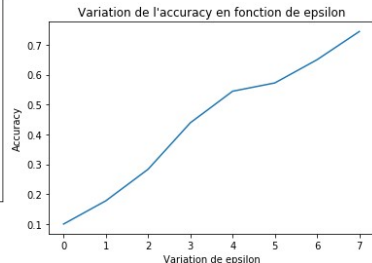
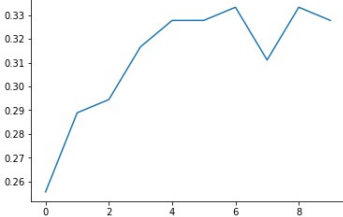


I. Régression supervisée

I. Prediction du vote_aveage d'un film à partir de vote_count, popularity et les catégories :

Classifieur KNN : Construction d'une base de données contenant pour chaque films ses catégories avec dummy-coding, les vote_count et la popularity et l'objectif étant de prédire le vote_aveage de chaque film :

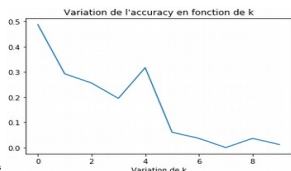
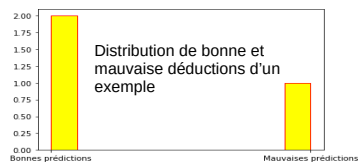


II. Classification supervisée

I. Classification des films pour un utilisateur (recommandé ou pas) :

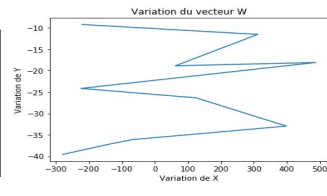
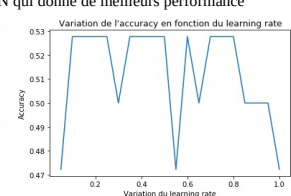
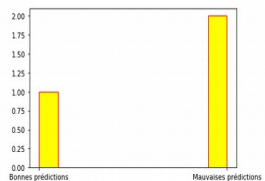
La construction du labeledSet est faite à partir des categories(dummy-coding), vote_count, popularity et vote average le but étant de prédire ce le film serait bon à recommander à un utilisateur :

Classifieur KNN :

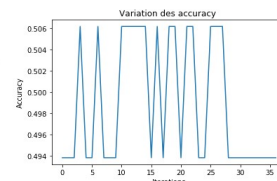
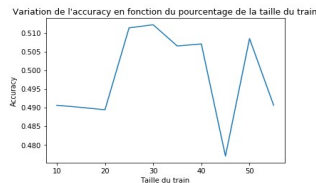


L'Accuracy du KNN est souvent bonne pour des valeurs de au dessous de 4, et le faite d'augmenter les voisins testés mène à fausser les résultats

Classifieur Perceptron : L'accuracy du perceptron varie entre 0.48 et 0.51 elle est donc moyenne, moins bonne que celle du KNN qui donne de meilleurs performance

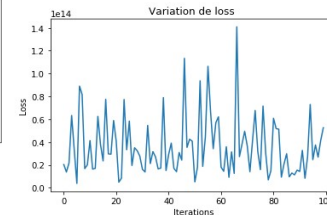


Variation de l'Accuracy :



La variation de l'Accuracy lors des train montre bien que celle ci ne se stabilise pas puisqu'elle dépend de l'ordre de visite des points

Le cout du Perceptron :



le Perceptron ne semble pas donner de bon résultats avec a un grand cout (dépassé les 50%) et ne stabilise pas.

III. Classification non supervisée :

Apprentissage Non Supervisé

Classifier des acteurs
Selon des critères
inconnus

Base de donnée utilisée :

	Categorie j de films
Acteur I	nombre de films

KMoyennes

Exprimenter pour
Trouver le meilleur k
Selon l'inertie

