

PRÉDICTION DES ÉVÉNEMENTS CLIMATIQUES EXTRÊMES À PARTIR DE DONNÉES MÉTÉOROLOGIQUES HISTORIQUES

REALISE PAR :

HOURRI CHAIMAE
CHEMCHAQ MARYEM
BENNISS ZINEB
REBBOUH HOUDA
SOU ABDELMOUNAIM

ENCADRE PAR:

PR. NIDAL LAMGHARI
PR. AMAL OURDOU
PR. NASSIMA SOUSSI
PR. HAMZA KHALFI

FILIÈRE : INFORMATIQUE ET INGÉIERIE DES DONNÉES





INTRODUCTION



Le changement climatique entraîne une hausse des événements météorologiques extrêmes à Marrakech, tels que canicules, vagues de froid et fortes précipitations. Ce projet utilise 15 ans de données historiques (2010–2024) pour développer un système intégrant Big Data, Deep Learning et ingénierie des connaissances, capable de prédire les anomalies climatiques et de fournir des alertes fiables via un tableau de bord interactif.

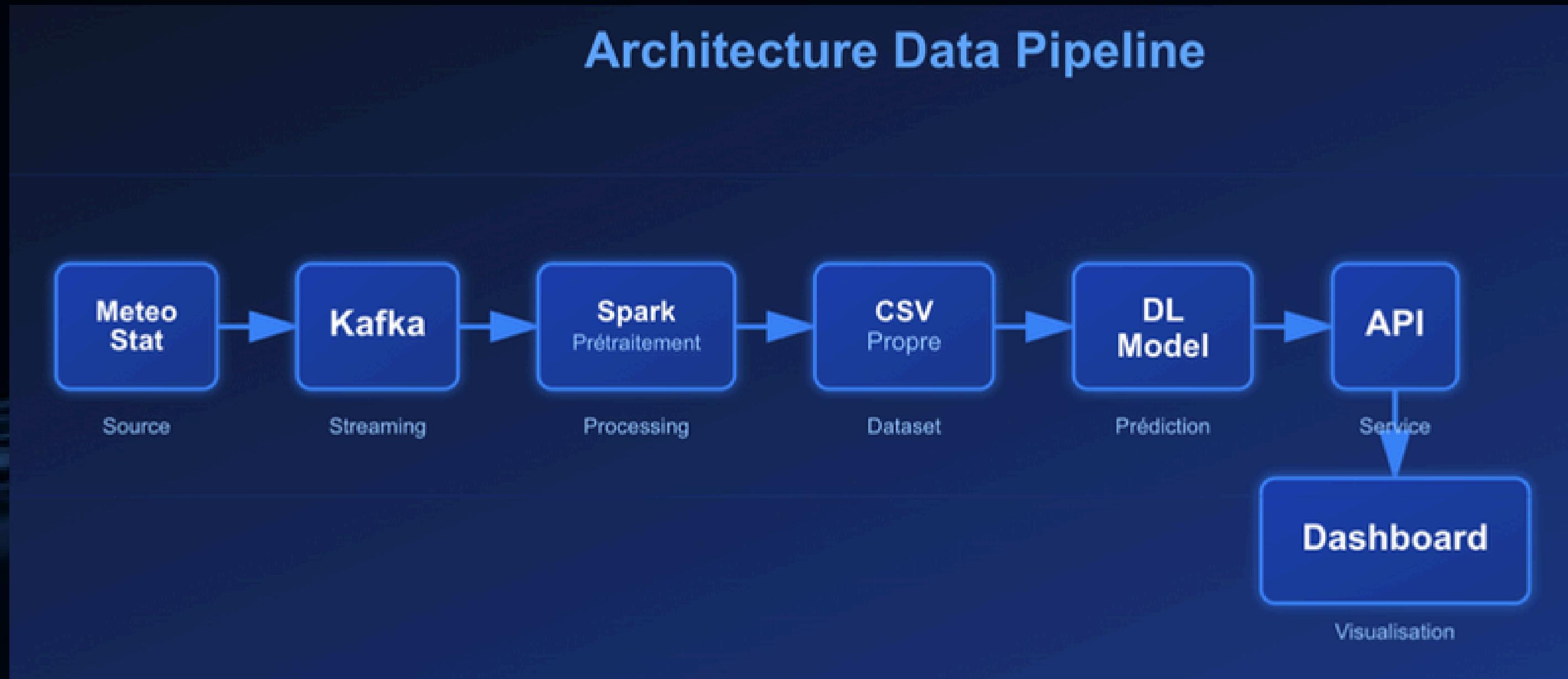
SOMMAIRE

- Démarche adoptée
- Architecture du système
- Module Big Data & ETL
- Deep Learning
- Système Expert / Ingénierie des connaissances
- Tableau de bord interactif
- Conclusion et perspectives

