



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

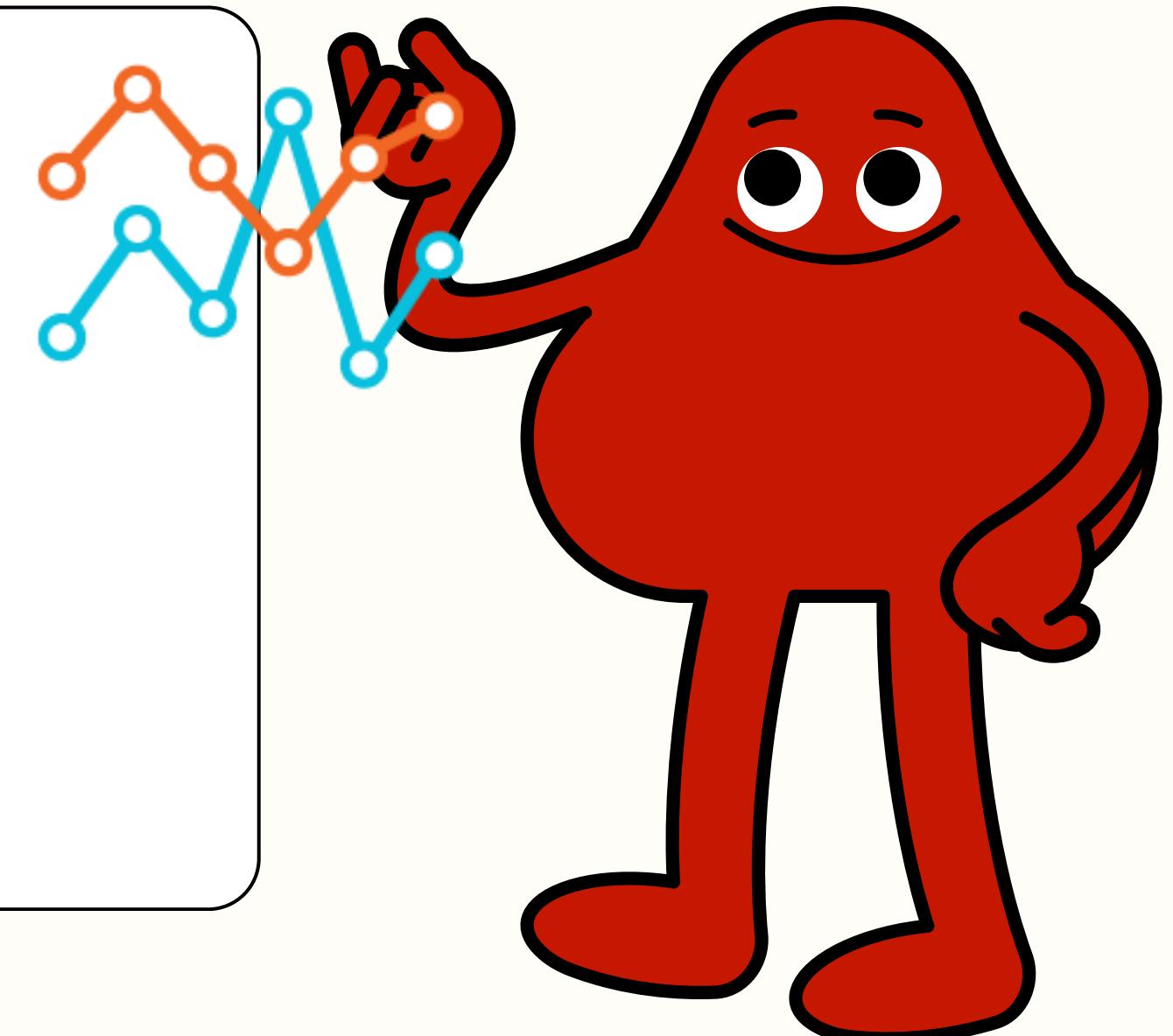


université
virtuelle
Burkina Faso



CITADEL SUMMER SCHOOL 2025

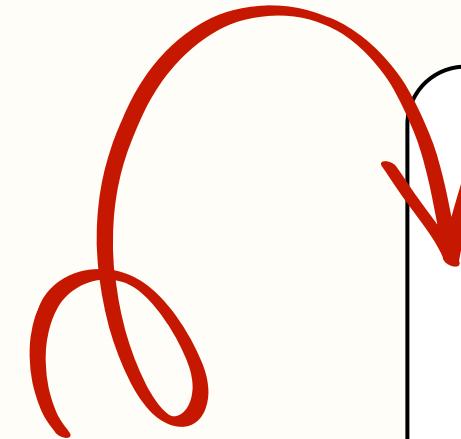
TIME SERIES



By Abdoul Aziz BONKOUNGOU



Bienvenue!



PROGRAMME

Découvrez la Révolution Numérique

1. Définition et importance
2. Problèmes fondamentaux des séries temporelles
3. Data preprocessing
4. Métriques d'évaluation
5. Décomposition d'une série temporelle
6. Prédiction des séries temporelles
7. TIME SERIES FOUNDATION MODELS
8. Utilisation des LLM dans l'analyse des time series

»» Plongeons ensemble dans l'univers des séries temporelles

DÉFINITION

Une **série temporelle** est une suite de valeurs numériques représentant l'évolution d'une quantité spécifique au cours du temps.

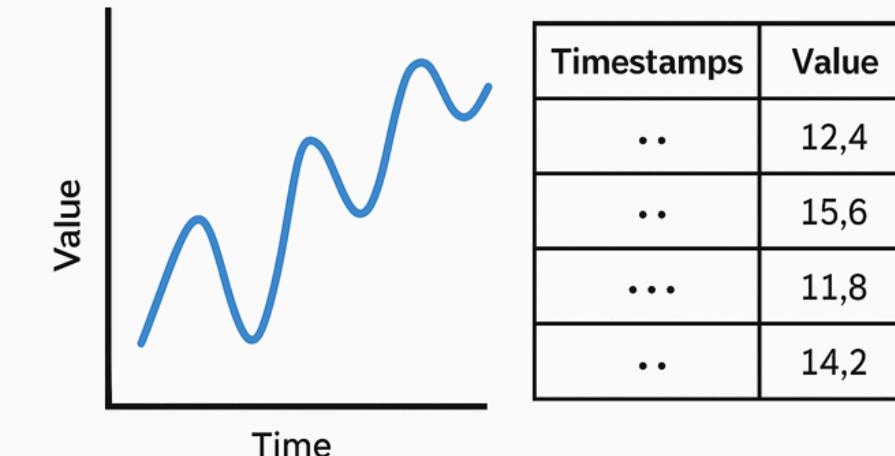
La formulation mathématique d'une série temporelle donne:

$$X = \{x_1(1), x_1(2), \dots, x_T(n)\}$$

$t \in [1, \dots, T]$ où T est la longueur de la série (nombre total d'observations)

$n \in \{1, \dots, C\}$ où C est le nombre de variables mesurées à chaque instant

- ➡ Une série temporelle peut être :
 - univariée ($C = 1$)
 - multivariée ($C > 1$)



DOMAINE D'APPLICATION

Économie & Finance

prix des actions, taux de change, inflation, prévisions économiques

Énergie & Industrie

consommation électrique, production, maintenance prédictive

Climat & Environnement

températures, précipitations, pollution, données satellitaires

Santé & Biologie

signaux ECG/EEG, suivi de patients, épidémiologie

Informatique & Réseaux

trafic réseau, détection d'anomalies, cybersécurité

Intelligence Artificielle & Data Science

prévision, détection de ruptures, séries issues de capteurs



Forecasting

Imputation

Classification

Détection
d'anomaly

- PROBLÈMES FONDAMENTAUX DES SÉRIES TEMPORELLES -

FORECASTING

Le forecasting ou la prévision consiste à prédire les valeurs futures d'une ou plusieurs variables à partir des observations passées.

Etant donné une série temporelle $X=\{x_1, x_2, \dots, x_T\}$ l'objectif de la prédiction est d'estimer $x_{T+1}, x_{T+2}, \dots, x_{T+H}$, H étant l'horizon de prédiction.

