce cours, vous avez dû recevoir une invitation pour vous enregistrer sur Gradescope sur votre boîte e-mail @umontreal.ca
- Using Gradescope is easy, but some of you will still get it wrong! Sad!
 L'utilisation de Gradescope est facile. Faites-le bien!



- To remedy this, we have an easy first "test" assignment to get you used to it.
 - It is still for marks but not as many, and we will not punish you for trouble with gradescope.

Un devoir 0 vous sera donné pour vous habituer à utiliser Gradescope. Vous ne serez pas pénalisés pour une mauvaise utilisation de la plate-forme.

- We <u>will</u> punish you for all assignments afterwards e.g.
 - not all answer pages selected for a question
 - not adding all your group members to a group submission

Pour les devoirs suivants, des points vous seront retirés si vous utilisez mal la plate-forme, (si vous oubliez de sélectionner des pages pour une réponse à une certaine question, ou si vous oubliez d'ajouter tous les membres de votre groupe à la soumission)



 There is a written assignment that you <u>should</u> do in LaTeX, but all you <u>need</u> to submit is a PDF.

Il y a une partie écrite des devoirs, que vous devriez idéalement faire sur LaTeX. Vous **devez** soumettre un PDF. (N'oubliez pas de sélectionner les pages pour chaque question)

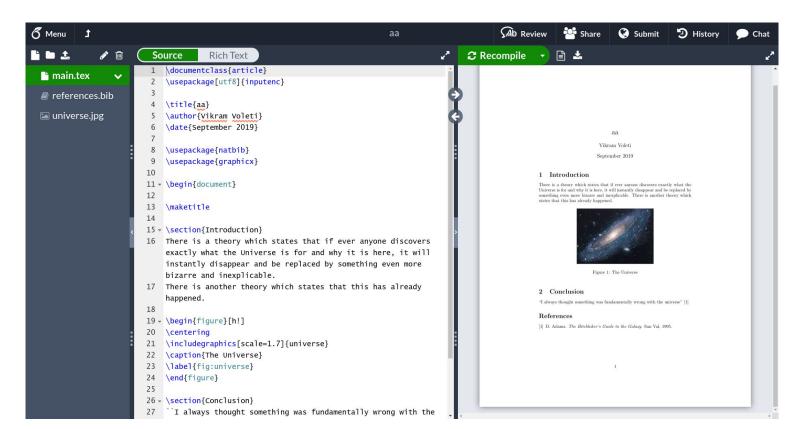
• There is a **programming** assignment that you <u>should</u> be using version control to keep track of (e.g. Github), but all you <u>need</u> to submit is a python file.

Il y a une partie programmation des devoirs, pour laquelle vous devriez idéalement utiliser Github. Vous **devez** soumettre un fichier python.

3. Overleaf (Latex)



overleaf.com



3. Overleaf (Latex)



LaTeX (pronounced "lay-tech"): Make a PDF using a program!

LaTeX: Créez un PDF avec un programme!

Extreme control, easy to use

Contrôle extrême. Facile à utiliser.

Latex CheatSheet: https://wch.github.io/latexsheet/



https://docs.conda.io/en/latest/miniconda.html

What is Miniconda?

Miniconda is the "mini" version of conda.

Ok, what is conda?

Conda is an open source package management system and environment management system that runs on Windows, macOS, and Linux. Conda quickly installs, runs and updates packages and their dependencies. Conda easily creates, saves, loads, and switches between environments on your local computer.

Miniconda est la version "mini" de conda. Conda permet de gérer vos environnements et vos librairies. Facile à installer, à utiliser, et permet de faire des mises-à-jour rapides de vos libraires. Il est facile de créer des environnements et d'alterner entre les différents environnements sur votre machine.



https://docs.conda.io/en/latest/miniconda.html

Why do you need it?

To run stuff in Python, many codes require Python packages (such as numpy, matplotlib) that come pre-installed with good functionalities! (For example, matplotlib can be used to visualize several kinds of data, like 2D points, 3D meshes, images, etc.)

Miniconda lets you install Python packages easily, and manage "environments".

Pour faire tourner votre code Python, vous aurez besoin de plusieurs librairies (qui viennent avec un tas de fonctionnalités très utiles). Miniconda vous permet d'installer ces librairies facilement, et de gérer vos environnements.



https://docs.conda.io/en/latest/miniconda.html

How to use it?

Install Miniconda using the link above

Installez Miniconda en utilisant le lien ci-dessus

Make a new virtual environment for your project: conda create -n proj1 python=3.7

Créez un nouvel environnement pour votre projet

• Activate this virtual environment: conda activate proj1

Activez cet environnement virtuel

- Install required Python packages, such as numpy and matplotlib: conda install numpy matplotlib
- Use it! (Launch python)

Installez les librairies Python nécessaires, puis utilisez votre environnement.



https://docs.conda.io/en/latest/miniconda.html

How to use it?

- Install Miniconda using the link above
- Make a new virtual environment for your project: conda create -n proj1 python=3.7
- Activate this virtual environment: conda activate proj1
- Install required Python packages, such as numpy and matplotlib: conda install numpy matplotlib
- Use it! (Launch python)

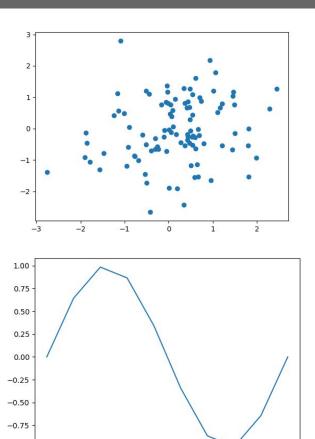
Good practice: Make a new virtual environment for each project - no conflicting packages among projects!

Pour éviter d'avoir des conflits entre vos librairies, créez un nouvel environnement virtuel pour chaque projet :)

5.python hello world.py



```
import matplotlib.pyplot as plt
import numpy as np
print("Hello World")
pts = np.random.randn(100, 2)
sudo mount -t ntfs-3g -o remove hiberfile /dev/sda2
/media/2A36EE3236EDFEA7
plt.show()
x = np.linspace(0, 2*np.pi, 10)
y = np.sin(x)
plt.plot(x, y)
```



-1.00

6. Jupyter notebook



https://jupyter.org/install.html

pip install jupyter

- Interactive Python "notebook"
- Very user-friendly for logging work flow of code
- Can plot graphs (and other visual data) within notebook itself!

7. Google Colab



https://colab.research.google.com/

Python notebook + GPU



Thank you.

