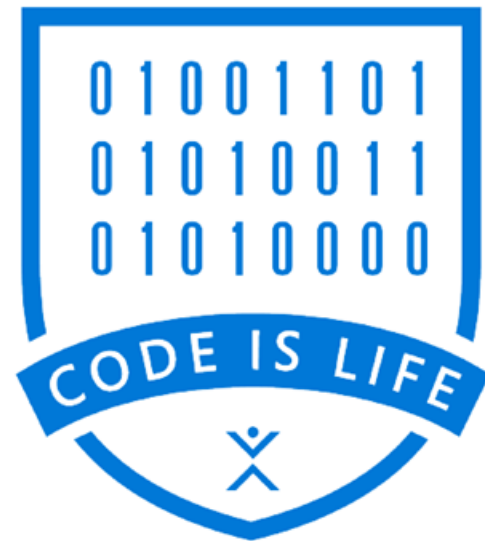


Microsoft Student Partners

Linear Algebra For AI [P3]

Nathan Bangwa



- Learning Model (rappel)
- C'est quoi la régression linéaire.
- Types de variables
- Interpolation & Extrapolation
- Cost and Optimization
- Applications

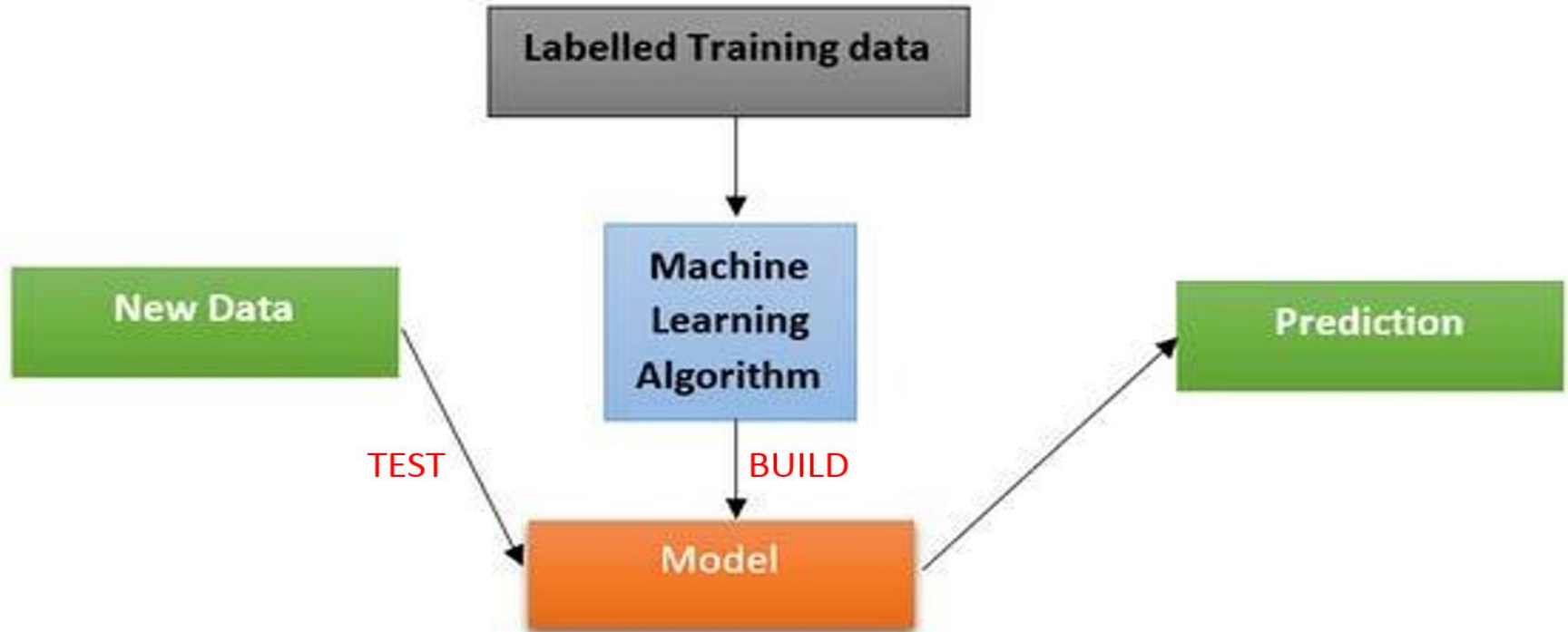
What's Learning Model



- Un modèle d'apprentissage est soit une fonction mathématique (ou logique) simple ou composée, soit une logique(schema) à suivre pour réaliser une tâche.
- C'est à la machine de trouver ce modèle d'apprentissage (fonction ou logique à suivre) partant de données qui lui sont fournies (supervised & unsupervised) ou celles qu'elle récupère dans un environnement donné (Reinforcing).

Cfr Linear Algebra For AI [P2]

Learning Schema



What's Linear Regression??



En statistiques, en économétrie et en apprentissage automatique, un modèle de régression linéaire est un modèle de régression qui cherche à établir une relation linéaire entre une variable, dite expliquée, et une ou plusieurs variables, dites explicatives. [Wikipédia](#)

MACHINE LEARNING

Apprentissage supervisé

Classification

Support Vector Machine (SVM)

Naive Bayes

Nearest Neighbor

Régression logistique

Arbre de classification

...

Régression

Régression linéaire

Arbre de régression

Réseaux neuronaux

