



MISE EN PLACE D'UNE PIPELINE MLOPS POUR UN MODÈLE DE PRÉDICTION DE TIR NBA

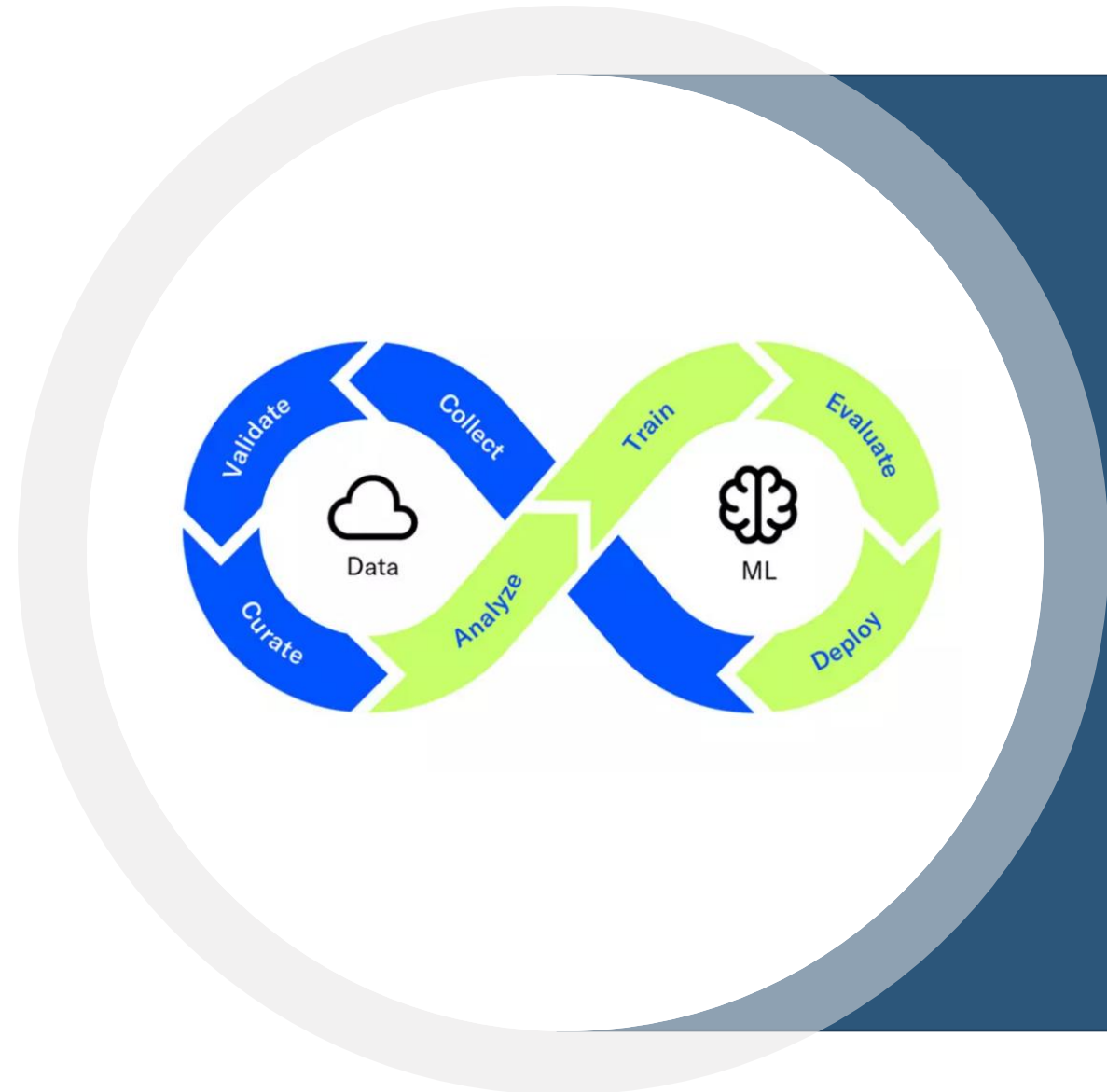
DEC24_MLOPS_PARIS_SPORTIF

P. COHEN, F. HEUZE, S. JIANYING

20/05/2025

SOMMAIRE

- 🏀 Présentation du projet
- 🏀 Exploration des données
- 🏀 Présentation de l'architecture
- 🏀 Description de l'ETL
- 🏀 Description de la pipeline ML
- 🏀 Présentation du Front end
- 🏀 Supervision
- 🏀 Drift monitoring
- 🏀 Demonstration live



PRÉSENTATION DU PROJET

- 🏀 Modèle qui prédit si un joueur NBA réussira un tir spécifique.
- 🏀 Le pipeline de données automatisé pour l'entraînement, le déploiement et le suivi du modèle.
- 🏀 Environnement de travail



LE DATASET



Explication dataset

Fichier csv listant les tirs en NBA entre 2017 et 2020



Variables initiales

Avant preprocessing	
Game ID	Shot Zone Range
Game Event ID	Shot Distance
Player ID	X Location
Player Name	Y Location
Team Name	Shot Made Flag
Period	Game Date
Minutes Remaining	Home Team
Seconds Remaining	Away Team
Action Type	Season Type
Shot Type	
Shot Zone Basic	
Shot Zone Area	