DÉMARCHE ADOPTÉE

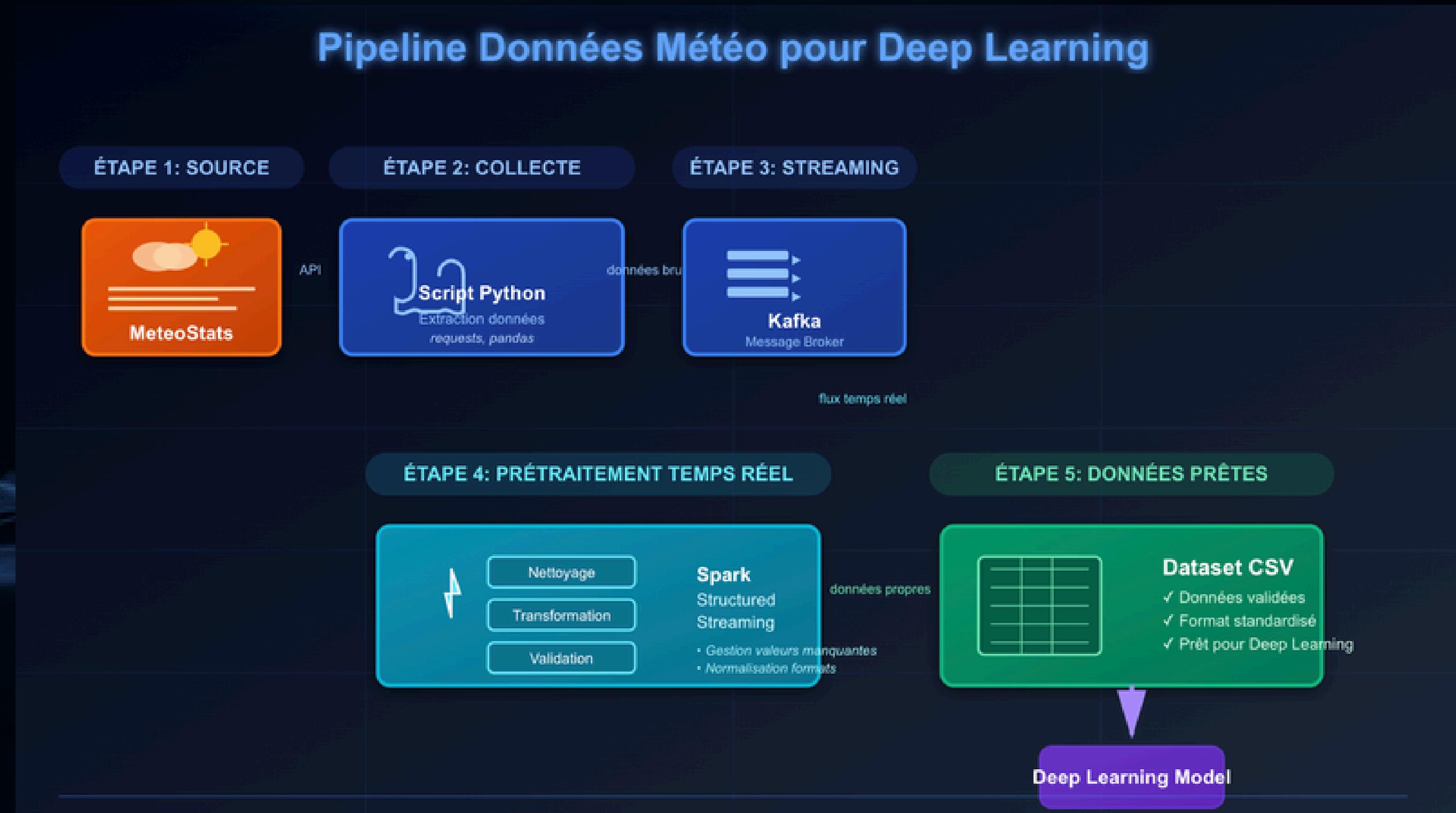


ARCHITECTURE GLOBALE



MODULE BIG DATA

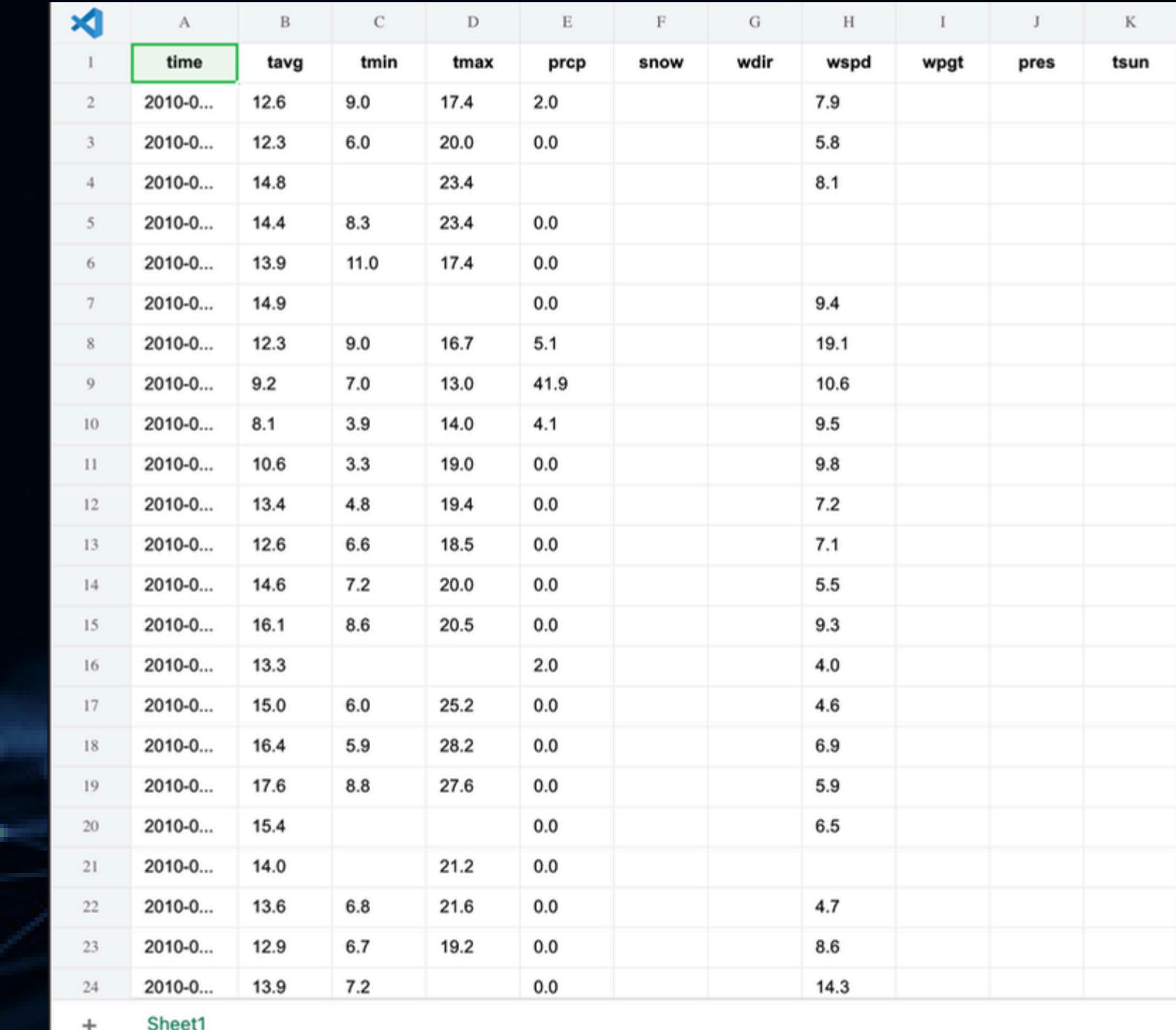
Pipeline Données Météo pour Deep Learning



1. SOURCE DES DONNÉES MÉTÉOROLOGIQUES

La bibliothèque Python Meteostat permet d'accéder facilement aux données météorologiques et climatiques via Pandas, en récupérant des données historiques provenant de sources publiques comme la NOAA et le service météorologique allemand (DWD).

