

Rapport sur le Processus MLOps pour la Détection des Débuts de Crises Épileptiques

Explication du Pipeline de Modèles

Ce rapport détaille le pipeline de modèles conçu spécifiquement pour la classification des données dans le cadre de la détection des débuts de crises épileptiques. Il met l'accent sur la gestion des déséquilibres de classes et l'intégration de divers modèles et ensembles d'hyperparamètres. L'ensemble du processus est suivi et enregistré à l'aide de MLflow, facilitant l'organisation, l'optimisation et la reproductibilité des expériences.

Suivi des Expériences

.

Que voulez-vous suivre pour chaque Essai d'Expérience ?

- Les données d'entraînement et de validation utilisées.
- Les hyperparamètres.
- Les métriques.
- Les modèles.

Pourquoi suivre ?

- Organisation.
- Optimisation.
- Reproductibilité.

Outil - MLFlow

MLFlow vous aide à organiser vos expériences en essais.

MLFlow conserve le suivi de :

- Balises.
- Paramètres.
- Métriques.
- Modèles.
- Artifacts.

Interface MLFlow pour le Suivi des Expériences

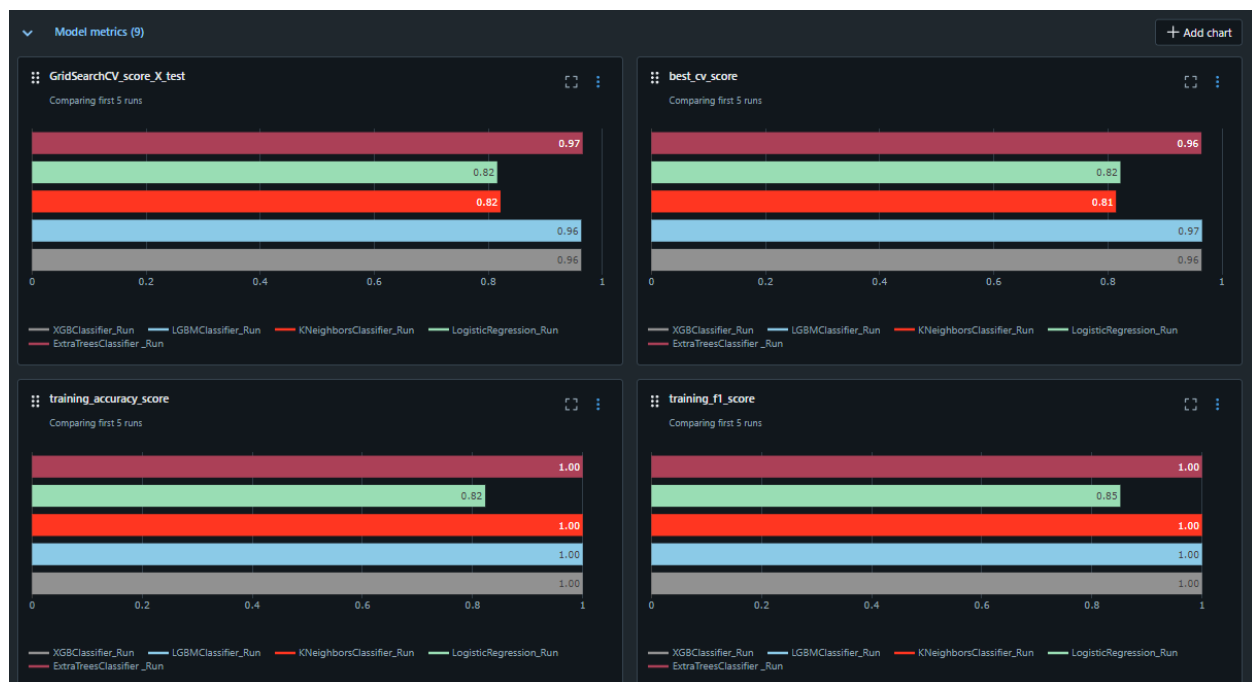
Dans cette section, nous utiliserons l'interface MLFlow pour suivre nos expériences, enregistrer les métadonnées et les modèles associés à chaque essai.