LE DATASET



Explication dataset

Fichier csv listant les tirs en NBA entre 2017 et 2020

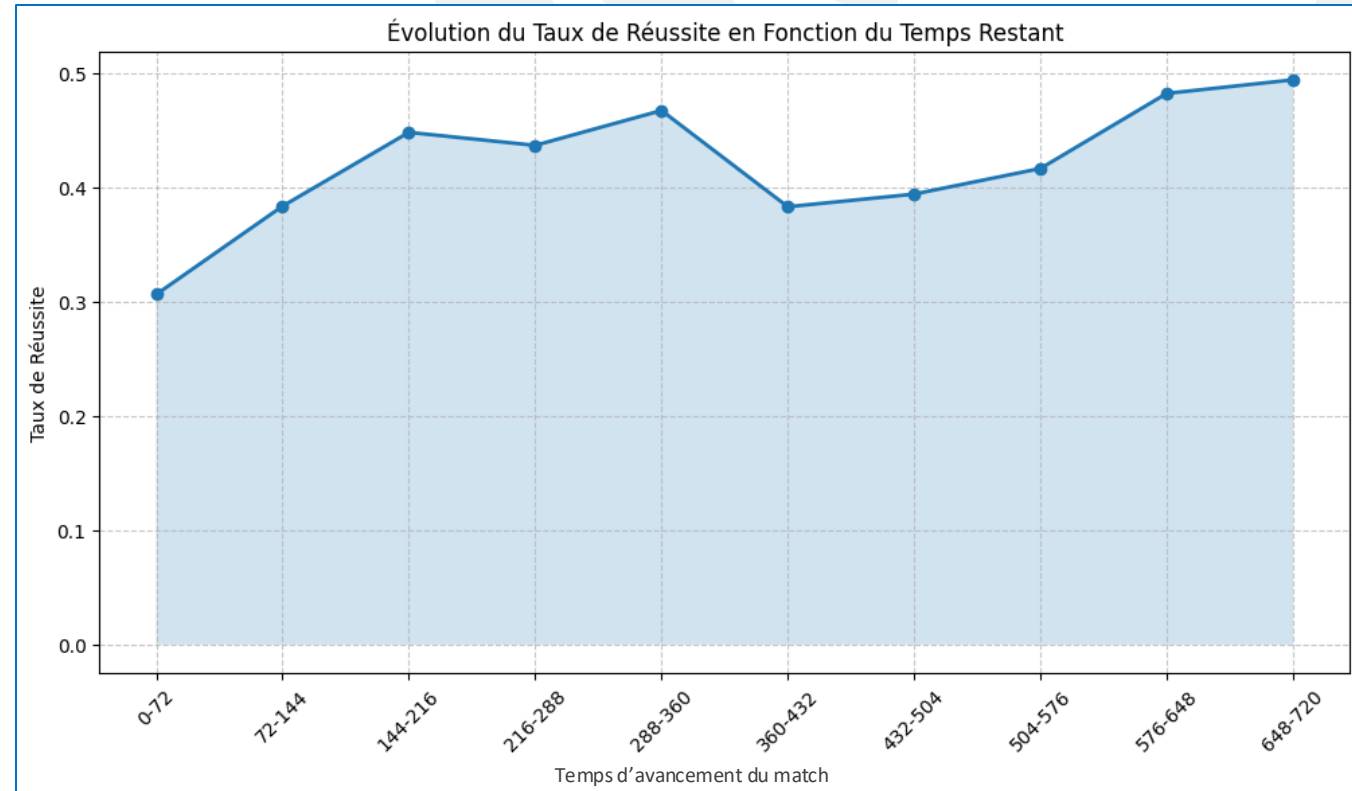
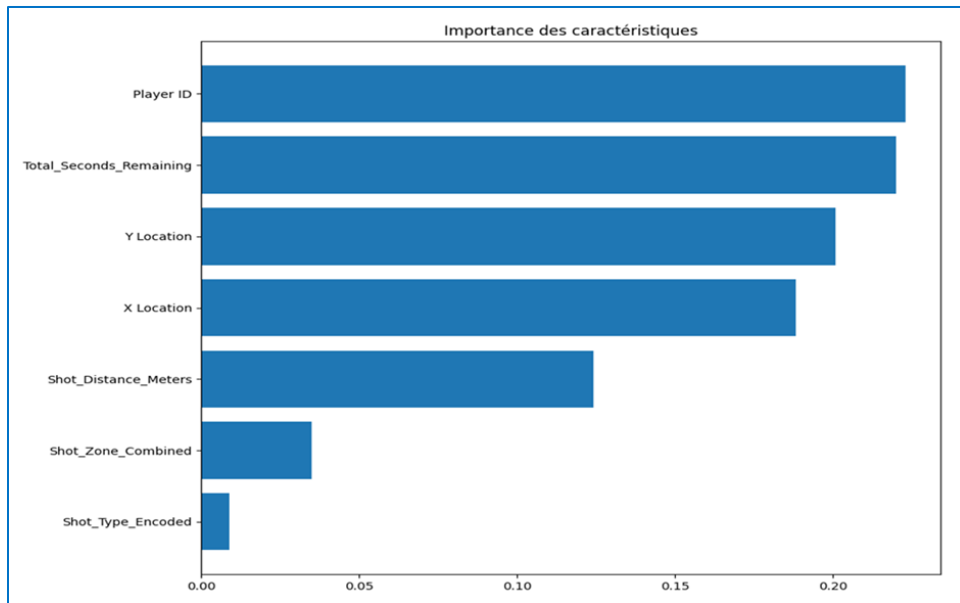


Variables sélectionnées

Colonne	Description
Player_ID	Identifiant du joueur
X_Location	Position X sur le terrain
Y_Location	Position Y sur le terrain
Total_Seconds_Remaining	Temps total restant
Shot_Type_Encoded	Type de tir (2 points ou 3 points)
Shot_Distance_Meters	Distance en mètres
Shot_Zone_Combined	Zone de tir (7 zones)
Shot_Made_Flag	Indicateur de réussite du tir



EXPLORATION DES DONNÉES - EDA



ARCHITECTURE

VM



Une machine virtuelle est utilisée pour la mise en production du projet et dispose des ressources suivantes :

- RAM : 8G
- CPU : 4 (2 Sockets, 2 Cores)
- Stockage : SSD 128G
- OS : Ubuntu Server 24.04
- Reverse-proxy configuré pour un accès public
- Hyperviseur : proxmox
- Backup de bas niveau hebdomadaire

```
DEC24_MLOPS_PARIS_SPORTIFS
├── api
├── app
├── assets
├── docker-compose.yml
├── homer
├── Jenkinsfile
├── LICENSE
├── monitoring
├── notebooks
├── README.md
├── reports
├── requirements.txt
├── src
└── virtualbetsapi
```

parivision.heuzef.com est un portail statique servant de point d'entrée du projet pour lister et accéder rapidement aux différents services.