Plus de 5000 lignes
de 2010 jusqu'à 2024



	A	B	C	D	E	F	G	H	I	J	K
1	time	tavg	tmin	tmax	prcp	snow	wdir	wspd	wpgt	pres	tsun
2	2010-0...	12.6	9.0	17.4	2.0						7.9
3	2010-0...	12.3	6.0	20.0	0.0						5.8
4	2010-0...	14.8		23.4							8.1
5	2010-0...	14.4	8.3	23.4	0.0						
6	2010-0...	13.9	11.0	17.4	0.0						
7	2010-0...	14.9			0.0						9.4
8	2010-0...	12.3	9.0	16.7	5.1						19.1
9	2010-0...	9.2	7.0	13.0	41.9						10.6
10	2010-0...	8.1	3.9	14.0	4.1						9.5
11	2010-0...	10.6	3.3	19.0	0.0						9.8
12	2010-0...	13.4	4.8	19.4	0.0						7.2
13	2010-0...	12.6	6.6	18.5	0.0						7.1
14	2010-0...	14.6	7.2	20.0	0.0						5.5
15	2010-0...	16.1	8.6	20.5	0.0						9.3
16	2010-0...	13.3			2.0						4.0
17	2010-0...	15.0	6.0	25.2	0.0						4.6
18	2010-0...	16.4	5.9	28.2	0.0						6.9
19	2010-0...	17.6	8.8	27.6	0.0						5.9
20	2010-0...	15.4			0.0						6.5
21	2010-0...	14.0		21.2	0.0						
22	2010-0...	13.6	6.8	21.6	0.0						4.7
23	2010-0...	12.9	6.7	19.2	0.0						8.6
24	2010-0...	13.9	7.2		0.0						14.3



