ADVANCED MACHINE LEARNING

Mini-Project Overview /
In Class Paper Reading

Lecture : Prof. Aude Billard (aude.billard@epfl.ch)

Teaching Assistants :
Lukas Huber, Bernardo Fichera, ,Thomas Pethick

Two Options (50 % of the Grade)

- In class paper reading + presentation

- Coding mini-project + report

In Class Paper Reading

Individual Project

- Pick from the list of paper provided
- Read the paper & understand it in depth (context, impact, implications)
- Present the paper to the class
- Q&A session to paper and related topics

In Class Paper Reading

Topics for survey will entail :

- Dimensionality Reduction
- Manifold learning
- Clustering
- Classification
- Regression
- Reinforcement Learning
- Time Series Analysis
- Deep Learning
- General Machine Learning Topics,
  e.g. Large scale learning, Imbalanced Datasets

Mini-Project (2 Students)

- Review an algorithm you have seen in class
- Learn about an algorithm you haven’t seen in class
- Test your understanding of machine learning concepts by comparing these two algorithms
- Write a qualitative and quantitative assessment of your comparison

What does this mean?

What you should do:

- Pick relevant dataset
  Apply the algorithms to these datasets.
- Discuss the hyper-parameter selection, training, CV...
- Note and explain the differences you observe.
- Talk about what conditions lead to an algorithms being better than the other.
- Use metrics! If you see a difference try to measure it!

What you should not do:

- Implement everything from scratch.
- Copy reports from previous years. (yes, we do check)

In Class Paper Reading

Main idea: compare two machine learning algorithms

Topics for survey will entail :

- Dimensionality Reduction / Manifold learning
- Clustering
- Classification
- Regression

Mini-Projects Requirements

Coding:
Self-contained piece of code in:
- Matlab
- Python
- C/C++

Including:
- Demo scripts
- Datasets
- Systematic assessment.

Report:
Algorithm analysis, including but not limited to:
- Number/sensitivity to hyper-parameters
- Computational costs train/test
- Growth of computation cost w.r.t. dataset dimension
- Sensitivity to non-uniformity/non-convexity in data.
- Precision of regression
- Benefits/disadvantages of algorithm w.r.t. to different types of data/applications.
- …

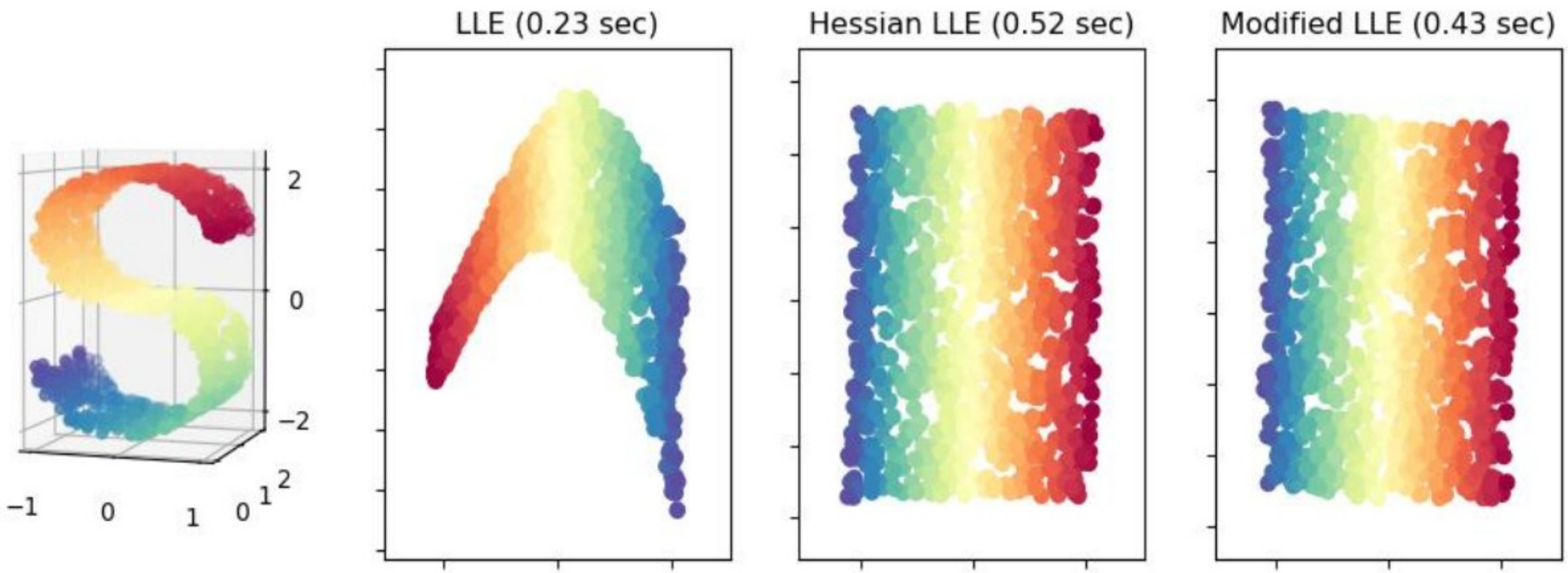
Find relevant datasets to study your algorithms (for code projects)