Support Vector Regression

...

Apprentissage non-supervisé

Association

A priori

...

Clustering

Réseaux neuronaux

K-means

Hierarchique

A priori

...

En statistiques, en économétrie et en apprentissage automatique, un modèle de régression linéaire est un modèle de régression qui cherche à établir une relation linéaire entre une variable, dite expliquée, et une ou plusieurs variables, dites explicatives.

[Wikipédia](#)

Variables

- Une **variable indépendante** dans un problème est un paramètre du problème qui varie sans être influencé par les autres paramètres du problème. Cela correspond le plus souvent aux paramètres exogènes ou imposés par la nature. [Wikipédia](#)
- Une **variable dépendante** dans un problème est un paramètre du problème qui varie sous l'influence d'autres paramètres du système. Cela correspond le plus souvent aux paramètres endogènes, qu'on cherche à caractériser. [Wikipédia](#)

Variables Exemples



| Variable indépendante Ou explicative | Variable dépendante Ou expliquée | X : data abscisses | Y : label ordonnées |
|---|-------------------------------------|--------------------|---------------------|
| Experience (âge d'une personne) dans l'entreprise | Salaire obtenu | 1 | 3 |
| Surface d'un terrain | Prix | 2 | 5 |
| Investissement | Benefices | 3 | 7 |
| ... | ... | | |