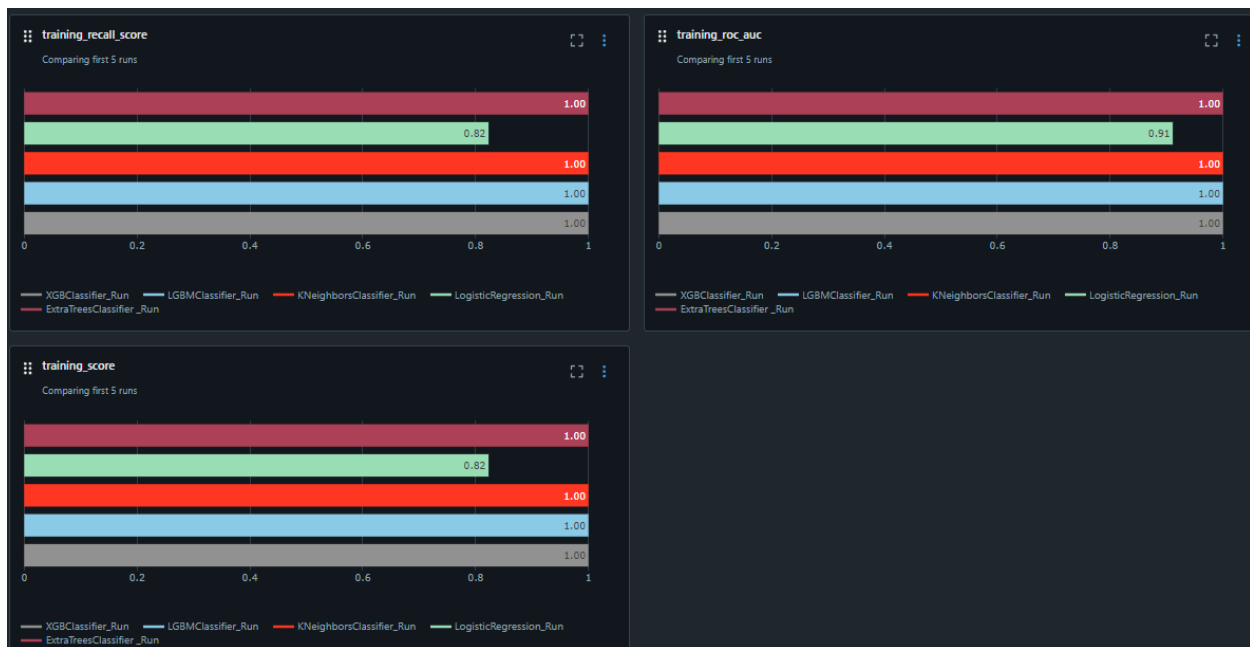
Run Name	Created	Duration	Source	Models	Metrics
XGClassifier_Run	3 hours ago	2.9min	C:\Users\...	Epileptic_.../1, 1 more	0.963011031...
LGBMClassifier_Run	3 hours ago	3.0min	C:\Users\...	Epileptic_.../2, 1 more	0.963011031...
KNeighborsClassifier_...	3 hours ago	38.4s	C:\Users\...	Epileptic_.../3, 1 more	0.821544451...
LogisticRegression_R...	3 hours ago	17.3s	C:\Users\...	Epileptic_.../4, 1 more	0.815704088...
ExtraTreesClassifier_...	3 hours ago	1.6min	C:\Users\...	Epileptic_.../5, 1 more	0.965606748...