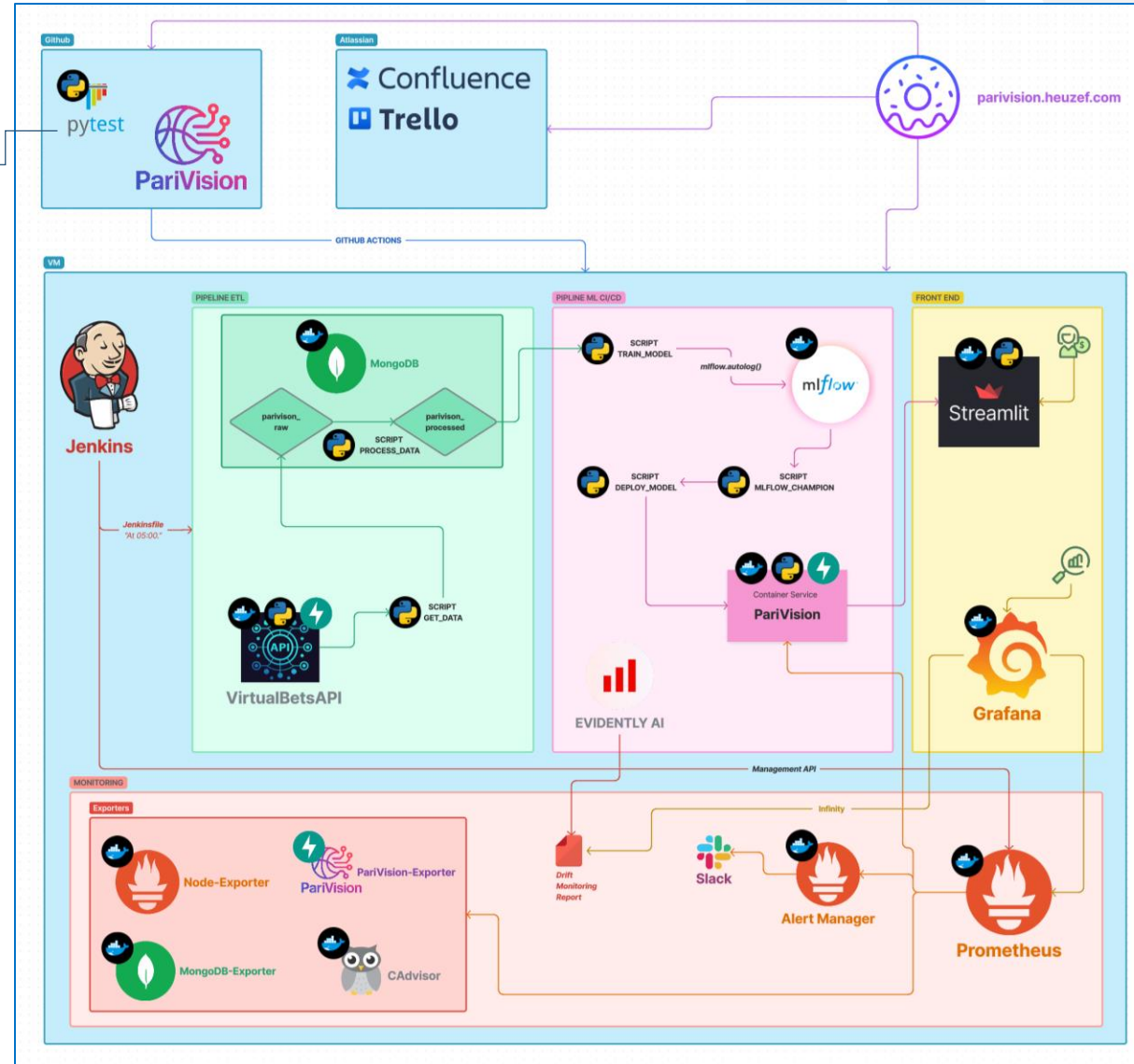
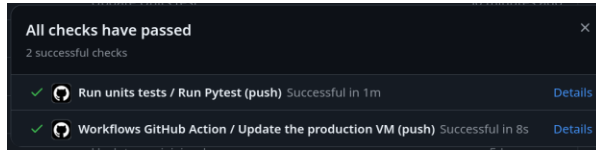
La configuration du Dashboard s'effectue directement sur le dépôt GIT du projet.



