

TP - Apprentissage machine sur des données d'usinage

Le livrable final devra contenir un rapport expliquant votre méthodologie, ainsi que le code correspondant à chaque partie. Vos propos doivent être illustrés par des graphiques.

Partie 0 : Configuration du logiciel *Geany* pour *Python* et installation des packages nécessaires.

Le logiciel *Geany* sera utilisé lors des séances de travaux pratiques pour les développements *Python*. Avant de commencer il faut configurer *Geany* (ou vérifier qu'il est correctement configuré) pour le langage *Python* en effectuant les étapes suivantes :

1. Ouvrir *Geany* et créer un fichier principal de code *Python*, en s'assurant que l'extension du fichier est .py
2. Ouvrir la fenêtre de configuration dans le menu "Construire" → "Définir les commandes de construction"
3. Remplacer la commande d'exécution par `C:\ProgramData\Miniconda3\python.exe "%f"`
4. Tester la configuration avec un code simple : `print("test")`, l'exécution se lance avec la touche F5 du clavier

Plusieurs packages *Python* additionnels sont nécessaires pour effectuer ces travaux pratiques (numpy, matplotlib, scikit-learn ...), pour les installer il faut rechercher dans la barre de menu Windows "miniconda" et exécuter le programme "Anaconda Powershell Prompt (Miniconda)", et lancer la commande suivante dans ce programme : `pip install --user numpy matplotlib scikit-learn scipy`

Partie 1 : La base de données.

4225

La base de données a été constituée lors d'une campagne d'acquisition sur un usinage de 169 trous effectués avec un foret de 6.36mm (soit 6360 microns). Pour chaque trou il y a 25 mesures sur le perçage (9 de diamètres, 14 de rugosité, et 2 de position) et 13 lors du processus (6 d'effort de coupe, 6 de moment de coupe, et une mesure énergétique).

Mesures de diamètre :	Mesures de rugosité :
D_In : Diamètre moyen d'entrée du trou	Rz_In_max : Rugosité maximale Rz à l'entrée
D_In_max : Diamètre maximal d'entrée du trou	Rz_In : Rugosité moyenne Rz à l'entrée
D_In_min : Diamètre minimal d'entrée du trou	Rz_Out_max : Rugosité maximale Rz à la sortie
D_Out : Diamètre moyen de sortie du trou	Rz_Cly_max : Rugosité maximal Rz sur la paroi du trou
D_Out_max : Diamètre moyen de sortie du trou	Rz_Cyl : Rugosité moyenne Rz sur la paroi du trou
D_Out_min : Diamètre minimal de sortie du trou	Rt_In : Rugosité moyenne Rt à l'entrée
D_Cyl : Diamètre moyen sur la paroi du trou	Rt_Out : Rugosité moyenne Rt à la sortie
D_Cyl_max : Diamètre maximal sur la paroi du trou	Rt_Cyl : Rugosité moyenne Rt sur la paroi du trou
D_Cyl_min : Diamètre minimal sur la paroi du trou	Rq_In : Rugosité moyenne Rq à l'entrée
	Rq_Out : Rugosité moyenne Rq à la sortie
Mesures de position :	Rq_Cyl : Rugosité moyenne Rq sur la paroi du trou
P(x) : Position en x du trou	Ra_In : Rugosité moyenne Ra à l'entrée
P(y) : Position en y du trou	Ra_Out : Rugosité moyenne Ra à la sortie
Mesures de énergétique :	Ra_Cyl : Rugosité moyenne Ra sur la paroi du trou
Et_Consumee : Energie totale consommée	

Mesures des efforts coupe :		Mesures des moments coupe :	
Moy_F_T1&T2	: Effort moyen de coupe entre T1 et T2	Moy_M_T1&T2	: Moment moyen de coupe entre T1 et T2
Moy_F_T2&T3	: Effort moyen de coupe entre T2 et T3	Moy_M_T2&T3	: Moment moyen de coupe entre T2 et T3
Moy_F_T3&T4	: Effort moyen de coupe entre T3 et T4	Moy_M_T3&T4	: Moment moyen de coupe entre T3 et T4
Max_F_T1&T2	: Effort maximal de coupe entre T1 et T2	Max_M_T1&T2	: Moment maximal de coupe entre T1 et T2
Max_F_T2&T3	: Effort maximal de coupe entre T2 et T3	Max_M_T2&T3	: Moment maximal de coupe entre T2 et T3
Max_F_T3&T4	: Effort maximal de coupe entre T3 et T4	Max_M_T3&T4	: Moment maximal de coupe entre T3 et T4