Table	Chart	Evaluation	Experimental							
									Metrics	
		Run Name	Created	Dataset	Duration	Source	Models	GridSearchCV_x	mean_test_score	mean_train_score
		XGBClassifier_Run	3 hours ago	dataset (ed0a33f0) Train	2.9min	C:\Users\...	Epileptic_1, 1 more	0.96301103...	-	-
		LGBMClassifier_Run	3 hours ago	dataset (ed0a33f0) Train	3.0min	C:\Users\...	Epileptic_2, 1 more	0.96301103...	-	-
		KNeighborsClassifier...	3 hours ago	dataset (ed0a33f0) Train	38.4s	C:\Users\...	Epileptic_3, 1 more	0.82154445...	-	-
		LogisticRegression_R...	3 hours ago	dataset (ed0a33f0) Train	17.3s	C:\Users\...	Epileptic_4, 1 more	0.81570408...	-	-
		ExtraTreesClassifier_...	3 hours ago	dataset (ed0a33f0) Train	1.6min	C:\Users\...	Epileptic_5, 1 more	0.96560674...	-	-
		rumbling-stoat-...	3 hours ago	-	1.6min	C:\Users\...	-	-	0.90605165...	0.92223744...
		defiant-mouse-7...	3 hours ago	-	1.6min	C:\Users\...	-	-	0.88155130...	0.89400439...
		defiant-crab-196	3 hours ago	-	1.6min	C:\Users\...	-	-	0.89761496...	0.91891129...
		nosy-eel-198	3 hours ago	-	1.6min	C:\Users\...	-	-	0.88430959...	0.89769562...
		clumsy-mink-975	3 hours ago	-	1.6min	C:\Users\...	-	-	0.94969994...	0.99987829...
		victorious-shrew...	3 hours ago	-	1.6min	C:\Users\...	-	-	0.95456744...	1
		orderly-gnu-368	3 hours ago	-	1.6min	C:\Users\...	-	-	0.96349153...	1
		overjoyed-jay-203	3 hours ago	-	1.6min	C:\Users\...	-	-	0.89858872...	0.91181229...
		abrasive-conch-...	3 hours ago	-	1.6min	C:\Users\...	-	-	0.88220065...	0.89489636...
		treasured-crab-2...	3 hours ago	-	1.6min	C:\Users\...	-	-	0.96397802...	1
		suave-snake-559	3 hours ago	-	1.6min	C:\Users\...	-	-	0.90118678...	0.92122275...
		salty-squid-810	3 hours ago	-	1.6min	C:\Users\...	-	-	0.96235583...	1
		tasteful-dog-504	3 hours ago	-	1.6min	C:\Users\...	-	-	0.96316659...	1
		traveling-elk-178	3 hours ago	-	1.6min	C:\Users\...	-	-	0.90377997...	0.92252157...

Affichage des Métriques

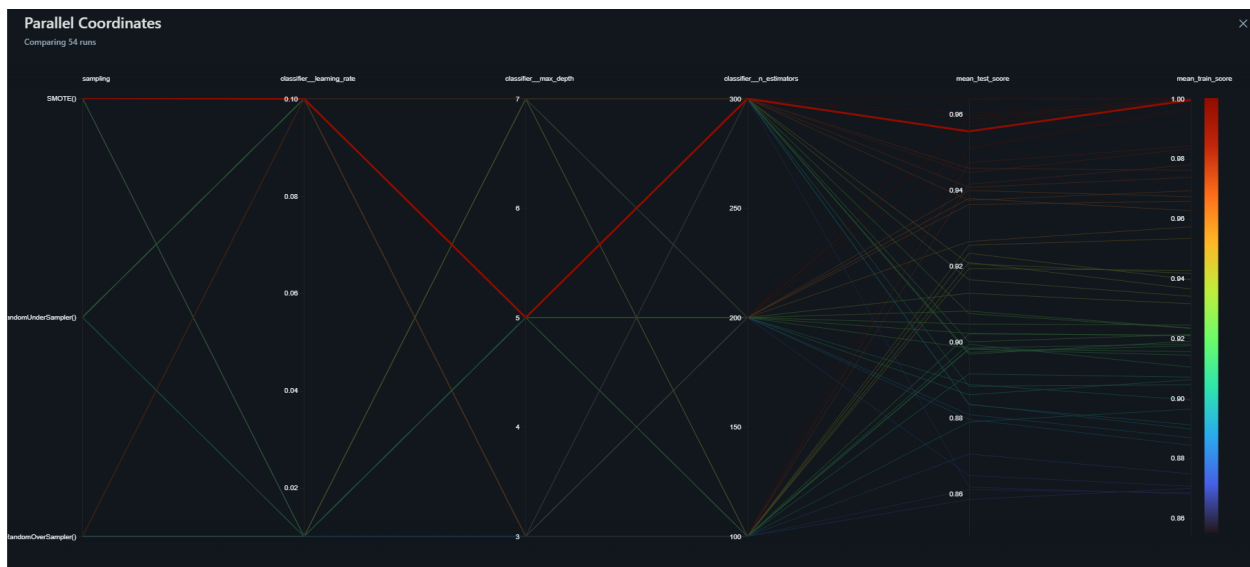
Des graphiques de métriques sont générés pour visualiser les performances des modèles, permettant une analyse détaillée et une comparaison entre différents essais



