

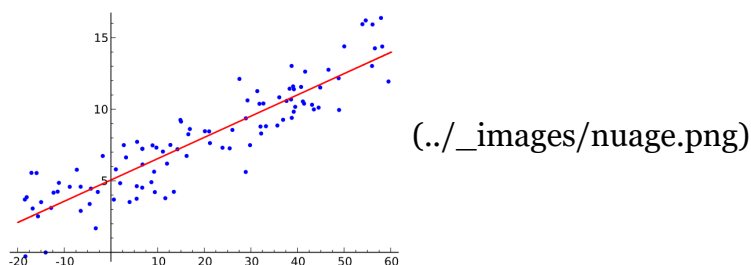
# Lectures sur le machine learning

Le machine learning ([https://fr.wikipedia.org/wiki/Apprentissage\\_automatique](https://fr.wikipedia.org/wiki/Apprentissage_automatique)) répond à de plus en plus de problématiques, beaucoup d'entre elles sont citées sur ce cours Python pour un Data Scientist / Economiste

([http://www.xavierdupre.fr/app/ensae\\_teaching\\_cs/helpsphinx3/td\\_2a.html](http://www.xavierdupre.fr/app/ensae_teaching_cs/helpsphinx3/td_2a.html)) proposé à l'ENSAE (<http://www.ensae.fr/ensae/fr/>) qui comprend des exercices, des articles et les modules qui implémentent quelques algorithmes. Quelques thèmes liés au machine learning ([https://fr.wikipedia.org/wiki/Apprentissage\\_automatique](https://fr.wikipedia.org/wiki/Apprentissage_automatique)) seront abordés en partant d'un jeu de données pour explorer quelques points récurrents ou techniques.

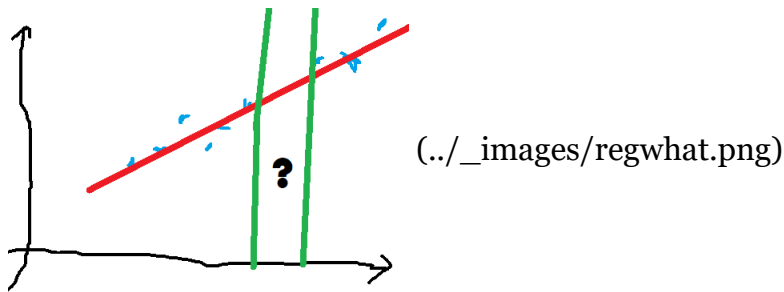
- Classification et régression (regclass.html)
- Comprendre un modèle (digmodel.html)
- Prétraitement des données, données textuelles (preprocessing.html)
- Pipeline de traitement de données (pipeline.html)
- Valeurs manquantes, problèmes déséquilibrés, overfitting (overfitting\_unbalanced\_missing.html)
- Machine Learning non supervisé (nonsupervise.html)
- Quelques autres problèmes de machine learning (otherml.html)
- Gradient et méthodes ensemblistes (gradienttree.html)
- Séries temporelles (seriestemp.html)
- Images et deep learning (imagedeep.html)
- Cartes (cartes.html)

Le machine learning cache bien des choses mais au final il s'agit de calculer une prédiction pour un événement jamais observé. Ces techniques synthétisent les observations passées pour en déduire une réponse probable sur des observations futures. Le nuage de points suivant représente ce qui a été observé jusqu'à présent.



Le machine learning ([https://fr.wikipedia.org/wiki/Apprentissage\\_automatique](https://fr.wikipedia.org/wiki/Apprentissage_automatique)) propose une façon de construire une fonction qui résume ou modélise les observations passées. Le résultat obtenu fait en quelque sorte une synthèse du passé pour inférer une réponse probable à des événements jamais observés. Dans le cas du nuage de point précédent, on souhaite calculer une prédiction  $y$  pour un  $x$

différent. Le machine learning prolonge un savoir fait de résultat d'expérience. Autrement dit, le modèle de machine learning pourra retourner une prédiction pour l'intervalle  $x \in [50.105, 50.108]$  alors que nous ne connaissons la réponse qu'aux extrémités de cet intervalle.



Source (../\_sources/lectures/index.rst.txt)

[Back to top](#)

Mis à jour le 2019-10-24.