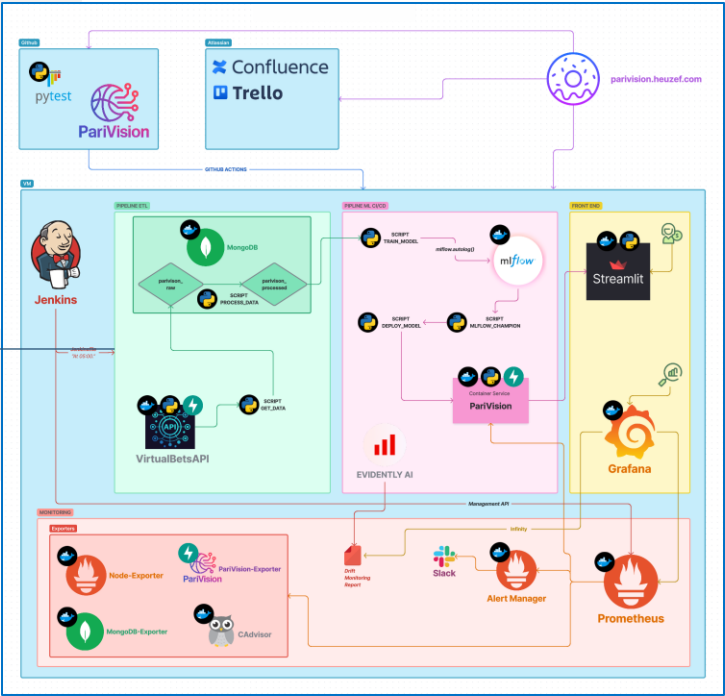
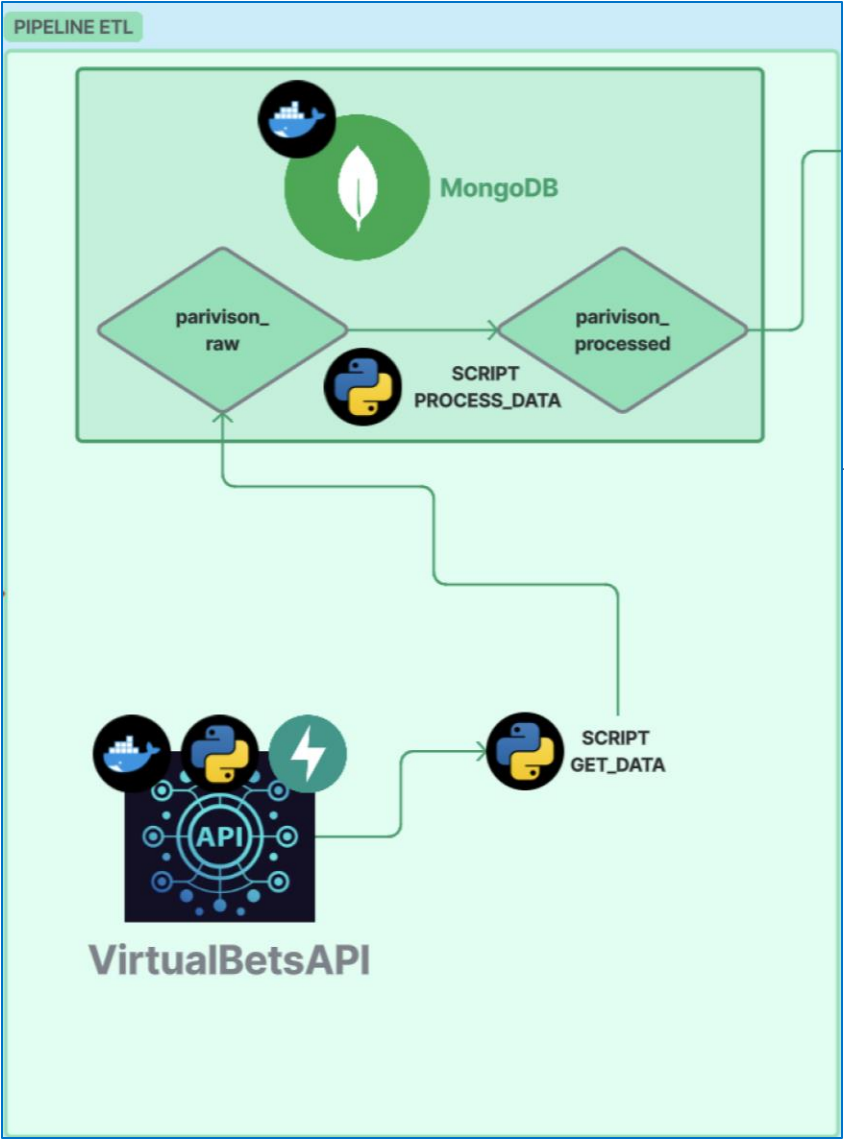
GITHUB

