



# Seminar: Advanced Topics in Machine Learning and Data Science

Fernando Perez-Cruz

Feb 26th, 2020



# 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
- Lecture days: 11/03, 18/03, 25/03, 08/03, 22/03, 29/03, 06/05, 13/05, 20/05 and 20/05.
- You are expected to attend all lectures and present in one.
- First presentation will be in two weeks.
- No lectures on **04/03 and 01/04**.
- I have reserved two 30-min slots to discuss with each one of you before your 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?

# 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/3foWpxtLinAnBvCLA>.
- 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 Friday Feb 28th.
- I will send you an invite to the lecture and an invite to two 30-mins meetings prior to the presentation.
- Luis Salamanca and Christian Donner will help with some papers.

# Papers 1

- Integer Discrete Flows and Lossless Compression.
- Universal Domain Adaptation.
- Asymmetric Tri-training for Unsupervised Domain Adaptation.
- Adaptive and Safe Bayesian Optimization in High Dimensions via One-Dimensional Subspaces.
- Combinatorial Bayesian Optimization using the Graph Cartesian Product.
- Variational Inference: A Review for Statisticians.
- Uniform convergence may be unable to explain generalization in deep learning.
- In Search of the Real Inductive Bias: On the Role of Implicit Regularization in Deep Learning.
- The Role of Over-Parametrization in Generalization of Neural Networks.
- Fairness Constraints: Mechanisms for Fair Classification.

## Papers 2

- Overparameterized Neural Networks Can Implement Associative Memory.
- Only Bayes should learn a manifold.
- Prescribed Generative Adversarial Networks.
- Federated Optimization: Distributed Machine Learning for On-Device Intelligence.
- Differentially Private Regression with Gaussian Processes.
- A Fast Learning Algorithm for Deep Belief Nets.
- Set Transformer: A Framework for Attention-based Permutation-Invariant Neural Networks.
- Geometric Deep Learning: Going beyond Euclidean data.
- Gaussian Process Behaviour in Wide Deep Neural Networks.
- A Kernel Two-Sample Test.