SCRIPT TIPE

**Slide1**

Bonjour…

**Slide2**

A travers ce travail de tipe nous allons essayer de répondre à ces différentes questions :

* …

**Slide3**

Contexte : fil rouge à voir…

**Slide4**

**Slide5**

On a choisit de travailler sur un dataset comportant la consommation électrique aux états unis :

**Slide6**

Graphiquement le set de données ressemble à ça : consommation en MegaWatt selon le temps

On remarque graphiquement que on a une série temporelle avec des motifs qui se répètent

**Slide7**

Pour pouvoir réaliser des modèles de machine learning la première étape concernant le dataset est de séparer les données en deux :

* La partie train qui va permettre d’entrainer notre modèle sur nos données
* La partie test qui va être là pour confronter nos prédictions aux données réelles

On a choisi de séparer notre jeu de données en prenant 80% des données pour la partie entrainement et 20% pour la partie test

Mais comment évaluer nos prédictions ?

**Slide8**

Eh bien en utilisant une erreur qui se comprend bien graphiquement :

* En vert vous avez les données réelles
* En bleu les prédictions du modèle
* L’erreur va être la distance entre ces deux courbes (ici en jaune)

Plusieurs types d’erreurs existent, nous avons choisi de façon complétement arbitraire l’erreur RMSE

Le choix est peu important ce qui compte en revanche c’est de conserver la même erreur tout au long des différents modèles

**Slide9**

1er modèle pour prévoir la consommation électrique :

On va travailler sur deux dataset extrait de notre dataset initial : on compresse notre dataset en faisant la somme par jour et par semaine

**Slide10**

Ce premier modèle simples est purement statistique/mathématique : cela consiste à faire la moyenne des années précédentes

**Slide11**

Graphiquement on obtient les résultats suivants

L’erreur du modèle est de 279583 MW sur en moyenne une consommation moyenne par semaine de 2 560 000

On a donc une perte de 279 583 mW à stocker ce qui est difficile $

De plus on remarque une surestimation de notre modèle : on explique cela par la decroissance de la consommation au cours des 12 dernières années aux Etats Unis -> notre modèle a donc appris sur des données plus élevés : C’est donc cohérent qu’on ait des surestimation

**Slide 12**

On réalise le même calcul pour la consommation par jour et on obtient une erreur de 12%

Consommation moyenne d’un ménage aux états unis : 10 766 kWh soit 10,7 MWh

Soit en une journée 256,8 MW en moyenne

Or 54819/256.8=180 -> 180 ménages sont privées d’électricité (si le modèle se trompe dans le négatif)

Transition :

**Slide13**

Maintenant que nous avons fait un modèle simple statistique passons à des modèles de machine learning

On commence par un modèle linéaire en utilisant la bibliothèque sklearn sur python

// à préciser et faire une meilleurs régréssion

**Slide14**

//mettre les nouveaux résulats

On a une erreur de \_\_\_\_\_

On en conclue que le machine learning ne réalise pas toujours des prédictions meilleures qu’avec un modèle statistique

**Slide15**

Vous l’aurez remarqué la régréssion linéaire n’est pas du tout adapté à nos données pour palier à cela on utilise une régréssion polynomiale qui va épouser la forme de la courbe de nos données afin d’améliorer nos prédictions (cf ma schéma)

**Slide16**

Pour la régréssion polynomiale on augmente le degré progressivement

Pour le degré 5 on obtient les meilleurs résultats

On obtient graphiquement :…

**Slide17**

Lorsqu’on pousse la régression jusqu’au degré 6 on remarque que l’erreur raugmente, le modèle sur apprend sur les particularités des données et devient donc moins précis

Fil rouge

De nos jours, on vit une transition énergétique : les énergies renouvelables commencent à prendre une part assez grande dans la production énergétique.

Le solaire ou l’éolien en sont deux mais comme vous pouvez le voir le solaire ne produit que la journée

Or d’ici 2050 le pic de consommation énergétique en France sera la nuit

On a donc besoin de stocker cette électricité,

Or le stockage de l’électricité est difficile et engendre des pertes.

Il faut donc limiter le plus possible le stockage. Une façon de le limiter est d’anticiper la consommation électrique. C’est donc ce que l’on va essayer de faire ici.

Consommation moyenne d’un menage aux etats unis : 10 766 kwh soit 10,7 MWh

Soit en une journée 256,8 MW

Or 54819/256.8=180 -> 180 ménages sont privées d’électricité (si le modèle se trompe dans le négatif)