https://github.com/DataScientest-Studio/DEC24_MLOPS_PARIS_SPORTIFS

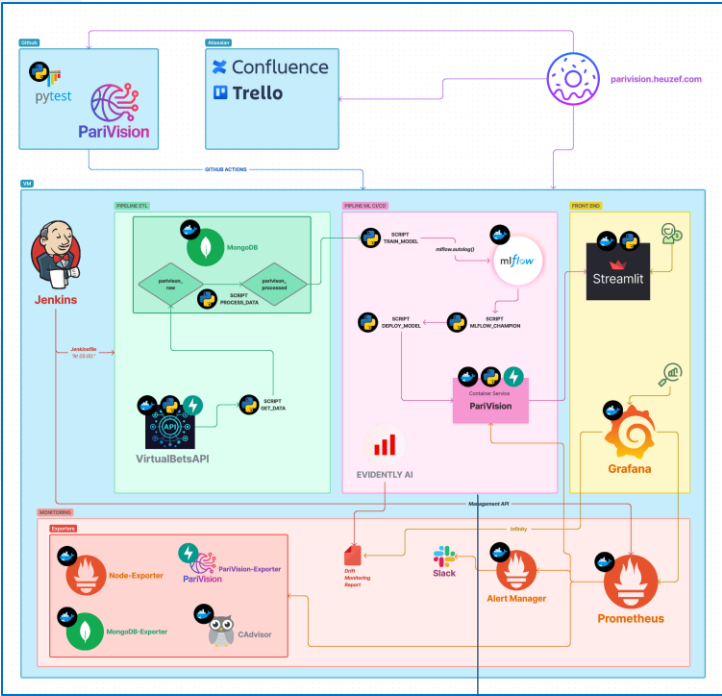
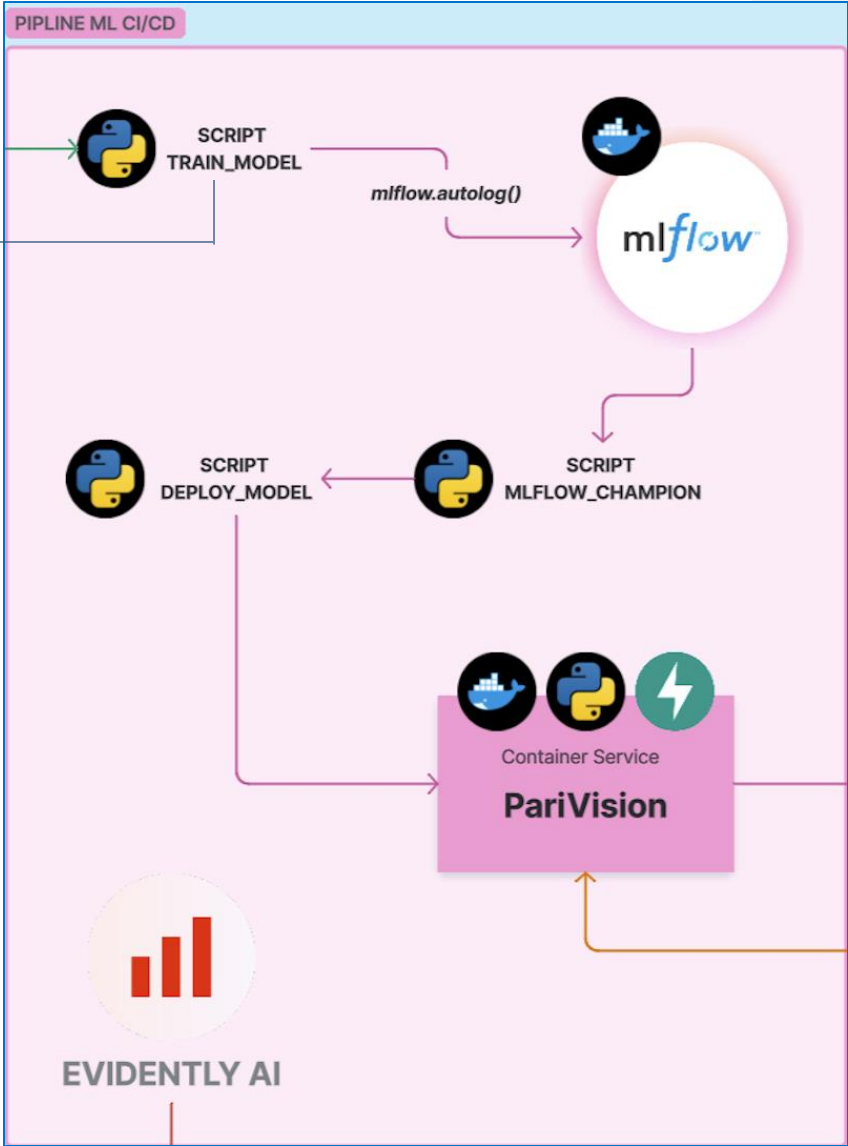
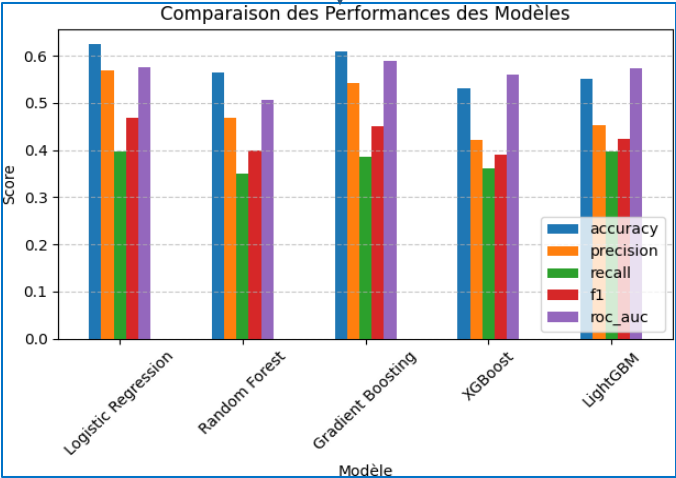
PRÉSENTATION DE LA PIPELINE



DESCRIPTION DE L'ETL



DESCRIPTION DE LA PIPELINE ML



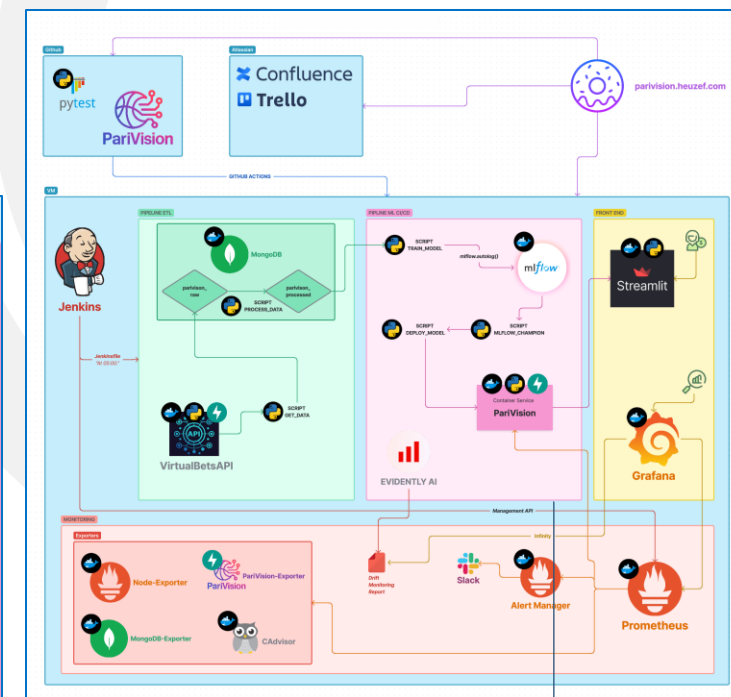
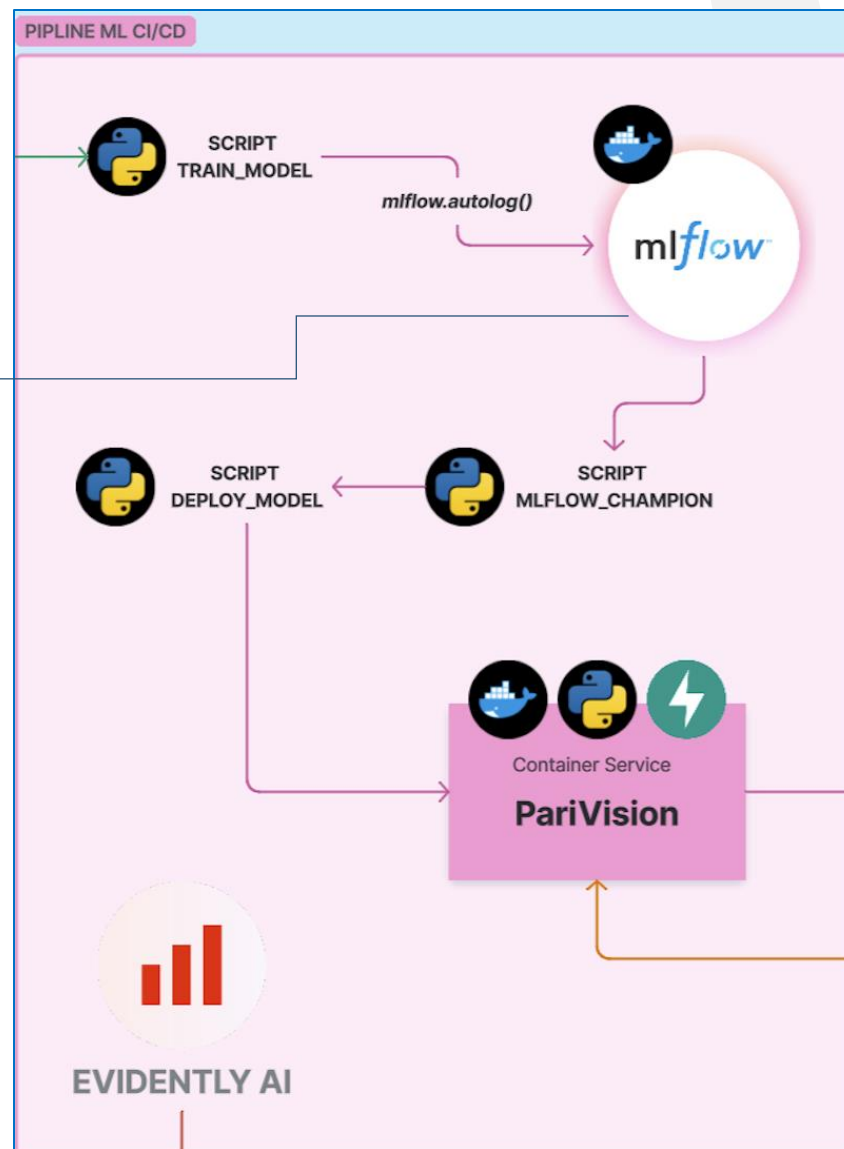
DESCRIPTION DE LA PIPELINE ML

