

ft\_linear\_regression

Une introduction au machine learning

Résumé: Dans ce projet vous implementerez votre premier algorithme de machine learning.

# Table des matières

1	Preambule	2
II	Introduction	3
III	Objectif	4
IV	Instructions Generales	5
$\mathbf{V}$	Ce que vous devez faire	6
VI	Bonuses	7
VII	Peer-evaluation	8

# Chapitre I Préambule

Ce que je pense etre la meilleur definition du machine learning :

"On dit d'un programme informatique qu'il apprend d'une expérience E en respectant les règles d'une tâche T et en mesurant les performances via P si ses performances à la tâche T, telles que mesurées par P, s'améliorent avec l'expérience E."

Tom M. Mitchell

#### Chapitre II

#### Introduction

Machine learning est un des champs de l'informatique qui est en pleine expansion. Ca à l'air compliqué et reservé aux mathematiciens au premier abord. Vous avez surement entendu parler de réseaux de neurone ou bien k-moyennes et vous ne comprenez pas encore comment ca marche ou comment developper de tel algorithme.

Ne vous inquietez pas, on va commencer par les bases, un simple algo de machine learning

# Chapitre III Objectif

Le but de se projet est de vous faire une introduction au concepte de base du machine learning. Pour ce projet vous devrez creer un programme qui predit le prix d une voiture en utilisant l'entrainement par fonction lineaire avec un algorithme du gradient.

Nous travaillerons sur un exemple precis pour ce projet, mais une fois que vous aurez fini vous pourrez l'utiliser sur d'autre jeu de donnée.

## Chapitre IV

#### Instructions Generales

Dans ce projet vous etre libre d'utiliser n'importe quel langage de programmation tant qu'il ne font pas tout le travaille pour vous.

Par exemple, python et numpy.polyfit sera considéré comme de la triche.



Utilisez un language ou il est facile de visualisé des données ca vous aidera à debugguer.

#### Chapitre V

#### Ce que vous devez faire

Implementez un algorithme de regression lineaire sur un seul element, en l'occurence le kilométrage d'une voiture Pour ce faire vous devez faire 2 programmes :

• Le premier programme sera utilisé pour predire le prix d une voiture en fonction de son kilométrage. Quand vous lancerez le programme, celui ci vous demandera le kilométrage et devrait vous donner un prix approximatif de la voiture en utilisant l'hypothese suivante :

$$prixEstime(kilom\acute{e}trage) = \theta_0 + (\theta_1 * kilom\acute{e}trage)$$

Avant de lancer le programme d'entrainement, theta0 et theta1 auront pour valeur 0.

• Le second program sera utilisé pour entrainer votre modele. Il lira le jeu de donné et fera une regression lineaire sur ces données.

Une fois la regression lineaire fini, vous sauvegarderez la valeur de theta0 et theta1 pour pouvoir l'utiliser dans le premier programme.

Vous utiliserez la formule suivante :

$$tmp\theta_0 = ratioDApprentissage * \frac{1}{m} \sum_{i=0}^{m-1} (prixEstime(kilom\acute{e}trage[i]) - prix[i])$$

$$tmp\theta_1 = ratioDApprentissage* \frac{1}{m} \sum_{i=0}^{m-1} (prixEstime(kilom\acute{e}trage[i]) - prix[i])*kilom\acute{e}trage[i] + prix[i])*kilom\acute{e}trage[i] + prix[i] * kilom\acute{e}trage[i] + prix[i] * k$$

Je vous laisse devinez ce que m est :)

Veuillez noter que le prixEstime est la meme chose que dans notre premier programme, mais ici il utilise vos valeures temporaires afin de calculer theta0 et theta1. Attention a bien mettre a jour theta0 et theta1 en meme temps.

# Chapitre VI

#### Bonuses

Voici des bonus qui pourrait etre utils :

- Visualisé les données sur un graph avec leur repartition.
- Affichez la ligne resultant de votre regression lineaire sur ce meme graphe et voir si ca marche!
- Un programme qui verifie la precision de votre algorithme.

... Et n importe quoi qui pourrait rendre cet exercice encore meilleur.

## Chapitre VII

### Peer-evaluation

Votre programme ne sera corrigé que par des humains, vous pouvez donc organiser votre code comme bon vous semble.

Voici les points qui seront verifié par votre peer-correcteur :

- L'absence de librairie ou autre qui fasse tout le travail a votre place.
- L'utilisation de l'hypothese donnée plus haut
- L'utilisation de la fonction d entrainement donnée plus haut.