Machine learnging :

**Steps Apprentissage supervisé :**

* Collecte data: Même nature
* Nettoyage des données
* Normalisation
* Affecter une Valeur vectorial aux sorties

**Modèle** :

* Fonction de coût quantifie l’erreur entre la sortie et l’entré. Il commence par donner des valeurs aléatoire à teta..
* Ensuite à partir de l’erreur un algorithme d’optimisation va essayer de proposer des nouvelles valeurs de teta..

🡪 Objectif c’est de trouver un modèle parfait avec des paramètre optimal

🡪 Structure générale

La fonction de coût change en fonction du problème ou type d’apprentissage.

Ex : gradient descent, parmis les plus utilisés, pour tous les types d’apprentissages. Le plus important c’est la fonction de coût qui dépend du dataset.

**Test des données :**

Matrice de confusion 🡪 Classification

(FP : false positive..)

* Accuracy = TN+TP/total samples
* Precision = TP/TP+FP
* Recall = TP/TP+FN
* F1 score = 2\*recall\*precision / recall + precision

In case of multi classification (moto, car, airplane) 🡪accuracy calcul different.

R2 score 🡪 Regression

[**Underfitting vs overfitting:**](https://www.google.com/search?client=safari&sca_esv=0946cc8a818a9103&sca_upv=1&rls=en&sxsrf=ADLYWIJjwCuXT3tpdwFlG_o8RxUXUlqqkg:1715931631667&q=Underfitting+vs+overfitting:&tbm=isch&source=lnms&prmd=ivnbz&sa=X&ved=2ahUKEwi72PS6l5SGAxXQUaQEHbmdBLwQ0pQJegQIBRAB&biw=1440&bih=820&dpr=2#imgrc=dEH0tdHT0uWCwM&imgdii=SJHJXJVfZz9yrM)



Methods pour y pallier :

* Régularisation (overfitting): Prendre en compte les paramètre dans la function de coût (Teta) 🡪 penalization de la fonction de coût

🡪Regularisation L2 (ridge regression)

Elle donne le même poids à tous les paramètre du modèle (teta : 0.12 ; 0.1)

🡪L1 (Lasso régularisation)

Elle sélectionne les paramètres du modèle. (teta : 0.8 ; 0.01)

3ème méthodes : combinaison de L1/L2

ANN : artificiel neural network

Y = phi(somme wi xi +b)

Ou

Y = phi(Z)

🡪 Neurone

Fonction MSE = fonction de cout utilisé dans la régression

Fonction de coût cross entropie = basé sur le logarithme de la vraisemblance