2. INGESTION DES DONNÉES AVEC KAFKA

- Kafka Producer Python
- Topic : weather-marrakesh
- Messages JSON
- Streaming simulé

```
' : 15.3, 'tmax': 32.1, 'prcp': 0.0}
Sent: {'date': '2024-04-19', 'tavg': 18.4, 'tmin': 14.0, 'tmax': 24.2, 'prcp': 0.0}
Sent: {'date': '2024-04-20', 'tavg': 16.6, 'tmin': 13.0, 'tmax': 23.1, 'prcp': 1.5}
Sent: {'date': '2024-04-21', 'tavg': 17.9, 'tmin': 12.0, 'tmax': 26.1, 'prcp': 0.0}
Sent: {'date': '2024-04-22', 'tavg': 21.9, 'tmin': 13.2, 'tmax': 29.8, 'prcp': 0.0}
Sent: {'date': '2024-04-23', 'tavg': 23.2, 'tmin': 15.3, 'tmax': 31.2, 'prcp': 0.0}
Sent: {'date': '2024-04-24', 'tavg': 23.0, 'tmin': 15.0, 'tmax': 31.8, 'prcp': 0.0}
Sent: {'date': '2024-04-25', 'tavg': 19.7, 'tmin': 11.6, 'tmax': 27.7, 'prcp': 0.0}
Sent: {'date': '2024-04-26', 'tavg': 18.3, 'tmin': 11.3, 'tmax': 26.0, 'prcp': 0.0}
Sent: {'date': '2024-04-27', 'tavg': 17.2, 'tmin': 13.5, 'tmax': 23.7, 'prcp': 0.0}
Sent: {'date': '2024-04-28', 'tavg': 16.0, 'tmin': 13.7, 'tmax': 21.7, 'prcp': 0.0}
Sent: {'date': '2024-04-29', 'tavg': 15.1, 'tmin': 11.3, 'tmax': 21.4, 'prcp': 0.0}
Sent: {'date': '2024-04-30', 'tavg': 17.3, 'tmin': 11.0, 'tmax': 24.3, 'prcp': 0.0}
Sent: {'date': '2024-05-01', 'tavg': 18.8, 'tmin': 11.0, 'tmax': 27.0, 'prcp': 0.0}
Sent: {'date': '2024-05-02', 'tavg': 19.8, 'tmin': 13.5, 'tmax': 27.5, 'prcp': 0.0}
Sent: {'date': '2024-05-03', 'tavg': 22.6, 'tmin': 13.6, 'tmax': 32.7, 'prcp': 0.0}
Sent: {'date': '2024-05-04', 'tavg': 24.4, 'tmin': 14.9, 'tmax': 34.0, 'prcp': 0.0}
Sent: {'date': '2024-05-05', 'tavg': 23.1, 'tmin': 14.9, 'tmax': 33.8, 'prcp': 0.0}
Sent: {'date': '2024-05-06', 'tavg': 23.4, 'tmin': 16.5, 'tmax': 32.0, 'prcp': 0.0}
Sent: {'date': '2024-05-07', 'tavg': 25.3, 'tmin': 17.8, 'tmax': 35.0, 'prcp': 0.0}
Sent: {'date': '2024-05-08', 'tavg': 26.9, 'tmin': 18.0, 'tmax': 36.0, 'prcp': 0.0}
Sent: {'date': '2024-05-09', 'tavg': 28.1, 'tmin': 19.9, 'tmax': 35.8, 'prcp': 0.0}
Sent: {'date': '2024-05-10', 'tavg': 25.1, 'tmin': 19.3, 'tmax': 29.2, 'prcp': 0.0}
Sent: {'date': '2024-05-11', 'tavg': 23.4, 'tmin': 16.9, 'tmax': 31.6, 'prcp': 0.0}
Sent: {'date': '2024-05-12', 'tavg': 23.4, 'tmin': 15.9, 'tmax': 32.1, 'prcp': 0.0}
Sent: {'date': '2024-05-13', 'tavg': 23.9, 'tmin': 15.9, 'tmax': 32.6, 'prcp': 0.0}
Sent: {'date': '2024-05-14', 'tavg': 21.6, 'tmin': 15.1, 'tmax': 30.0, 'prcp': 0.0}
Sent: {'date': '2024-05-15', 'tavg': 19.6, 'tmin': 13.6, 'tmax': 28.9, 'prcp': 0.0}
Sent: {'date': '2024-05-16', 'tavg': 20.4, 'tmin': 12.9, 'tmax': 29.6, 'prcp': 0.0}
Sent: {'date': '2024-05-17', 'tavg': 20.7, 'tmin': 12.5, 'tmax': 29.1, 'prcp': 0.0}
Sent: {'date': '2024-05-18', 'tavg': 19.6, 'tmin': 16.9, 'tmax': 24.7, 'prcp': 0.0}
Sent: {'date': '2024-05-19', 'tavg': 18.3, 'tmin': 14.0, 'tmax': 23.8, 'prcp': 0.0}
Sent: {'date': '2024-05-20', 'tavg': 19.9, 'tmin': 13.5, 'tmax': 26.9, 'prcp': 0.0}
Sent: {'date': '2024-05-21', 'tavg': 20.4, 'tmin': 13.5, 'tmax': 28.6, 'prcp': 0.0}
Sent: {'date': '2024-05-22', 'tavg': 21.1, 'tmin': 13.8, 'tmax': 30.5, 'prcp': 0.0}
Sent: {'date': '2024-05-23', 'tavg': 21.8, 'tmin': 14.2, 'tmax': 30.5, 'prcp': 0.0}
Sent: {'date': '2024-05-24', 'tavg': 20.9, 'tmin': 13.5, 'tmax': 29.6, 'prcp': 0.0}
Sent: {'date': '2024-05-25', 'tavg': 22.3, 'tmin': 15.0, 'tmax': 29.0, 'prcp': 0.0}
Sent: {'date': '2024-05-26', 'tavg': 24.3, 'tmin': 17.0, 'tmax': 33.3, 'prcp': 0.0}
Sent: {'date': '2024-05-27', 'tavg': 25.4, 'tmin': 17.7, 'tmax': 35.0, 'prcp': 0.0}
Sent: {'date': '2024-05-28', 'tavg': 26.4, 'tmin': 17.5, 'tmax': 36.8, 'prcp': 0.0}
Sent: {'date': '2024-05-29', 'tavg': 27.3, 'tmin': 18.5, 'tmax': 38.6, 'prcp': 0.0}
Sent: {'date': '2024-05-30', 'tavg': 26.6, 'tmin': 18.5, 'tmax': 37.0, 'prcp': 0.0}
Sent: {'date': '2024-05-31', 'tavg': 25.1, 'tmin': 18.0, 'tmax': 34.8, 'prcp': 0.0}
Sent: {'date': '2024-06-01', 'tavg': 26.6, 'tmin': 18.6, 'tmax': 37.0, 'prcp': 0.0}
Sent: {'date': '2024-06-02', 'tavg': 25.4, 'tmin': 17.4, 'tmax': 35.8, 'prcp': 0.0}
Sent: {'date': '2024-06-03', 'tavg': 22.5, 'tmin': 16.4, 'tmax': 30.7, 'prcp': 0.0}
Sent: {'date': '2024-06-04', 'tavg': 23.4, 'tmin': 18.0, 'tmax': 31.9, 'prcp': 0.0}
Sent: {'date': '2024-06-05', 'tavg': 25.4, 'tmin': 17.9, 'tmax': 34.1, 'prcp': 0.0}
Sent: {'date': '2024-06-06', 'tavg': 23.2, 'tmin': 18.0, 'tmax': 31.0, 'prcp': 0.0}
Sent: {'date': '2024-06-07', 'tavg': 20.5, 'tmin': 17.7, 'tmax': 26.3, 'prcp': 0.0}
Sent: {'date': '2024-06-08', 'tavg': 22.1, 'tmin': 16.0, 'tmax': 29.3, 'prcp': 0.0}
```