parivision [Provide Feedback](#) [Add Description](#)

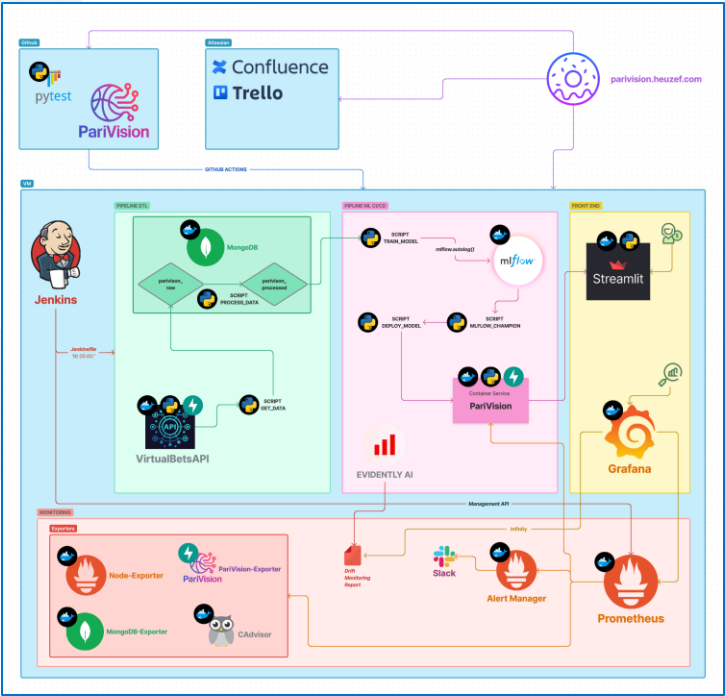
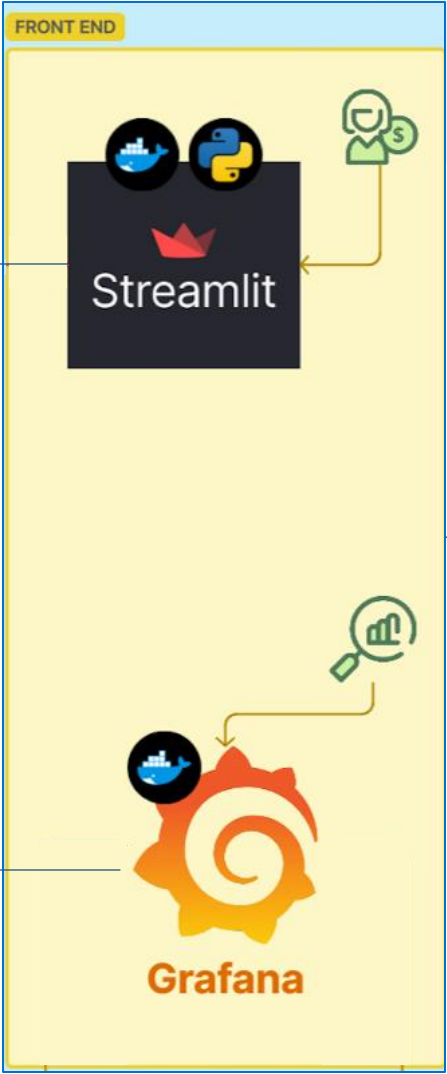
Runs Evaluation **Experimental** Traces

metrics.rmse < 1 and params.model = "tree"