$$\text{Label} = 2 * \text{data} + 1$$

$$Y = aX + b$$

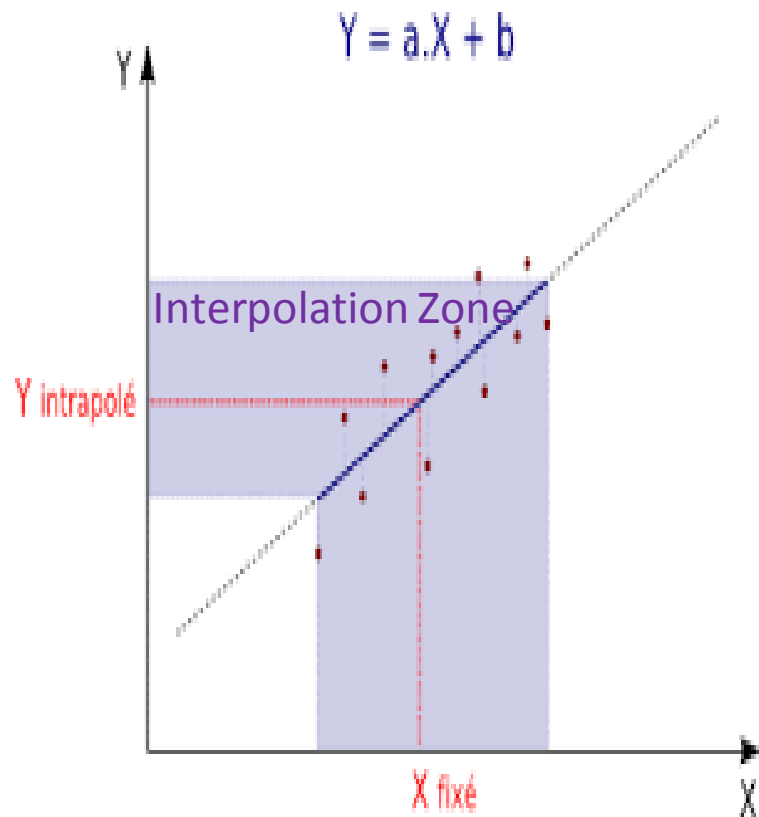
$$Y = f(X)$$

Interpolation & Extrapolation

Simple Linear Regression



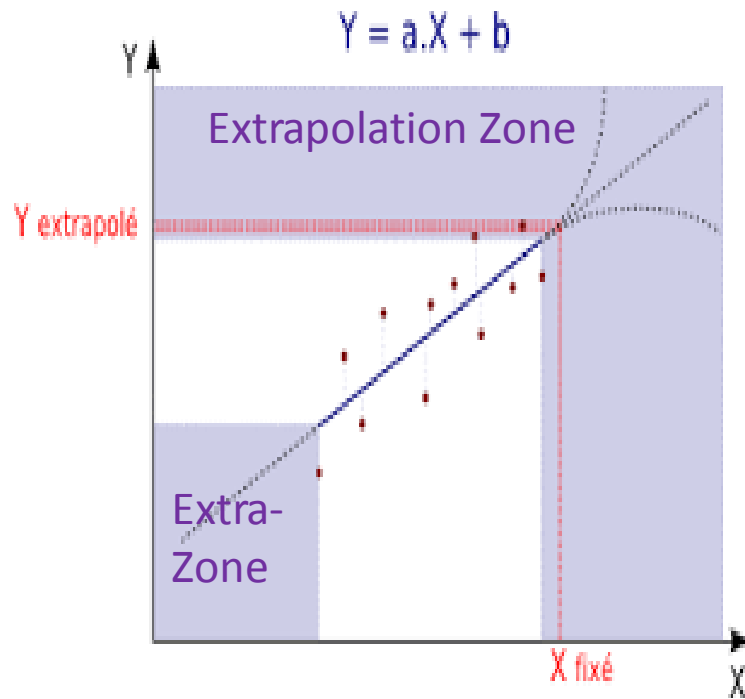
- L'interpolation linéaire est la méthode la plus simple pour estimer la valeur prise par une fonction continue entre deux points déterminés. Elle consiste à utiliser pour cela la fonction affine passant par les deux points déterminés. [Wikipédia](#)