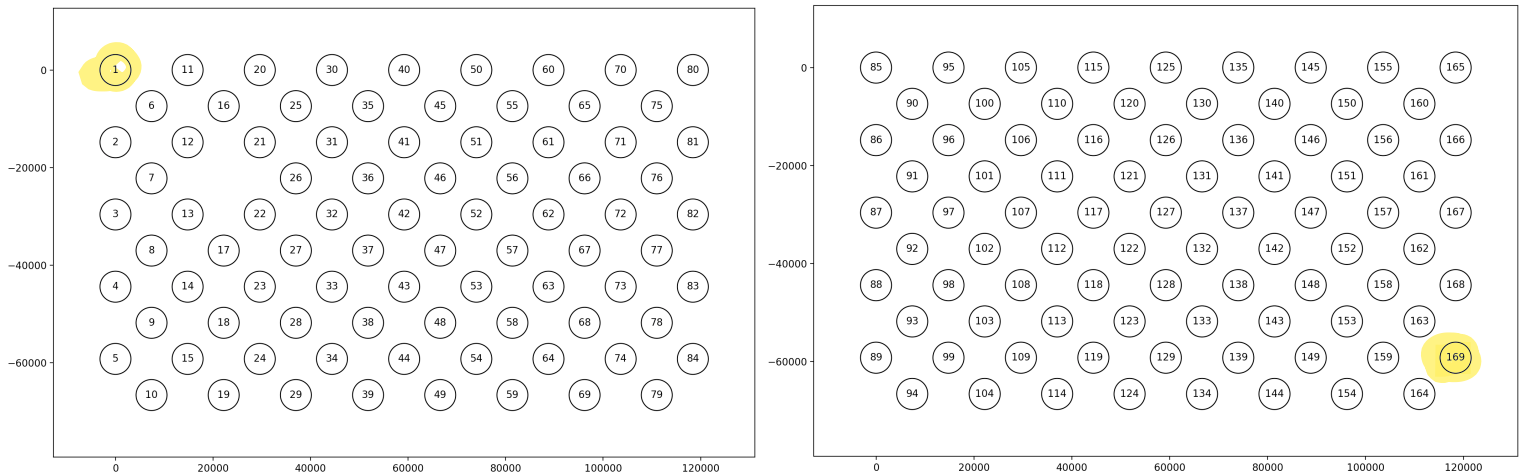


FIGURE 1 – Positions des trous (**P(x)** et **P(y)**), sur les deux plaques.

Les données sont dans le fichier `data_capteurs.py` disponible sur SAVOIR, pour y accéder il suffit de l'importer avec `import data_capteurs`, puis d'accéder aux données avec la structure `data_capteurs.data`, voici plusieurs exemples d'utilisation de cette structure :

- `data_capteurs.data["D_In"][0]` retourne le diamètre moyen d'entrée du premier trou (d'indice 0).
- `data_capteurs.data["Et_Consummee"]` retourne la liste de toutes les valeurs de l'énergie consommée.
- `data_capteurs.data["Moy_F_T1&T2"][41]` retourne l'effort moyen entre T1 et T2 pour le trou d'indice 41.
- `[data_capteurs.data["P(x)"][0], data_capteurs.data["P(y)"][0]]` : position du premier trou
- etc ...

Un exemple d'affichage d'une courbe des données est également donné sur SAVOIR (`exemple.py`), il vous permettra d'avoir un point de départ pour découvrir la base.

Une première étape indispensable va être de **se familiariser avec la base de données**, en la manipulant et en comprenant quel est son contenu.

Effectuez **un affichage des différentes données** sous forme de **courbes**, ou tout **autre graphique** qui vous semblerait pertinent pour **mieux comprendre ce jeu de données**.

Partie 2 : Développement de différentes problématiques.

Cette base de données riche offre de nombreuses perspectives d'exploration de l'analyse de données et de l'apprentissage machine. Pour chacun de ces points développez la problématique en vous appuyant sur des expériences, en illustrant vos propos par des graphiques, des calculs, des résultats d'apprentissage ou d'analyses issues d'autres méthodes. Le volume de donnée étant faible, il sera nécessaire d'utiliser une méthode de validation adaptée : la validation croisée, qui permet d'avoir des score plus pertinent qu'une validation classique lorsque l'on a peu de données (à découvrir/rechercher en autonomie).

1. Lors d'une campagne de perçage d'un grand nombre de trous, le foret va s'user petit à petit, avec une erreur sur le diamètre du trou potentiellement de plus en plus grande. Le premier objectif que l'on se propose est de prédire l'évolution de l'erreur de la taille du trou, connaissant le nombre de trous déjà effectués par le foret. Tentez de déterminer des valeurs dans la plage de l'expérimentation, puis des valeurs en dehors de la plage d'expérimentation, que remarquez-vous ?
2. La consommation d'énergie et les informations d'effort / moment de coupe sont fortement liées, peut-on prédire l'un à partir de l'autre ? Si l'on considère qu'un effort plus grand est nécessaire si l'usure du foret est plus importante, en prenant la chute de diamètre de perçage comme une mesure de l'usure du foret, peut-on estimer cette usure en s'appuyant sur la consommation d'énergie ?
prédire la ct E en se basant sur M et F
Prédire Diamètre de perçage / ct E
3. On peut supposer que la rugosité de la zone d'usinage augmente avec l'usure de l'outil, car la coupe est moins précise, est-il envisageable de prédire la rugosité en connaissant le nombre de coupes déjà effectuées ? En rajoutant d'autres informations ?
4. L'utilisation de N données précédentes pour prédire la suivante est une piste envisageable sur différentes valeurs, cette valeur de N correspond au nombre de trous utilisés pour prédire une valeur suivante, l'objectif est donc d'expérimenter plusieurs valeurs de N sur plusieurs caractéristiques, individuellement ou sur un ensemble.
5. La position du trou dans la plaque a-t-elle une influence dans les données ? Comment pourrait-on vérifier cette information ?
corrélation
6. Bonus : à vous de proposer une étude intéressante sur ces données !