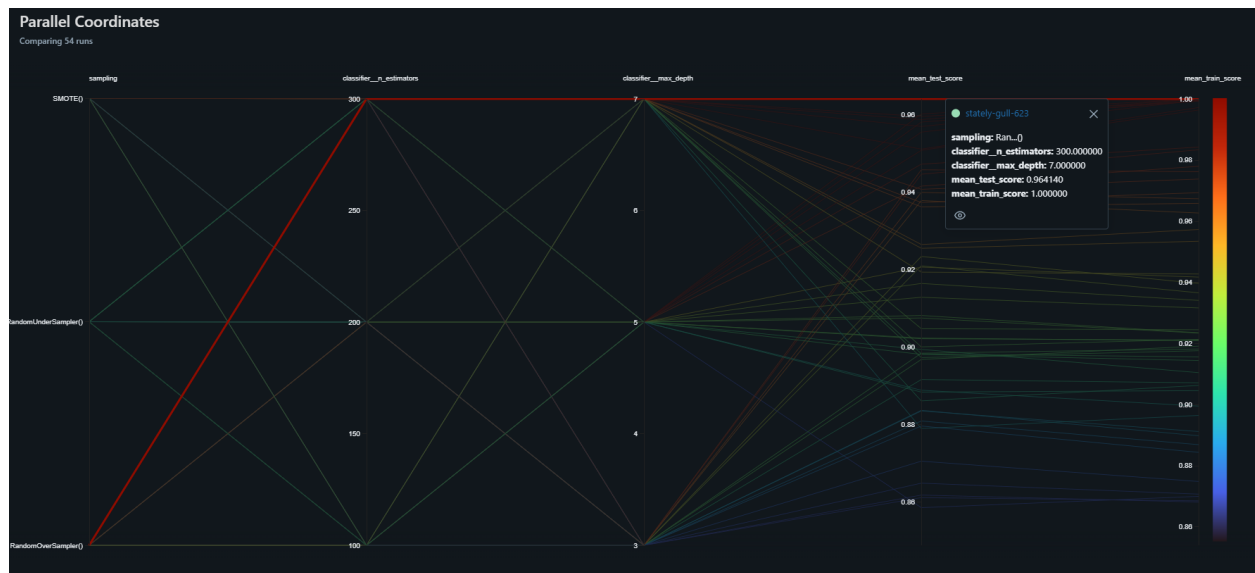


Parallel Coordinates des Modèles

Les graphiques de Parallel Coordinates pour XGBoost et ExtraTrees offrent une vue d'ensemble des performances et des configurations d'hyperparamètres, aidant à identifier les configurations les plus performantes.