3. PRÉTRAITEMENT BIG DATA AVEC APACHE SPARK

Spark consomme les données depuis Kafka et effectue le nettoyage et la transformation. Nous avons converti les types, standardisé les dates et ajouté des variables saisonnières pour capturer la périodicité climatique. Spark permet de traiter efficacement de grands volumes de données.

Spark Batch Processing

- 1 Spark Session
`SparkSession.builder()`

- 2 Read Kafka in BATCH
`spark.read.format("kafka")`

- 3 Convert to String
`CAST(value AS STRING)`

- 4 Parse JSON Schema
`from_json() → date, tavg, tmin, tmax, prcp`

- 5 Data Cleaning
`cast(DoubleType) + to_date()`

- 6 Add Seasonal Features
`dayofyear, sin_doy, cos_doy`

- 7 Save Cleaned Data
`write.parquet() + write.csv()
✓ Prêt pour Deep Learning`

4. STOCKAGE DU DATASET FINAL

A	B	C	D	E	F	G	H	
1	date	tavg	tmin	tmax	prcp	dayof...	sin_doy	cos_doy
2	2010-0...	12.6	9.0	17.4	2.0	1	0.0172133561558346...	0.99985183920911...
3	2010-0...	12.3	6.0	20.0	0.0	2	0.03442161162274574	0.9994074007397...
4	2010-0...	14.8	NaN	23.4	NaN	3	0.0516196672232537...	0.998666816288476
5	2010-0...	14.4	8.3	23.4	0.0	4	0.06880242680231986	0.9976303053065...
6	2010-0...	13.9	11.0	17.4	0.0	5	0.08596479873744647	0.9962981749346...
7	2010-0...	14.9	NaN	NaN	0.0	6	0.10310169744743485	0.99467081991152...
8	2010-0...	12.3	9.0	16.7	5.1	7	0.1202080448993527	0.9927487224577...
9	2010-0...	9.2	7.0	13.0	41.9	8	0.13727877211326478	0.9905324521322...
10	2010-0...	8.1	3.9	14.0	4.1	9	0.15430882066428117	0.9880226656636...
11	2010-0...	10.6	3.3	19.0	0.0	10	0.1712931441814776	0.9852201067560...
12	2010-0...	13.4	4.8	19.4	0.0	11	0.1882267098432442	0.9821256058680...
13	2010-0...	12.6	6.6	18.5	0.0	12	0.2051044998686192	0.9787400799669...
14	2010-0...	14.6	7.2	20.0	0.0	13	0.2219215130041655	0.9750645322571...
15	2010-0...	16.1	8.6	20.5	0.0	14	0.2386727660059501	0.97110005188295...
16	2010-0...	13.3	NaN	NaN	2.0	15	0.255353295116187	0.9668478136052...
17	2010-0...	15.0	6.0	25.2	0.0	16	0.2719581575341055	0.9623090774541...
18	2010-0...	16.4	5.9	28.2	0.0	17	0.288482432880609	0.9574851883550...
19	2010-0...	17.6	8.8	27.6	0.0	18	0.30492122465628907	0.9523775757303...
20	2010-0...	15.4	NaN	NaN	0.0	19	0.3212696616923644	0.9469877530760...
21	2010-0...	14.0	NaN	21.2	0.0	20	0.3375228995941133	0.9413173175128...
22	2010-0...	13.6	6.8	21.6	0.0	21	0.3536761221763716	0.9353679493131...
23	2010-0...	12.9	6.7	19.2	0.0	22	0.3697245428906731	0.92914141140317...
24	2010-0...	13.9	7.2	NaN	0.0	23	0.38566340624360707	0.9226395488404...

- Colonnes : tavg, tmin, tmax, prcp, sin_doy, cos_doy
- Après le prétraitement, les données sont sauvegardées dans un data lake sous format CSV . Ce dataset propre est directement utilisé pour l'entraînement du modèle Deep Learning.