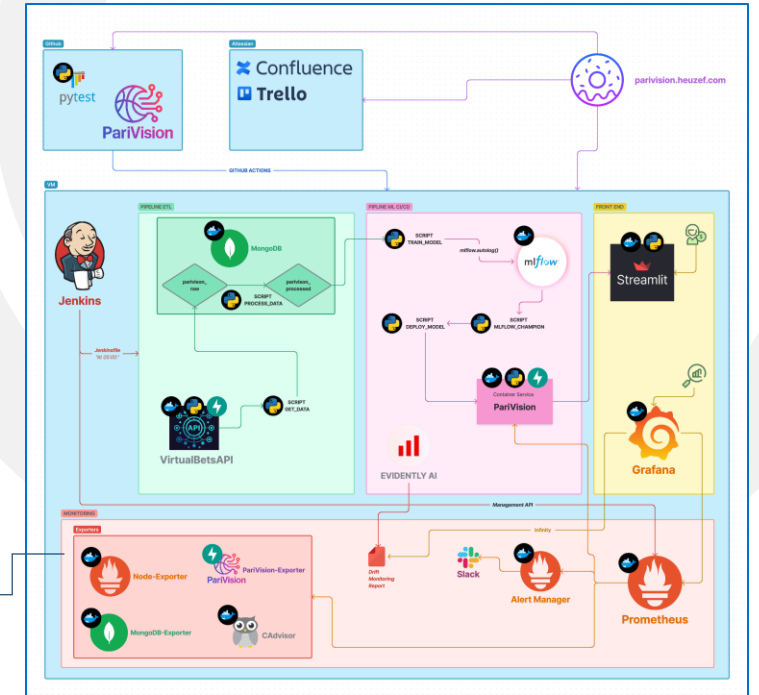
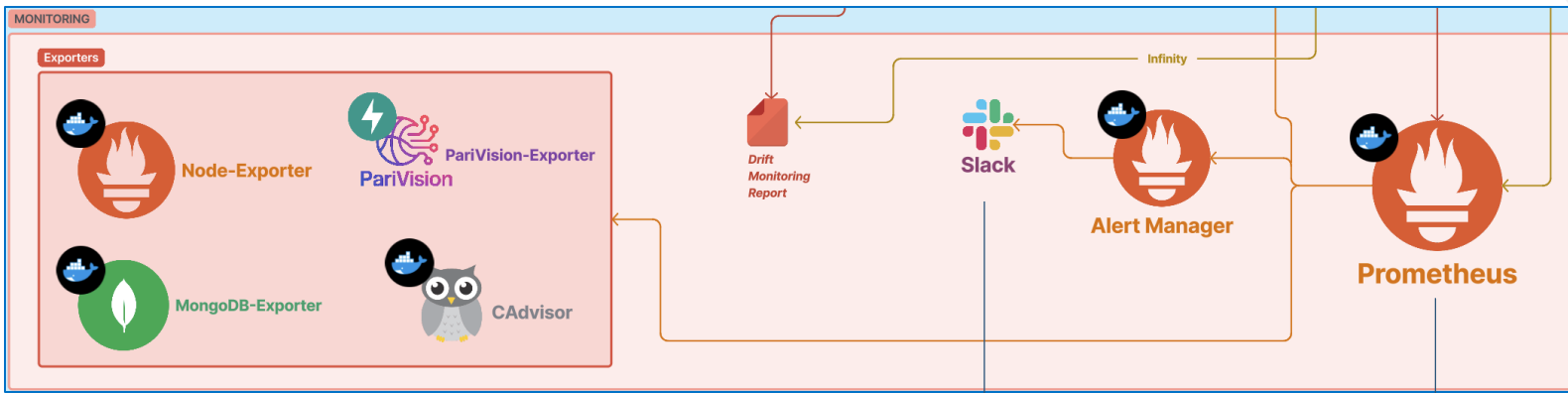
	Run Name	Created	Duration	training_score
<input type="checkbox"/>	unique-yak-156	2 hours ago	10.4s	0.6259703...
<input type="checkbox"/>	worried-trout-367	10 hours ago	10.4s	0.6316698...
<input type="checkbox"/>	legendary-koi-978	1 day ago	10.8s	0.6287529...
<input type="checkbox"/>	respected-shad-971	2 days ago	11.2s	0.6453145...
<input type="checkbox"/>	chill-robin-520	3 days ago	11.2s	0.6538181...
<input type="checkbox"/>	colorful-mole-443	4 days ago	11.1s	0.6292677...
<input type="checkbox"/>	mysterious-sloth-327	5 days ago	10.7s	0.6268672...
<input type="checkbox"/>	polite-finch-818	6 days ago	10.8s	0.6264558...
<input type="checkbox"/>	colorful-roo-484	7 days ago	10.7s	0.6264683...
<input type="checkbox"/>	dazzling-tern-571	8 days ago	10.9s	0.6358299...
<input type="checkbox"/>	adventurous-elk-590	9 days ago	11.2s	0.6302226...
<input type="checkbox"/>	puzzled-crow-644	10 days ago	11.0s	0.6258435...
<input type="checkbox"/>	victorious-hog-940	11 days ago	10.7s	0.6483982...





PRESENTATION DU FRONT END




DESCRIPTION DE LA SUPERVISION



 **PariVision** APPLI 23 h 19

 **[FIRING:1]** HighLoad (node-exporter:9100 node-exporter parivision warning)

 **[RESOLVED]** HighLoad (node-exporter:9100 node-exporter parivision warning)

```
monitoring/
├── alertmanager
│   └── alertmanager.yml
├── evidently
│   ├── html
│   ├── main.py
│   └── workspace
├── grafana
│   ├── provisioning
│   │   ├── dashboards
│   │   └── datasources
├── prometheus
│   ├── prometheus.yml
│   └── rules
│       ├── alerting.rules
│       └── recording.rules
```

DRIFT MONITORING - EVIDENTLY

Drift is detected for 0.0% of columns (0 out of 9).