Example of sites to look for datasets:

- UCI: https://archive.ics.uci.edu/ml/datasets.html
- Kaggle: https://www.kaggle.com/datasets

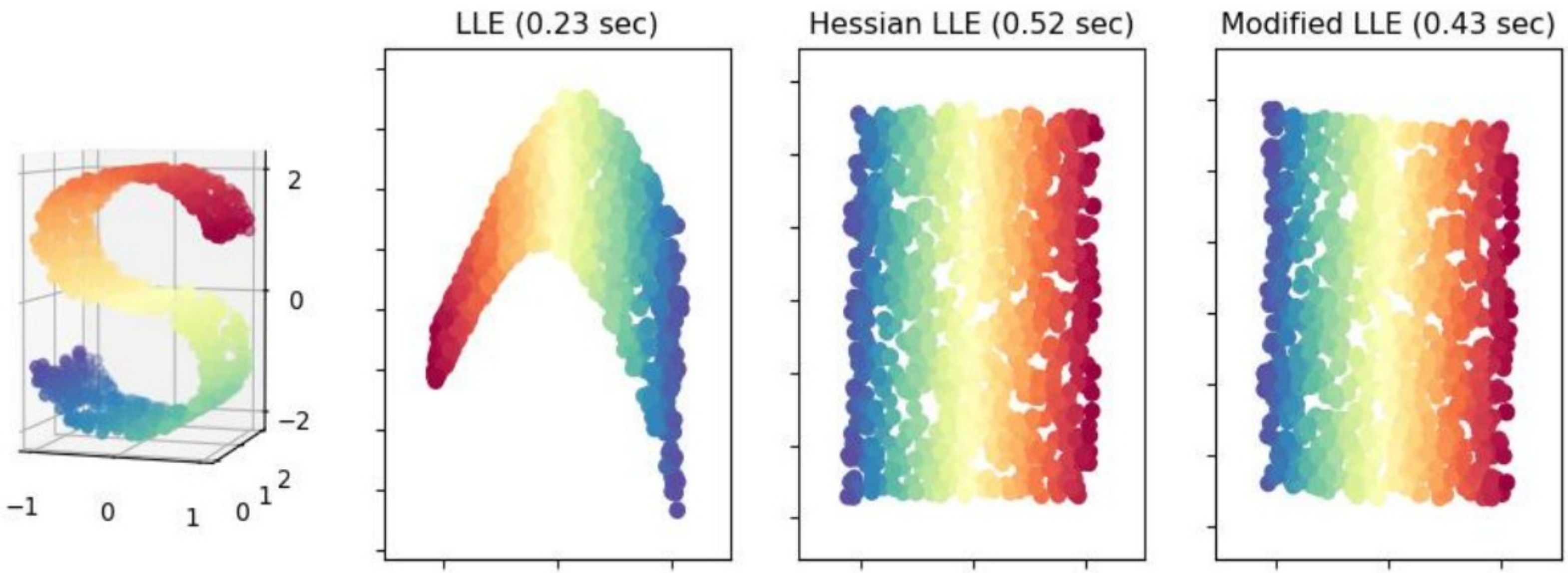
Dimensionality reduction / Manifold learning projects(1)

- **LLE** (Locally Linear Embedding)
  Roweis, Sam T., and Lawrence K. Saul. "Nonlinear dimensionality reduction by locally linear embedding." science 290.5500 (2000): 2323-2326.
- **MLLE** (Modified LLE)
  Zhang, Zhenyue, and Jing Wang. "MLLE: Modified locally linear embedding using multiple weights." Advances in neural information processing systems. 2007.
- **HLLE** (Hessian LLE)
  Donoho, David L., and Carrie Grimes. "Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data." Proceedings of the National Academy of Sciences 100.10 (2003): 5591-5596.