On modélise la relation avec $x_{T+h} = f(x_T, x_{T-1}, \dots, x_1)$, $\forall h \in \{1, \dots, H\}$

Les modèles de prévision sont utilisés pour :

- anticiper l'évolution d'un phénomène,
- aider à la prise de décision,
- optimiser la planification et la gestion des ressources.



Forecasting

Imputation

Classification

Détection d'anomaly

- PROBLÈMES FONDAMENTAUX DES SÉRIES TEMPORELLES -

RESSOURCES

- [A comprehensive survey of deep learning for time series forecasting – Kim et al. \(May 2025\)](#)
 - Revue approfondie des architectures de deep learning pour la prévision de séries temporelles, incluant RNN, CNN, Transformers et modèles hybrides.
- [Deep Time Series Models: A Comprehensive Survey and Benchmark - Wang et al. \(September 2025\)](#)
 - Analyse comparative des modèles récents de séries temporelles, accompagnée d'un benchmark expérimental sur des jeux de données standards.
- [25 years of time series forecasting: a selective guide to the literature – De Gooijer & Hyndman \(January 2006\)](#)
 - Article fondateur retraçant l'évolution des méthodes classiques de prévision, notamment ARIMA et les techniques de lissage exponentiel.
- [A Review of Time Series Forecasting Methods – Shah \(April 2024\)](#)
 - Revue synthétique comparant les modèles statistiques traditionnels, les approches de machine learning et les modèles hybrides.

Scannez-moi





Forecasting

Imputation

Classification

Détection
d'anomaly

- PROBLÈMES FONDAMENTAUX DES SÉRIES TEMPORELLES -

IMPUTATION

L'imputation consiste à estimer et remplacer les valeurs manquantes dans une série temporelle

$X=\{x_1, x_2, \dots, x_{t-1}, \text{NaN}, x_t+1, \dots, x_T\}$ où NaN représente une valeur manquante à l'instant t.
L'objectif est d'estimer \tilde{x}_t , une approximation de la valeur manquante t, à partir des observations disponibles.

$$\tilde{x}_t = f(x_{t-1}, x_{t-2}, \dots) \text{ ou } \tilde{x}_t = g(x_{t+1}, x_{t+2}, \dots)$$

Types de données manquantes

- MCAR (Missing Completely At Random)
- MAR (Missing At Random)
- MNAR (Missing Not At Random)

La nature du manque influence fortement la qualité de l'imputation.

L'imputation est souvent une étape préalable essentielle avant la prévision, la classification ou la détection d'anomalies.



Forecasting

Imputation

Classification

Détection
d'anomaly

- PROBLÈMES FONDAMENTAUX DES SÉRIES TEMPORELLES -

RESSOURCES

- [Deep Learning for Multivariate Time Series Imputation: A Survey - Wang et al. \(IJCAI 2025\)](#)
 - Revue complète des méthodes d'imputation pour séries temporelles multivariées, proposant une taxonomie basée sur l'incertitude d'imputation et les architectures de deep learning.

Scannez-moi





Forecasting

Imputation

Classification

Détection
d'anomaly

- PROBLÈMES FONDAMENTAUX DES SÉRIES TEMPORELLES -

CLASSIFICATION

La classification consiste à attribuer une étiquette à une série temporelle (ou à un segment de série) parmi un ensemble de classes prédéfinies en fonction de sa dynamique temporelle.

Soit une série temporelle X, la classification cherche à apprendre une fonction :

$$y = f(X)$$

où :

- X est une série temporelle (univariée ou multivariée),
- $y \in \{1, \dots, K\}$ est une classe parmi K possibles.

Examples:

- Analyse de signaux médicaux (ECG ou EEG): Classes :rythme normal, arythmie, crise épileptique
- Surveillance industrielle. Classes: fonctionnement normal,panne mécanique, usure
- Informatique & cybersécurité. Classes :trafic normal, attaque,surcharge



Forecasting

Imputation

Classification

Détection
d'anomaly

- PROBLÈMES FONDAMENTAUX DES SÉRIES TEMPORELLES -

RESSOURCES

- [Time Series Classification: A Review of Algorithms and Implementations - Faouzi \(March 2024\)](#)
 - Panorama des principales approches de classification des séries temporelles, des méthodes classiques aux modèles profonds, avec ressources open-source.

