



Seminar: Advanced Topics in Machine Learning and Data Science

Fernando Perez-Cruz

Feb 24th, 2021



Introduction

Goal of this Seminar Course

- Learn about advanced topics in machine learning.
- Share what you have learnt by giving a scientific presentation.

Today

- Present the 20 papers.
- Algorithm for assigning the papers.
- Course organization and grading.

Course Organization

- Lectures on Wednesdays 16:00 – 18:00 @ LFW E13 Zoom: 939 1921 5655
- Lecture days: 10/03, 17/03, 24/03, 31/03, 14/04, 21/04, 05/05, 12/05, 19/05 and 26/05.
- You are expected to attend all lectures and present in one.
- First presentation will be in two weeks.
- No lectures on **03/03 and 28/04**.
- I have reserved two 30-min slots to discuss with each one of you before your presentation:
 - Wednesdays at 10:00 and 10:30. One week before the presentation.
 - Mondays at 16:00 and 16:30. Two days before the presentation.

Grading

- Your grade will be determined based on your talk, as well as participation in the discussion.
- Criteria:
 - ① **Structure:** how well is your talk organized?
 - ② **Understandability:** how understandable is your oral presentation and slide design?
 - ③ **Completeness:** how well do you provide right background, and manage to focus on what is important and relevant?
 - ④ **Engagement:** how engaged are you in class and in the talk preparation?
 - ⑤ **Independence:** how independent are you in preparing the presentation, and in reflecting on the paper?
 - ⑥ **Class participation:** Questions during the lectures?

Presentation details

- Use electronic slides (ppt, pdf, ...).
- Talk length: **30 min** + 15 min discussion.
- The talk should provide sufficient background to be understandable to someone who has taken an ML class.
- You should present:
 - ① Main Contributions.
 - ② Results.
 - ③ Critic.

Meta Presentation

- **Introduction:** Motivation.
- **Problem statement:** Notation and main contribution.
- **Technical contribution:** details of the contribution.
- **Experiments** (if applicable).
- **Your take:** Why did you like the paper? (or not) what are you missing? What's wrong? ...

Paper Assignment

- You have a list of papers in:
<https://github.com/fernandoperezc/Advanced-Topics-in-Machine-Learning-and-Data-Science>.
- Please select at least five papers in this survey:
<https://forms.gle/VEXC9v3GaRxTSWoVA>.
- You will also need to select at least 5 prefer dates.
- I will assign the papers first to the students that select the earlier dates.
- I will try to maximize your preferred paper ranking.
- I might move you to a later date, but not to an earlier date.
- I will finalize the assignment on Sunday Feb 28th.
- I will send you an invite to the lecture and an invite to two 30-mins meetings prior to the presentation.

Papers 1

- Reconciling modern machine-learning practice and the classical bias-variance trade-off.
 - A Kernel Two-Sample Test.
 - Natural Graph Networks.
 - On Calibration of Modern Neural Networks.
 - Density estimation using Real NVP.
 - Do Deep Generative Models Know What They don't Know?
 - SurVAE Flows: Surjections to Bridge the Gap between VAEs and Flows.
 - Flows for simultaneous manifold learning and density estimation.
 - Nonparametric graphical model for counts.
 - Bayesian Learning of Sum-Product Networks.

Papers 2

- Fair Decisions Despite Imperfect Predictions.
- Towards A Rigorous Science of Interpretable Machine Learning.
- Elements of Causal Inference (Chapters 1-2 and 3-4).
- Language Models are Few-Shot Learners.
- Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR.
- Shopper: A Probabilistic Model of Consumer Choice with Substitutes and Complements.
- Using Deep Neural Networks as Cost-Effective Surrogate Models for Super-Parameterized E3SM Radiative Transfer.
- Learning to Detect Sepsis with a Multitask Gaussian Process RNN Classifier.
- Inferring causation from time series in Earth system sciences.