

## LISTE DE COMMANDES

Depuis le terminal, vous pouvez lancer une ligne de commande en spécifiant les arguments suivants :

Argument	Options (résultat souhaité)	Description
--data		Chemin vers le fichier CSV contenant les données (obligatoire)
--colonne_cible		Nom de la colonne cible à analyser pour la classification (obligatoire)
--k_max		Valeur maximale de k à tester pour les courbes d'erreur (optionnel, défaut : 10)
--k		Valeur spécifique de k à tester pour le taux d'erreur (optionnel, défaut : 3)
--n_repetitions		Nombre de répétitions pour obtenir des résultats fiables (optionnel, défaut : 10)
--test_size		Fraction des données réservée au test (optionnel, défaut : 0.3)
plot_KNN_error_rate		Action à effectuer (obligatoire), parmi
--action		Tracer les courbes de taux d'erreur en fonction des valeurs de k.
--action	charger_et_preparer_donnees	Charger et préparer un dataframe (but : séparer la variable dépendante des variables explicatives.
	KNN_error_rate	Calculer le taux d'erreur pour une valeur spécifique de k.
	plot_KNN_error_rate	Tracer les courbes de taux d'erreur en fonction des valeurs de k.
	k_rate_intuitive	Calculer étape par étape le taux d'erreur de manière intuitive.
	meilleur_k	Identifier la valeur optimale de k en minimisant le taux d'erreur.
	afficher_voisins	Afficher un graphique montrant un patient avec ses k plus proches voisins ainsi que les points d'entraînement.

Les arguments nécessaires pour chaque fonction (résultat escompté) qu'on veut exécuter :

Fonction	Arguments nécessaires
charger_et_preparer_donnees	--data --colonne_cible
KNN_error_rate	--data --colonne_cible --k --test_size
plot_KNN_error_rate	--data --colonne_cible --k_max --n_repetitions --test_size
k_rate_intuitive	--data --colonne_cible --test_size
meilleur_k	--data --colonne_cible --k_max --n_repetitions --test_size
afficher_voisins	--data --colonne_cible --k_max --n_repetitions --test_size

Exemple de lignes de commande :

Fonction	Exemple de ligne de commande
charger_et_preparer_donnees	python kcurve.py --data ../data/cancer.csv --colonne_cible <b>diagnosis</b> --action <b>charger_et_preparer_donnees</b>
KNN_error_rate	python kcurve.py --data ../data/cancer.csv --colonne_cible <b>diagnosis</b> --k <b>5</b> --test_size <b>0.3</b> --action <b>KNN_error_rate</b>
plot_KNN_error_rate	python kcurve.py --data ../data/cancer.csv --colonne_cible <b>diagnosis</b> --k_max <b>25</b> --n_repetitions <b>100</b> --test_size <b>0.3</b> --action <b>plot_KNN_error_rate</b>
k_rate_intuitive	python kcurve.py --data ../data/cancer.csv --colonne_cible <b>diagnosis</b> --test_size <b>0.3</b> --action <b>k_rate_intuitive</b>
meilleur_k	python kcurve.py --data ../data/cancer.csv --colonne_cible <b>diagnosis</b> --k_max <b>25</b> --n_repetitions <b>100</b> --test_size <b>0.3</b> --action <b>meilleur_k</b>
afficher_voisins	python kcurve.py --data ../data/cancer.csv --colonne_cible <b>diagnosis</b> --k_max <b>25</b> --n_repetitions <b>100</b> --test_size <b>0.3</b> --action <b>afficher_voisins</b>

**Remarques :**

- Veuillez diriger le terminal dans le bon répertoire avant d'exécuter les lignes de commande
- Le chemin du fichier contenant les données doit être bien spécifié
- Les parties des lignes de commandes en rouges sont celles qu'on peut varier selon les besoins