Simple Linear Regression



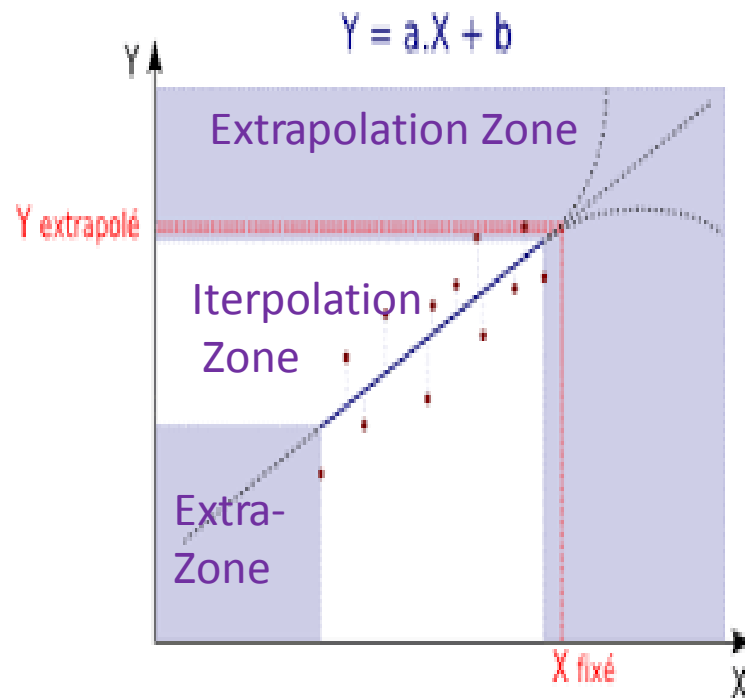
- En mathématiques, l'**extrapolation** est le calcul d'un point d'une courbe dont on ne dispose pas d'équation, à partir d'autres points, lorsque l'abscisse du point à calculer est au-dessus du maximum ou en dessous du minimum des points connus. [Wikipédia](#)



Simple Linear Regression



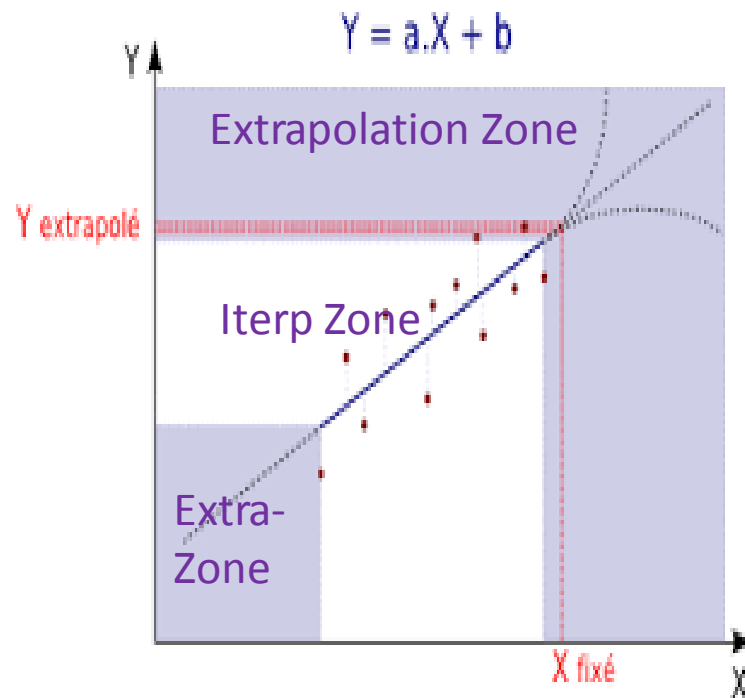
- **Interpolation:** faire une prediction dans la zone des valeurs connues.
- **Extrapolation:** faire une prediction en dehors de la zone de données connues.