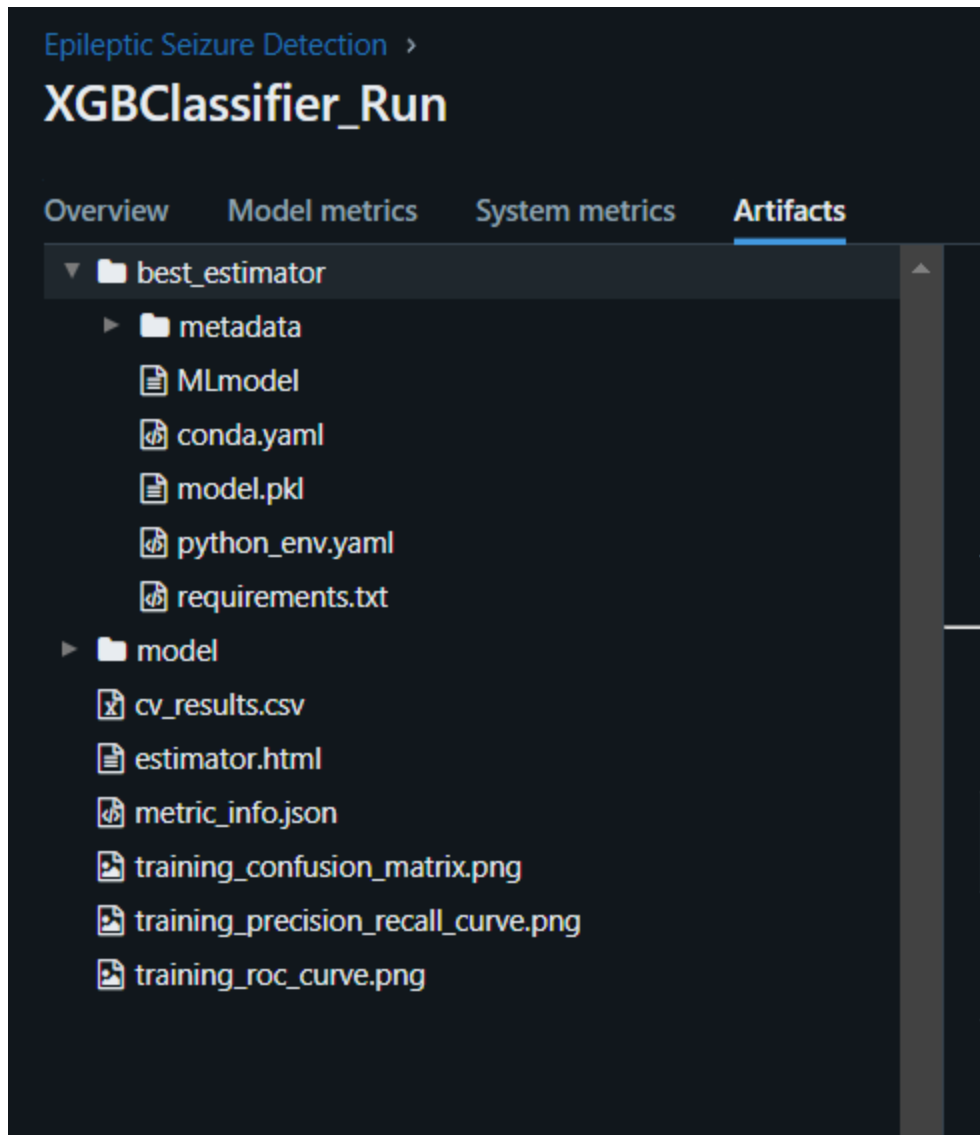


ExtraTrees :



Exemple de Model Artifacts

Les artifacts de modèles, comme ceux générés pour XGBoost, fournissent des informations supplémentaires et des fichiers nécessaires pour le déploiement ou l'analyse ultérieure des modèles.



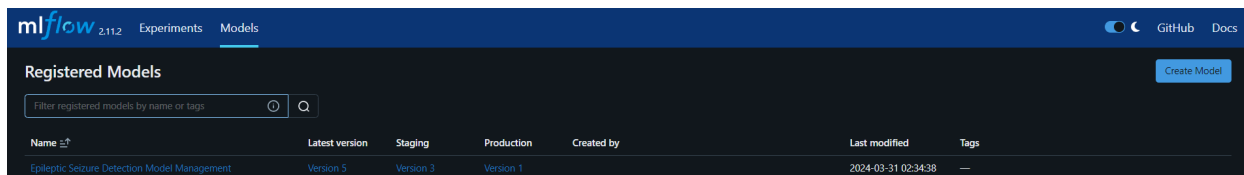
Registre de Modèles

Registre de Modèles

Le registre de modèles est une fonctionnalité centrale de MLFlow, supportant :

1. **Enregistrement de Modèles** : Pour conserver les modèles entraînés.
2. **Versionnage de Modèles** : Pour suivre les différentes versions des modèles.
3. **Transitions de Stades** : Pour gérer les transitions des modèles entre différents stades (Archivé, En Staged, Production).