Scannez-moi



Forecasting

Imputation

Classification

Détection
d'anomaly

- PROBLÈMES FONDAMENTAUX DES SÉRIES TEMPORELLES -

DÉTECTION D'ANOMALY

Une anomalie (ou outlier) est une observation qui s'écarte significativement du comportement majoritaire des données.
Ces observations suggèrent qu'un mécanisme différent de celui qui génère les données normales est à l'origine du phénomène observé.

Deux interprétations fondamentales des anomalies

Anomalies comme bruit (noise)

- dues à des erreurs de mesure ou de transmission,
- non informatives,
- généralement supprimées lors du nettoyage des données.

Anomalies comme nouveauté (novelty)

- événements rares mais significatifs,
- révélateurs de comportements inhabituels,
- souvent l'objectif principal de l'analyse.

Types d'anomalies

Anomalies ponctuelles
Valeurs isolées anormales
→ pic, chute brutale

Anomalies contextuelles
Valeurs anormales dans un contexte donné (temps, saison)
→ température élevée en hiver

Anomalies collectives
Sous-séquences anormales
→ dérive prolongée, changement de dynamique



Forecasting

Imputation

Classification

Détection
d'anomaly

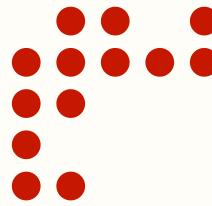
- PROBLÈMES FONDAMENTAUX DES SÉRIES TEMPORELLES -

DÉTECTION D'ANOMALY

- [Deep Learning for Time Series Anomaly Detection: A Survey - Darban et al.\(Obtober 2024\)](#)
 - Revue exhaustive des méthodes de deep learning pour la détection d'anomalies dans les séries temporelles, proposant une taxonomie structurée, une analyse des architectures et des benchmarks de référence.
- [Dive into Time-Series Anomaly Detection: A Decade Review – Paparrizos et al. \(2024/2025\)](#)
 - Revue de référence couvrant dix ans de recherche en détection d'anomalies, avec une taxonomie centrée sur les processus, une méta-analyse des tendances et une synthèse des défis ouverts.

Scannez-moi





COMPOSANTS D'UNE SÉRIE TEMPORELLE

Une série temporelle résulte de la superposition de plusieurs composantes qui influencent son comportement dans le temps, notamment une tendance, des variations saisonnières et du bruit aléatoire.

Formulation mathématique

$$X(t) = T(t) + S(t) + \varepsilon(t)$$

où :

- $T(t)$: tendance
- $S(t)$: saisonnalité
- $\varepsilon(t)$: bruit

L'objectif de l'analyse des séries temporelles est souvent d'isoler ces composantes pour mieux comprendre, modéliser et prédire les données.

Tendance (Trend)

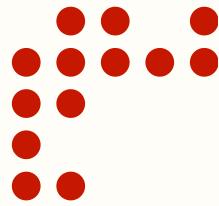
- Évolution globale à long terme de la série
- Représente la direction générale (croissante, décroissante, stable)

Saisonnalité (Seasonality)

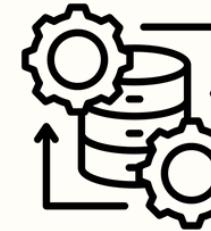
- Motifs qui se répètent périodiquement
- Liés à des effets calendaires

Bruit (Noise / Résidu)

- Variations aléatoires non expliquées
- Contient les erreurs de mesure et perturbations imprévues
- Remarque :
 - Le bruit limite la capacité de prévision



DATA PREPROCESSING



Le prétraitement des données est une étape essentielle en analyse de séries temporelles, visant à améliorer la qualité des données avant toute modélisation ou analyse.