DEEP LEARNING

FEATURE ENGINEERING & SEQUENCES

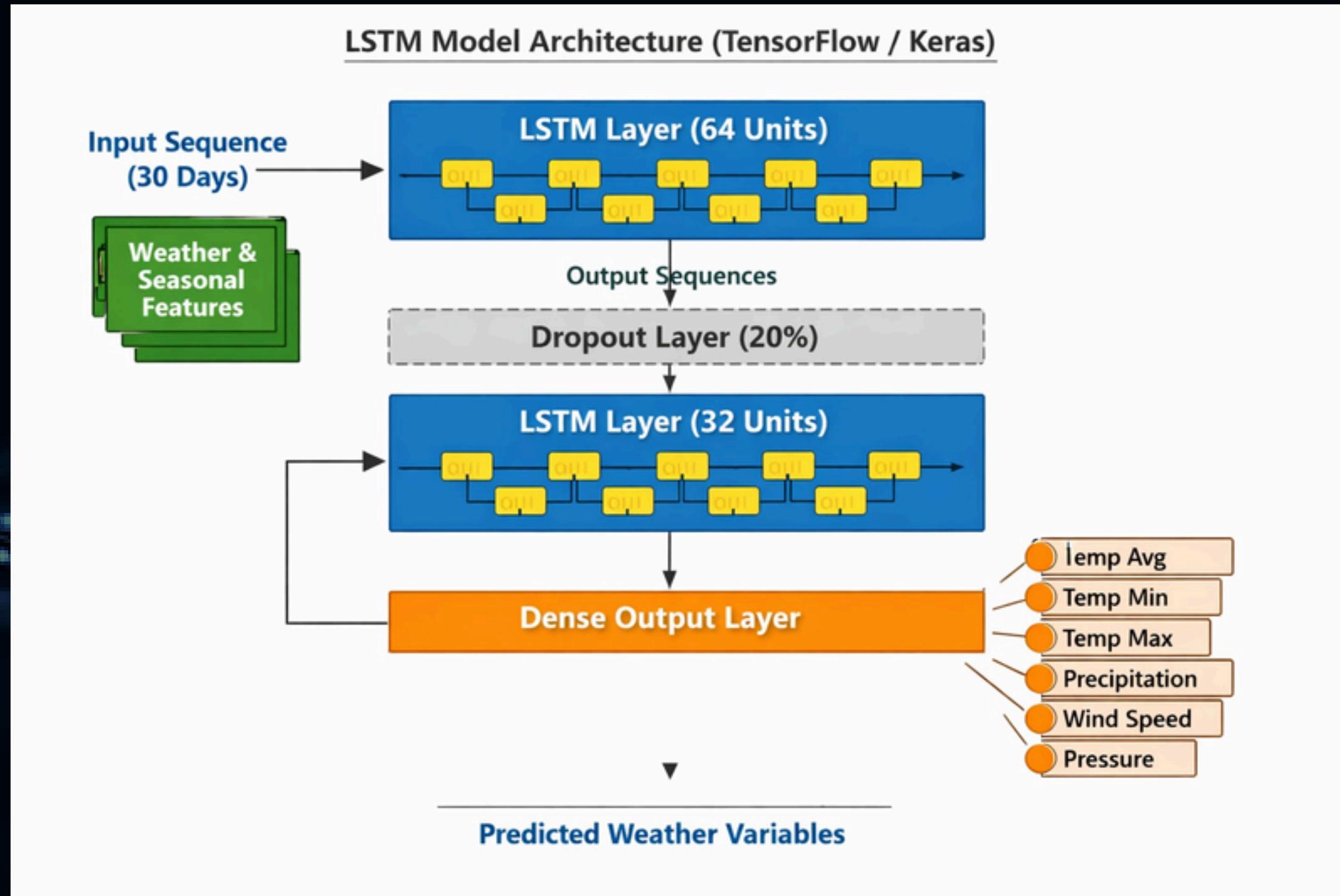
Normalisation : toutes les variables mises entre 0 et 1.

Séquences temporelles : 30 jours → entrée X, jour suivant → sortie y.