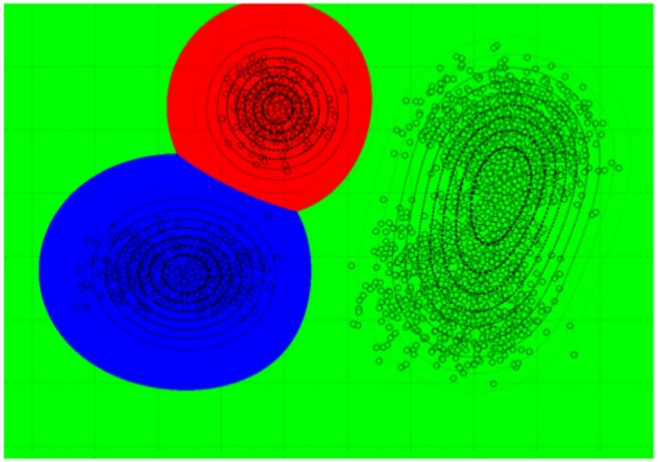
Dimensionality reduction / Manifold learning projects (2)

- **SNE** (Stochastic Neighbor Embedding)
  Hinton, Geoffrey E., and Sam T. Roweis. "Stochastic neighbor embedding." Advances in neural information processing systems. 2003.
- **tSNE** (t-distributed SNE)
  Maaten, Laurens van der, and Geoffrey Hinton. "Visualizing data using t-SNE." Journal of machine learning research 9.Nov (2008): 2579-2605.
- **GPLVM** (Gaussian Process Latent Variable Models)
  Lawrence, Neil. "Probabilistic non-linear principal component analysis with Gaussian process latent variable models." Journal of machine learning research 6.Nov (2005): 1783-1816.

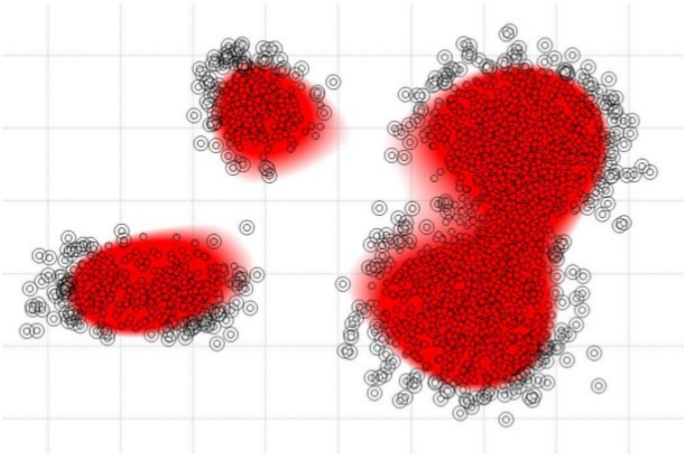


Clustering

- **Kernel K-means**
  Welling, Max. "Kernel K-means and Spectral Clustering." 2013-03-15]. http://www, ics. uci. edu/-welling/teaching/273 ASpring09/SpectralClustering. pdf. (Resource for the Advanced ML class)
- **SV Clustering** (Support Vector Clustering)
  Ben-Hur, Asa, et al. "Support vector clustering." Journal of machine learning research 2.Dec (2001): 125-137. (Resource for the Advanced ML class)



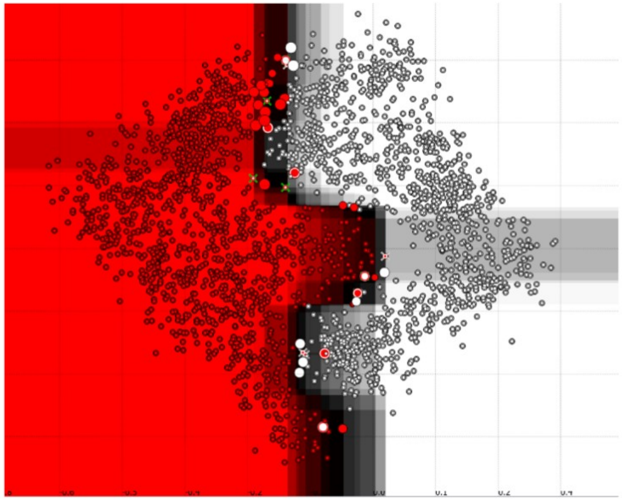
Kernel K-means
(RBF kernel,
all 2660 points used)



