Course Overview

Introduction to Machine Learning – GIF-7015

Professor: Christian Gagné

Week 1

Introduction à l'apprentissage automatique

- GIF-4101: Optional course in computer engineering, computer science, statistic, mathematic-informatic, software engineering, mechanical engineering, geomatics engineering, etc.
  - Requirements
    - STT-2920 – Probabilités pour ingénieurs (or equivalent)
    - MAT-2930 – Algèbre linéaire appliqué (or equivalent)
- GIF-7005: graduate course, professional Master's in Computer Science – Artificial Intelligence, Master's and PhD in Electrical Engineering, Computer Science
  - Open to students from other Master's and PhD programs (sciences and engineering)
  - No official requirements, but must have basic knowledge in:
    - Linear algebra
    - Probability and statistics
    - Calculus
    - Algorithms and programming

- GIF-7015: official English version of GIF-7005 (**new** this year)
  - Slides and video capsules available in English
  - Howeworks, quizzes and exam provided in English
  - Bilingual text in monPortail (when possible)
  - Discussions on PAX forum in both languages
- In-class discussions (Wednesday morning) can be in English
  - By default, explanations will be given in French, but I will answer in the language the questions are asked
  - Do not hesitate to ask for repeating an answer in English when you do not get the French explanations
- Everything else is the same than GIF-7005

Course Organization

- Theoretical content presented as video capsules
  - Detailled presentation of content in slides
  - Video capsules presenting the content every week
- Weekly discussions in hybrid synchrone mode, Wednesday from 8:30 to 9:45 AM
  - In-person at the VCH-3820 (106 seats available)
  - Simultaneous broadcasting on Zoom, with session recording for later listening
  - Explanations, technical presentations, questions and answers

Evaluations

- Five quizzes at the end of Wednesday sessions (9:45-10:30 AM), on the content of the 2-3 last weeks
- One written exam on theoretical content and mathematical developments (November 9 in the evening)

	Date	Heure	Pond.	Format
Quiz 1	September 28, 2022	9:45-10:30 AM	4 %	Online objective questionnaire
Quiz 2	October 12, 2022	9:45-10:30 AM	4 %	Online objective questionnaire
Quiz 3	October 26, 2022	9:45-10:30 AM	4 %	Online objective questionnaire
Quiz 4	November 16, 2022	9:45-10:30 AM	4 %	Online objective questionnaire
Quiz 5	November 30, 2022	9:45-10:30 AM	4 %	Online objective questionnaire
Exam	November 9, 2022	6:30-8:20 PM	20 %	Paper questionnaire with answers to develop

Practical Exercises

- Five homeworks done in the PAX environment (Jupyter notebooks)
  - Undergraduate: homeworks in team of 2 students
  - Graduate students: individual homeworks
- Project on open topics
  - Teams of 3-5 students (mixing between the groups is allowed)
  - Presentation during a poster session, on December 21 and 22 in the morning
- Programming done on Python with scikit-learn and PyTorch

Task	Date	Hour	Weighting
Homework 1	October 7, 2022	12:00	8 %
Homework 2	October 21, 2022	12:00	8 %
Homework 3	November 11, 2022	12:00	8 %
Homework 4	November 25, 2022	12:00	8 %
Homework 5	December 9, 2022	12:00	8 %
Project	December 21-22, 2022	morning	20 %

Grades scales

- GIF-4101 (undergraduates): pass at **50 %**

A+:	$[100, 90]$	A:	$]90, 85]$	A-:	$]85, 80]$
B+:	$]80, 76]$	B:	$]76, 72]$	B-:	$]72, 68]$
C+:	$]68, 64]$	C:	$]64, 60]$	C-:	$]60, 56]$
D+:	$]56, 53]$	D:	$]53, 50]$	E:	$]50, 0]$

- GIF-7005 and GIF-7015 (graduates): pass at **60 %**

A+:	$[100, 90]$	A:	$]90, 85]$	A-:	$]85, 80]$
B+:	$]80, 76]$	B:	$]76, 72]$	B-:	$]72, 68]$
C+:	$]68, 64]$	C:	$]64, 60]$	E:	$]60, 0]$

Content (1/2)

- **Introduction (2h)**: types of learning; optimization; regularization; generalization; methodology.
- **Parametric methods (6h)**: discriminant functions; Bayes' theorem; parametric methods; maximum likelihood estimation; bias/variance trade-off; multivariate methods; mixture density; linear regression; multivariate regression.
- **Nonparametric methods (3h)**: density functions; kernel density estimation; nearest neighbor ranking; distance measures.
- **Linear discriminants (3h)**: gradient descent; multi-class separation; Perceptron algorithm; least squares method; logistic regression.
- **Kernel methods (3h)**: kernel functions; geometric margin maximization; support vector machines; examples of other kernel methods.
- **Multilayer Perceptron (3h)**: neural network model; error backpropagation algorithm.

Content (2/2)

- **Deep learning (6h)**: learning representations; techniques for learning deep networks; examples of applications.
- **Ensemble methods (3h)**: majority voting; error correction codes; expert mixtures; bagging; boosting; decision trees; random forests.
- **Model preprocessing and configuration (3h)**: principal component analysis; greedy forward/backward feature selection; heterogeneous data; missing variables; imputation; cross validation; hyper-parameter optimization.
- **Clustering (3h)**: K-means algorithm; expectation-maximization algorithm; hierarchical clustering; variety learning; multidimensional positioning.
- **Design and analysis of learning experiments (3h)**: design of experiments; confusion matrix; ROC curves; Bootstrap; error estimation; statistical tests; comparison of algorithms.

Sources

- Several references available for further study
  - No book is required
  - Most of the references are available in digital format

📖 Ethem Alpaydin. *Introduction to Machine Learning*. 4th edition, MIT press, 2020.

📖 Christopher M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
`https://www.microsoft.com/en-us/research/people/cmbishop/prml-book/`

📖 Trevor Hastie, Robert Tibshirani et Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. 2nd edition, Springer, 2009.
`https://web.stanford.edu/~hastie/ElemStatLearn/`

📖 Ian Goodfellow, Yoshua Bengio et Aaron Courville. *Deep Learning*. MIT press, 2016.
`http://www.deeplearningbook.org/`

📖 Andriy Burkov. *The Hundred-Page Machine Learning Book en français*. 2019.
`http://themlbook.com/wiki/doku.php`

Professor

- Name: Christian Gagné
- Office: PLT-1138-F
- Email: christian.gagne@gel.ulaval.ca
- Web: `https://vision.gel.ulaval.ca/~cgagne`
- Course assistance:
  - Questions and discussions of general interest on the subject matter and evaluations: PAX course forum
  - More specific or personal considerations: email the professor

Web sites

- monPortail site: `https://sitescours.monportail.ulaval.ca/ena/site/accueil?idSite=146163`
  - Course overview
  - Slides and video capsules
  - Online quizzes
  - Grades for homeworks and exams
- PAX's website: `https://pax.ulaval.ca/GIF-4101-7005/A22/`
  - Homeworks (Jupyter notebook)
  - Discussion forum
- YouTube channel:
  `https://www.youtube.com/channel/UCkMXqYnOzhIlRseUOgq8_Xw`
  - Video capsules