Principales étapes du data preprocessing

- **Imputation des valeurs manquantes**
- **Resampling et agrégation**
- **Détection et traitement des valeurs aberrantes**
- **Normalisation / transformation des données**

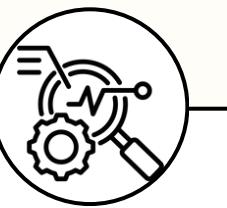
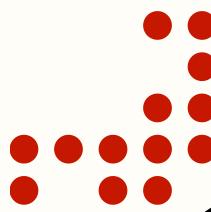
Imputation des valeurs manquantes

- Remplacer les observations absentes (NaN)
- Méthodes courantes :
 - interpolation (linéaire, spline)
 - moyenne / médiane locale
 - modèles statistiques ou ML

Resampling et agrégation

- Changer la fréquence temporelle :
 - horaire → journalier
 - journalier → mensuel
- Fonctions d'agrégation :
 - moyenne, somme, maximum, minimum
- **Détection et traitement des valeurs aberrantes**
 - Identifier et corriger ou supprimer des observations anormales pouvant fausser l'analyse et la modélisation.
- **Normalisation / transformation des données**
 - Mettre les données sur une échelle comparable ou appliquer des transformations (log, standardisation) pour faciliter l'apprentissage des modèles.

»» **Un bon prétraitement conditionne fortement la qualité des modèles**



METRIQUES D'EVALUATION

Les métriques d'évaluation permettent de quantifier l'erreur entre les valeurs observées et les valeurs prédites,

Métriques basées sur l'erreur absolue et quadratique

- MAE – Mean Absolute Error $MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|$
- MSE – Mean Squared Error $MSE = \frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2$
- RMSE – Root Mean Squared Error $RMSE = \sqrt{MSE}$

Métriques relatives (pourcentages)

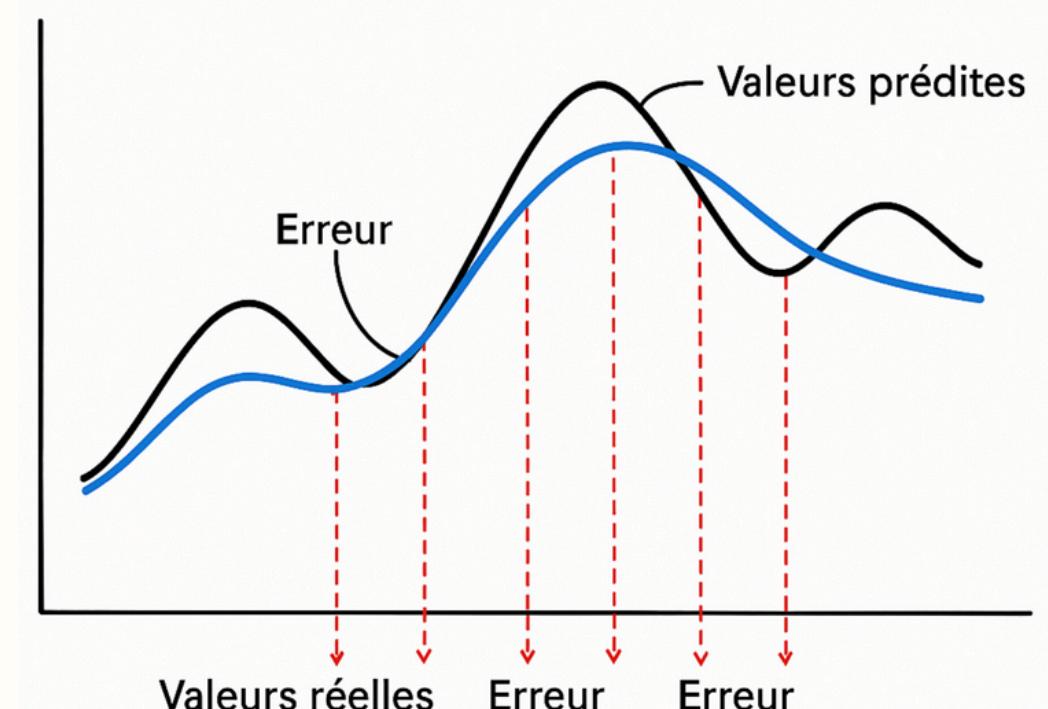
- MAPE – Mean Absolute Percentage Error $MAPE = \frac{100}{T} \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right|$
- SMAPE – Symmetric MAPE $SMAPE = \frac{100}{T} \sum_{t=1}^T \frac{|y_t - \hat{y}_t|}{(|y_t| + |\hat{y}_t|) / 2}$

Prévision probabiliste

- CRPS – Continuous Ranked Probability Score
 - Mesure la distance entre la distribution prédite et la valeur observée.
 - (plus faible = meilleure prévision)

$$CRPS(F, y) = \int_{-\infty}^{+\infty} (F(z) - \mathbb{1}_{\{z \geq y\}})^2 dz$$

où : F est la distribution prédite, y est la valeur observée.