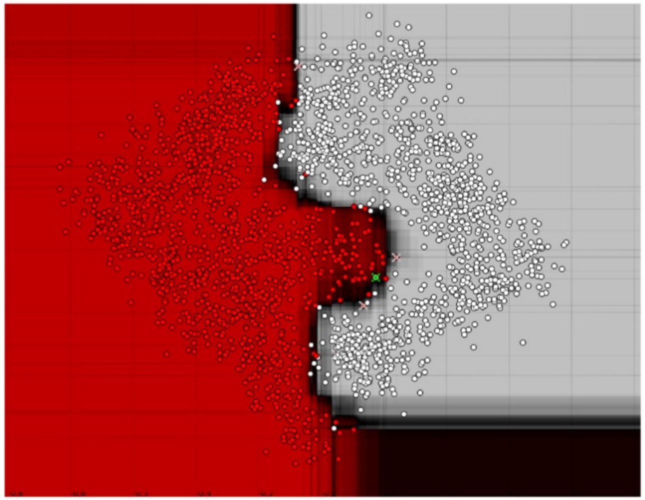
Support Vector Clustering
(RBF kernel,
355 SVs on 2660 points)

Classification

- **AdaBoost**
  Freund, Yoav, and Robert E. Schapire. "A decision-theoretic generalization of on-line learning and an application to boosting." Journal of computer and system sciences 55.1 (1997): 119-139.
- **RTF** (Random Tree Forests)
  Breiman, Leo. "Random forests." Machine learning 45.1 (2001): 5-32.
  `https://www.stat.berkeley.edu/users/breiman/RandomForests/cc_home.htm`
- **Gaussian Process Classification**
  Williams, Christopher KI. "Prediction with Gaussian processes: From linear regression to linear prediction and beyond." Learning in graphical models. Springer, Dordrecht, 1998. 599-621.
- **Feed-forward Neural Networks**
  Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors." nature 323.6088 (1986): 533-536.



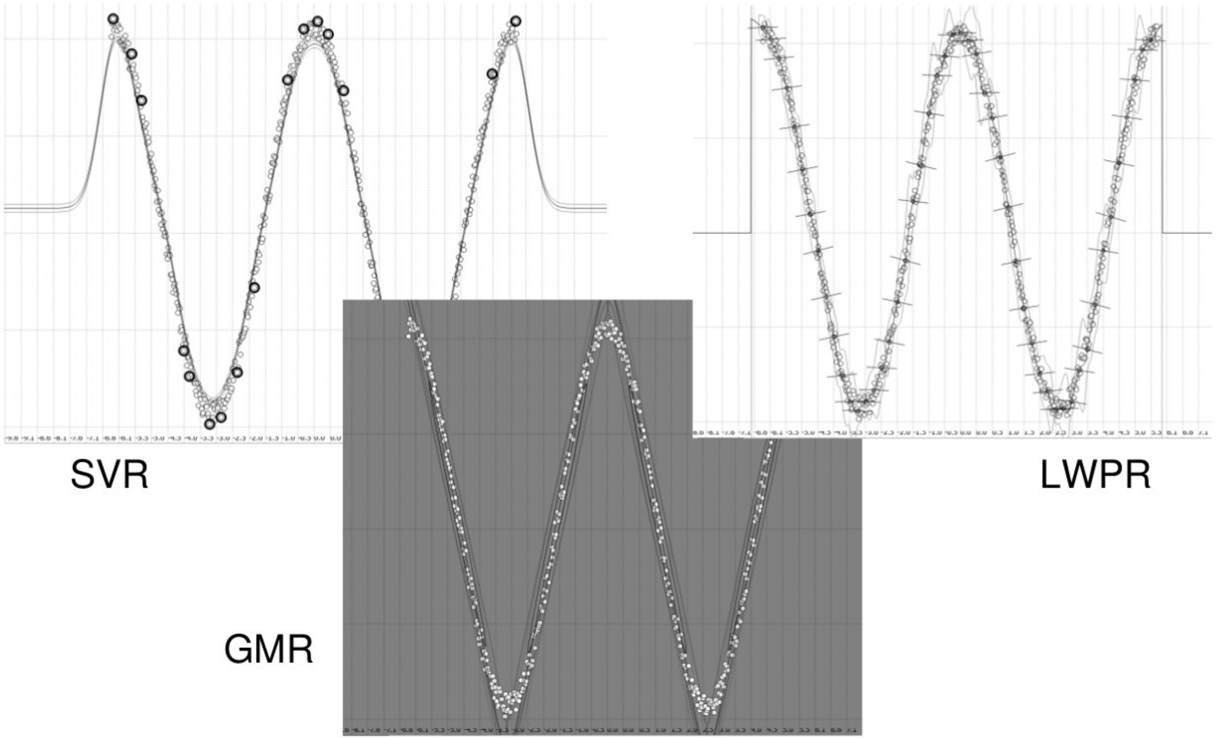
AdaBoost
(80 decision stumps)



