Pour ma présentation:

Ajouter les outils, mes références, mes livres de référence.

* + Data Science Games : *Data Science : Fondamentaux et études de cas. É. Biernat, M. Lutz, 2012, Eyrolles, 2016*
  + Outils : python sklearn, pandas & xgboost;   
    Stacknet *Marios Michailidis (2017), StackNet, StackNet Meta Modelling Framework, url <https://github.com/kaz-Anova/StackNet>*
* Intégration des identifiants dans la base d’apprentissage :
  + Séparation de la base d’apprentissage en 5, pour chaque partie, calculer la moyenne de la cible pour chaque identifiant (le nombre de chanson écouté jusqu’au bout/ nombre de chanson vues au total). Puis remplacer l’identifiant par cette valeur. Concaténer les 5 parties.
* Régression Quantile, package Scipy:
  1. *Algorithms for Adaptive Quantile Regression, 2006, Jan Kloppenborg Møller, Henrik, Aalborg , Nielsen and Henrik Madsen*
  2. *R. Koenker,2000, QUANTILE REGRESSION AN INTRODUCTION, Journal of Economic Perspectives*
* Electrolux:
  + Outils:
  + *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media TextC.J. HuttoEric Gilbert.*

Pour la thèse:

* Pour les données, parler à Bastien de l'exploitation des datas d'Engie pour thèse. Présentation et on verra fin Mai.
* Construire un exemple jouet. Type Minst
  + Essayer des choses, avec Keras En cours
  + Regarder kaggle avec Concept Drift Pas concluant.   
    Une Méthode revient cependant. Beaucoup de défis Kaggle propose des données d’entrainement et de test. La méthode consiste à considérer chacun des agrégats comme la cible tour à tour. Puis entrainer un modèle prédictif et sur les données d’entrainement sur l’agrégat. Si les prédiction sur le jeu de test sont sous un seuil alors on considère qu’un drift est présent dans les données.
  + Regarder IBM Open scaleregarder l'implémentation Pas d’information si ce n’est que c’est un modèle
* Regarder Google Scolar pour voir qui site les articles
  + N'importe qui peut mettre un article sur arxiv.org
  + Si conf c'est mieux
* Est-ce que notre problématique est vraiment celle du concept drift ? Oui.
  + 3 types de drift.
* Pour fin mai, faire présentation de 3-4 papiers, dont 1 où je vais plus loin.
* Meeting 7 Mai, même heure.