INTRODUCTION AUX METHODES STATISTIQUES

Moyenne mobile et Exponential Smoothing

Objectif : lisser une série temporelle pour réduire le bruit et faire ressortir la tendance sous-jacente.

- **Moyenne mobile (Moving Average)**

- Méthode de lissage consistant à calculer la moyenne sur une fenêtre de points voisins.
- SMA – Simple Moving Average
 - → moyenne des n dernières observations
- WMA – Weighted Moving Average
 - → moyenne pondérée : les observations récentes ont plus de poids
- (Idée clé : plus la fenêtre est large, plus le lissage est fort)

- **Lissage exponentiel (Exponential Smoothing)**

- Extension de la moyenne mobile
- Pondération exponentielle décroissante
- Les observations récentes ont beaucoup plus d'impact

»» Ces méthodes sont simples, interprétables et constituent une première approche du forecasting en séries temporelle



INTRODUCTION AUX MÉTHODES STATISTIQUES

ARIMA – AutoRegressive Integrated Moving Average

ARIMA est un modèle statistique qui combine dépendance temporelle et stationnarité pour réaliser des prévisions.

- Décomposition du modèle ARIMA(p, d, q)
 - AR (p) – AutoRegressive
 - → dépendance aux p valeurs passées
 - I (d) – Integrated
 - → nombre de différenciations pour rendre la série stationnaire
 - MA (q) – Moving Average
 - → dépendance aux q erreurs passées
- Limites
 - ARIMA suppose une série stationnaire
 - Peu adapté aux fortes saisonnalités
 - Paramètres parfois difficiles à choisir manuellement
 - ➡ motivation naturelle pour SARIMA et Auto-ARIMA

»» Lorsque la série présente une saisonnalité marquée, ARIMA devient insuffisant :
On introduit alors SARIMA.



INTRODUCTION AUX METHODES STATISTIQUES

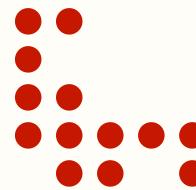
SARIMA – Seasonal ARIMA

- SARIMA étend ARIMA pour prendre en compte la saisonnalité des séries temporelles.

Auto-ARIMA

- Auto-ARIMA explore automatiquement plusieurs modèles ARIMA/SARIMA et sélectionne le meilleur. Il repose sur :
 - des tests de stationnarité (pour choisir d et D)
 - une recherche sur (p, q, P, Q)
 - des critères statistiques :
 - AIC
 - BIC
 - Le modèle retenu minimise la complexité tout en maximisant la qualité d'ajustement.