Random Forest
(50 trees)

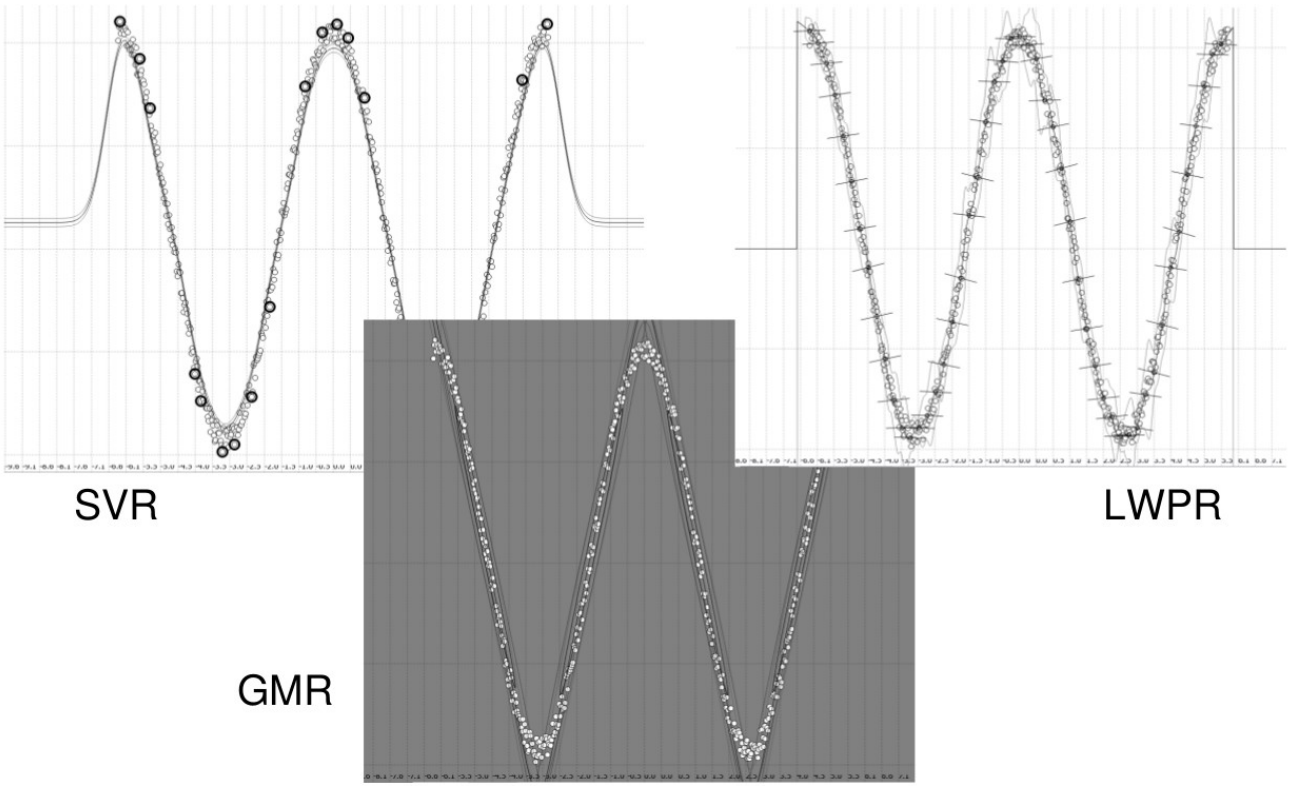
Regression (1)

- **SVR** (Support Vector Regression)
  Drucker, Harris, et al. "Support vector regression machines." Advances in neural information processing systems. 1997.
- **LWPR** (Locally Weighted Projection Regression)
  Vijayakumar, Sethu, and Stefan Schaal. "Locally weighted projection regression: An O (n) algorithm for incremental real time learning in high dimensional space." Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000). Vol. 1. 2000.



Regression (2)

- **GMR** (Gaussian Mixture Regression)
  Sung, Hsi Guang. Gaussian mixture regression and classification. Diss. Rice University, 2004.
- **GPR** (Gaussian Process Regression)
  Williams, Christopher KI. "Prediction with Gaussian processes: From linear regression to linear prediction and beyond." Learning in graphical models. Springer, Dordrecht, 1998. 599-621.



#TODO before March 6th

- Select a project from the available ones (code/in-class paper)
- Official project registration opens this Friday afternoon. We will send an forum post.
- Maximum of 2 teams per code project. / Maximum 1 person for each in-class paper.

> Choose dataset and start project