DEEP LEARNING

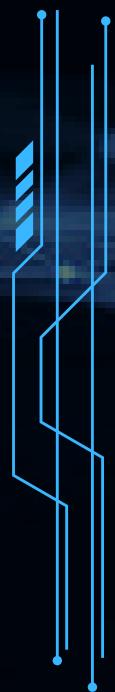
MODELE LSTM & ENTRAINEMENT



DEEP LEARNING

RESULTATS

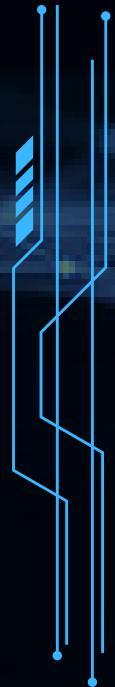
- Prédiction récursive des variables météo sur 365 jours (2026)
- Dénormalisation des résultats pour obtenir des valeurs réelles
- Modèle LSTM performant pour capter saisonnalité et dépendances temporelles
- Résultats exploitables pour analyse climatique et planification



DEEP LEARNING

METRIQUES DE PERFORMANCE

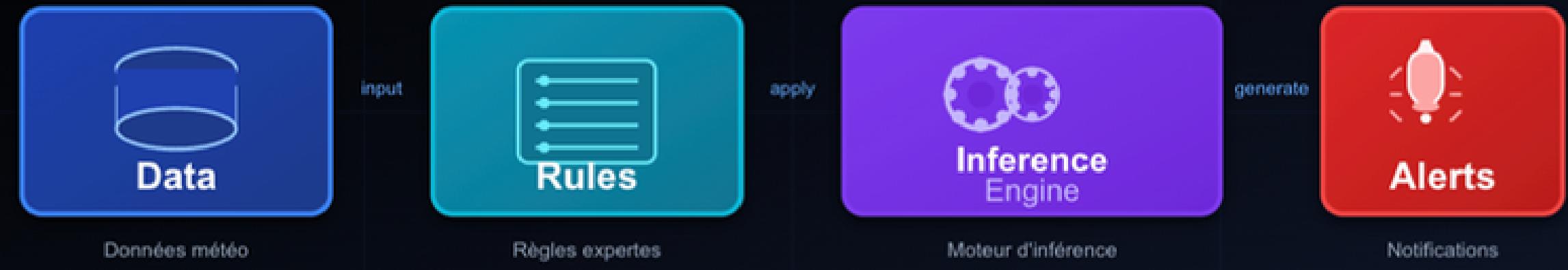
Métrique	Description	Valeur
RMSE	Erreurs moyennes entre valeurs réelles et prédictives (température)	≈ 20 °C
MAE	Erreurs absolues moyennes	≈ 15 °C
Accuracy	Taux global de bonne détection des événements extrêmes	≈ 75 %
Recall	Proportion de catastrophes réellement détectées	≈ 80 %
F1-score	Qualité globale des alertes (précision + rappel)	≈ 0.77



SYSTÈME EXPERT CLIMATIQUE

- Ontologie
- Base de règles
- Moteur d'inférence
- Simulation

Système Expert - Pipeline



RÈGLES IMPLÉMENTÉES

- Si $T_{max} \geq 42^{\circ}C$ → Alerte critique
- Si $T_{max} \geq 38^{\circ}C$ et Humidité < 20% → Risque élevé
- Si $T_{max} > 42^{\circ}C$ pendant 3 jours → Canicule persistante



TABLEAU DE BORD

MAR CLIMATE
MOROCCAN CLIMATE EXTREME EVENT PREDICTION SYSTEM

A cooperation framework between University Sultan My Slimane & ENSA Khouribga to improve awareness and prediction of extreme climate events in Morocco.

Admin Register/Login Contact Us

HOME ALERTS MAPS & IMAGERY KNOWLEDGE ABOUT

Map of disaster alerts: Morocco, 2026. The designations employed do not imply the expression of any opinion whatever on the part of the developers. [RCI - LSTN_v1]

ACTIVE ALERTS BY TYPE LAST 7 DAYS

HEATWAVES	WIND/STORM	DROUGHTS	FLOODS	VOLCANOES	FOREST FIRES
42.1°C 2026-02-03	No active alerts				
43.5°C 2026-02-04					

Maps and satellite imagery



CONCLUSION ET PERSPECTIVES

Le projet a permis de mettre en place un pipeline complet, depuis la collecte et le traitement des données jusqu'à la prédiction des événements climatiques extrêmes. Le modèle prédictif s'est montré efficace pour détecter les anomalies, et l'interface offre une visualisation claire et interactive des résultats. Ce travail fournit un support à la décision grâce à la combinaison de Deep Learning et d'un système expert explicable.

Perspectives futures :

- Intégrer des flux de données en temps réel via IoT ou satellites
- Améliorer continuellement les modèles pour plus de précision
- Déployer des alertes régionales plus fines pour mieux prévenir les risques



**MERCI
QUESTIONS ?**