4. **Collaboration Intra-Équipe** : Facilite la collaboration en permettant à l'équipe de partager et de discuter des modèles.



The screenshot shows the mlflow web interface with the 'Models' tab selected. It displays a table of 'Registered Models'. The table has columns for Name, Latest version, Staging, Production, Created by, Last modified, and Tags. One model is listed: 'Epileptic Seizure Detection Model Management' with versions 5, 3, and 1, last modified on 2024-03-31 at 02:34:38.

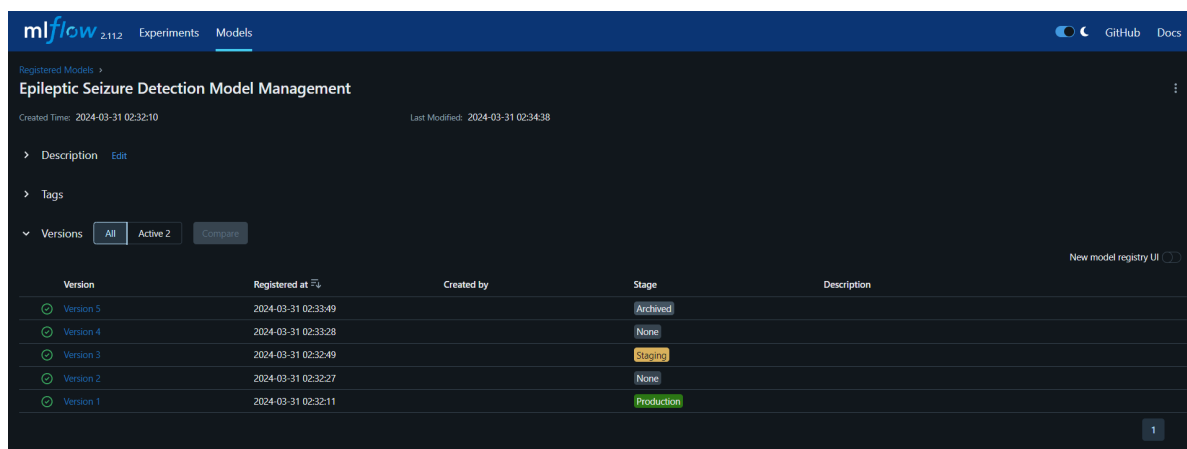
Name	Latest version	Staging	Production	Created by	Last modified	Tags
Epileptic Seizure Detection Model Management	Version 5	Version 3	Version 1		2024-03-31 02:34:38	—

Principales Fonctionnalités :

1. Enregistrement de Modèles
2. Versionnage de Modèles
3. Transitions de Stades
4. Collaboration Intra-Équipe

Versionnage de Modèles

1. **Archivé** : Ces versions ne sont plus en cours d'utilisation.
2. **En Staged** : Ces versions sont prêtes pour le déploiement en attente de validation finale.
3. **Production** : Ces versions sont activement utilisées par les utilisateurs dans un environnement de production en direct.



The screenshot shows the mlflow web interface with the 'Models' tab selected. It displays the details of a specific model, 'Epileptic Seizure Detection Model Management'. The page shows the model's description, tags, and a list of versions. The versions are listed in a table with columns for Version, Registered at, Created by, Stage, and Description. The stages are color-coded: Archived (grey), None (white), Staging (yellow), and Production (green).

Version	Registered at	Created by	Stage	Description
Version 5	2024-03-31 02:33:49		Archived	
Version 4	2024-03-31 02:33:28		None	
Version 3	2024-03-31 02:32:49		Staging	
Version 2	2024-03-31 02:32:27		None	
Version 1	2024-03-31 02:32:11		Production	

Conclusion

Le pipeline MLOps décrit dans ce rapport, appuyé par l'utilisation de MLFlow, offre une structure robuste pour la détection des débuts de crises épileptiques. Il assure une gestion efficace des expériences, une optimisation des modèles, et une collaboration fluide au sein de l'équipe, tout en facilitant la reproductibilité et l'organisation du travail.