â€‹ Search						
Column	Type	Reference Distribution	Current Distribution	Data Drift	Stat Test	Drift Score
> prediction	cat			Not Detected	Z-test p_value	0.141113
> Player ID	num			Not Detected	K-S p_value	0.887711
> X Location	num			Not Detected	K-S p_value	0.853919
> Shot_Distance_Meters	num			Not Detected	K-S p_value	0.559369
> Shot_Type_Encoded	num			Not Detected	Z-test p_value	0.520051
> Shot Made Flag	num			Not Detected	Z-test p_value	0.397001
> Total_Seconds_Remaining	num			Not Detected	K-S p_value	0.279666
> Y Location	num			Not Detected	K-S p_value	0.195326
> Shot_Zone_Combined	num			Not Detected	chi-square p_value	0.123966
Rows per page: 9 rows < 1-9 of 9 >						

Classification Model Performance. Target: 'Shot Made Flag'

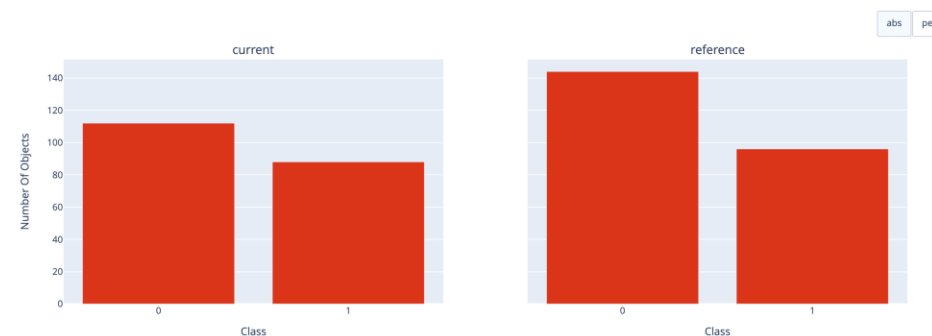
Current: Model Quality Metrics

0.57	0.516	0.375	0.434
Accuracy	Precision	Recall	F1

Reference: Model Quality Metrics

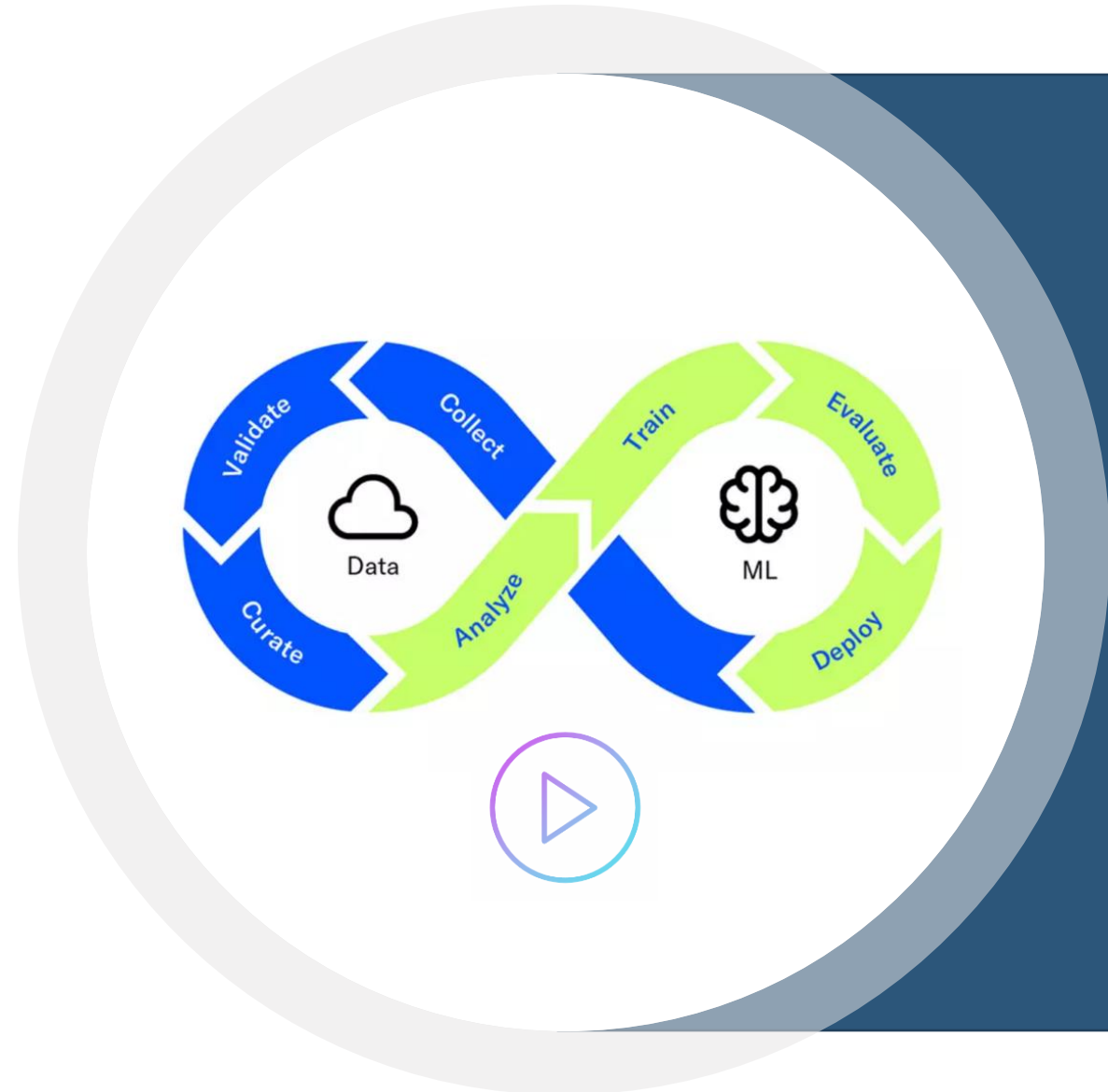
0.971	0.978	0.948	0.963
Accuracy	Precision	Recall	F1

Class Representation



DÉMONSTRATION LIVE

- 🏀 Dashboard / 2xAPIs / Supervision
- 🏀 Grafana / Drift Monitoring
- 🏀 Jenkins
- 🏀 MLFlow
- 🏀 Streamlit





PIERRE COHEN

Responsable informatique &
Développeur full stack
Datascientist



FLORENT HEUZE

Chef de projet IA
Heuzef.com



SHI JIANYING

Enseignante
Datascientist
Recherche : Les émotions,
intuitions, culture

Merci