

# Entraînement 2

Deuxième entraînement sur le jeux de données contenant 100 000 notes. Il y a 1643 films et 943 utilisateurs. Pour permettre une cross-validation les jeux de données sont séparés en 5 sets. Pour chaque set nous avons un set d'apprentissage avec 80% des données et un set de test avec les 20% restant.

---

## Configuration système

Les configurations système sont les suivantes:

- Ubuntu 14.04 LTS
- 1800 Mo de mémoire RAM
- CPU single core

### Note:

L'environnement d'exécution est une machine virtuelle

## Configuration de l'IA

Les paramètres de l'algorithme sont les suivants pour l'ensemble des 5 entraînements:

- $\alpha = 0.001$
- $\lambda = 1/100$
- seuil du taux d'erreur: 0.1

### Note:

Le seuil du taux d'erreur est la condition d'arrêt de la descente de gradient, dès que le taux d'erreur passe en dessous on s'arrête.

## Rapport

### U1

- erreur en fin d'apprentissage: 0.0952622
- temps d'exécution: 27.159s
- consommation mémoire: 50Mo (maximum)

## U2

- erreur en fin d'apprentissage: 0.0661727
- temps d'exécution: 18.938s
- consommation mémoire: 50 Mo (maximum)

## U3

- erreur en fin d'apprentissage: 0.0970682
- temps d'exécution: 12.189s
- consommation mémoire: 50 Mo (maximum)

## U4

- erreur en fin d'apprentissage: 0.0653083
- temps d'exécution: 17.727s
- consommation mémoire: 50 Mo (maximum)

## U5

- erreur en fin d'apprentissage: 0.0935517
- temps d'exécution: 14.901s
- consommation mémoire: 50 Mo (maximum)

# Cross Validation

## U1

- taux d'erreur: 60%

## U2

- taux d'erreur: 59%

## U3

- taux d'erreur: 60%

## U4

- taux d'erreur: 59%

## U5

- taux d'erreur: 60%

Le taux d'erreur globale pour cette phase de test est donc de 59%, c'est médiocre.