Simple Linear Regression

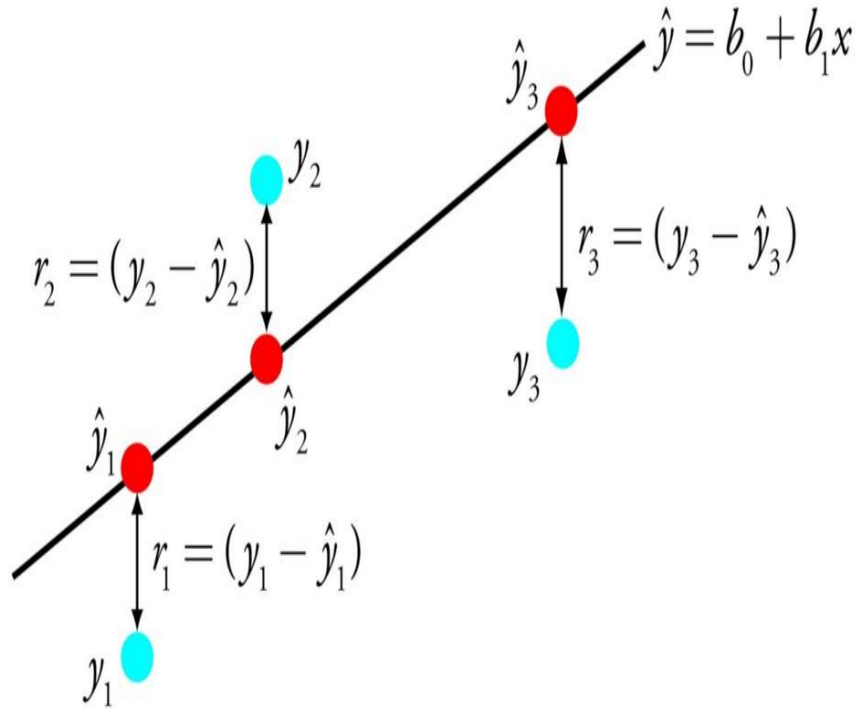


- Et donc, l'interpolation & l'extrapolation définissent la généralisation d'un modèle de régression linéaire.

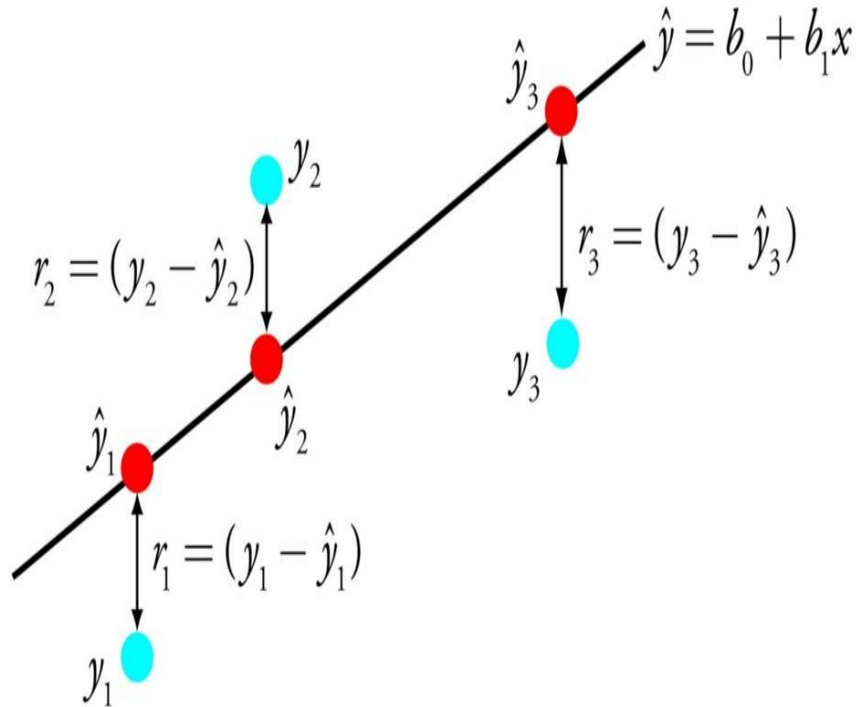


Cost & Optimization

- **Erreur**: la différence entre un point (donnée connue) est sa projection sur la droite.
- Quand on considère le carré de cette de l'erreur, on parle de **Erreur quadratique**.



- L'optimisation consiste à trouver une fonction qui puisse minimizer le cost moyen. [l'algorithme de moindres carrés est beaucoup plus utilisé]
- $\text{Cost} = \sum_{k=0}^n (y - f(x))^2$
 - $\text{Cost} = r_1 + r_2 + \dots + r_n$
- $\text{CostMoyen} = \text{cost} / n$



Application

Salary prediction with Linear Regression

- Les demos se trouvent dans le fichier jupyter notebook

Microsoft Student Partners