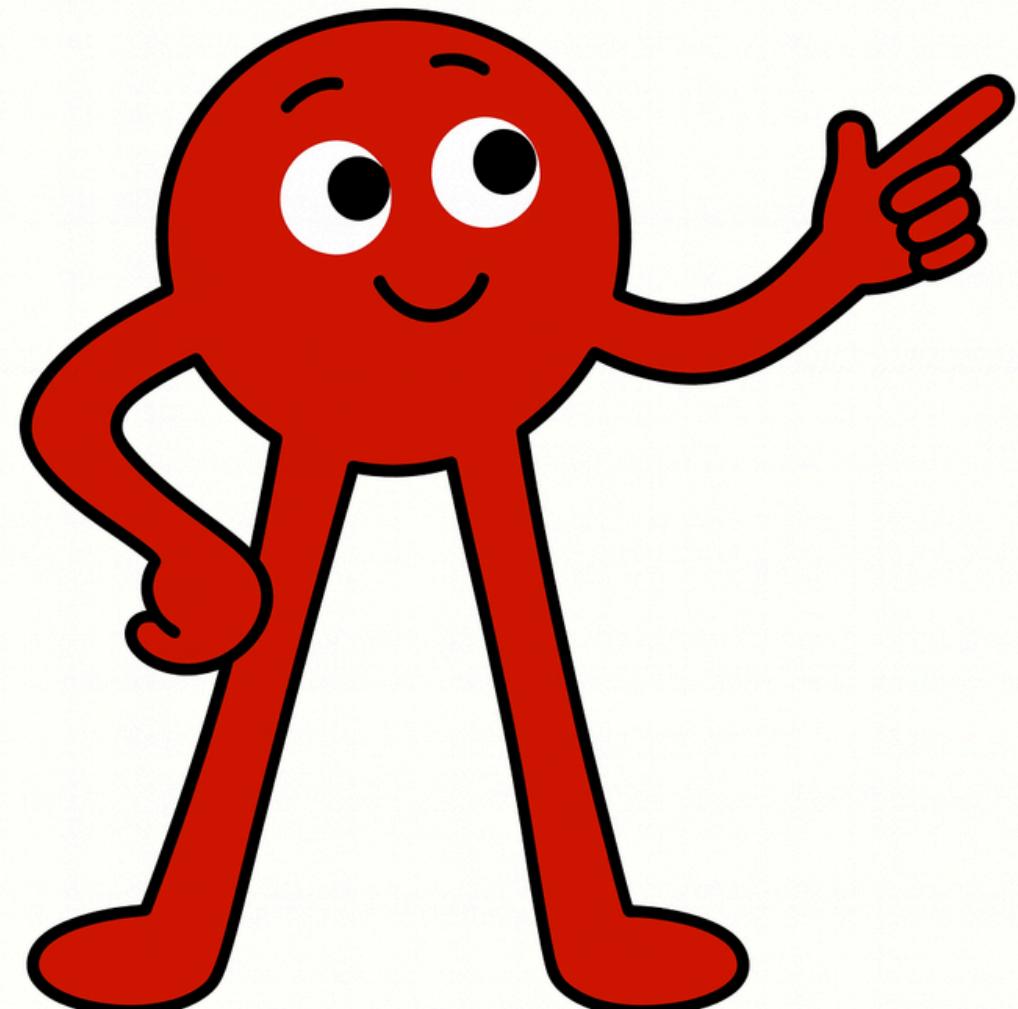




INTRODUCTION AUX METHODES STATISTIQUES

TRAVAUX PRATIQUES

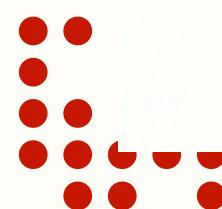
À vous de jouer

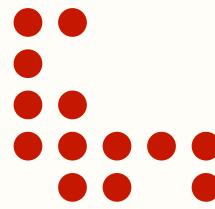


Scannez ce QR code pour
accéder aux travaux
pratiques (TP)
sur les modèles
statistiques.



Follow-me





Les modèles fondationnels pour les séries temporelles sont des modèles pré-entraînés à grande échelle, capables de généraliser à de nombreuses tâches et domaines temporels avec peu ou pas d'adaptation.

Chronos (Amazon)

- Modèle fondationnel pour le forecasting
- Séries temporelles tokenisées
- Architecture Transformer autoregressive
- Zero-shot / few-shot forecasting

Forces :

- Très bonne généralisation
- Simple à utiliser



Moirai / Uni2TS (Salesforce)

- Modèle fondationnel multi-tâches
- Un cadre unifié pour :
 - forecasting
 - imputation
 - classification
 - anomaly detection
- Séries multivariées et hétérogènes

Forces :

- Forte capacité de transfert
- Approche réellement universelle



Time-MoE

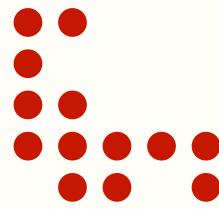
- Architecture Mixture-of-Experts
- Spécialisation automatique des experts
- Adapté aux données massives

Forces :

- Excellente scalabilité
- Très bonnes performances à grande échelle



TIME SERIES FOUNDATION MODELS



Passons maintenant à la pratique

Scannez ce QR code pour accéder aux travaux pratiques (TP)
sur les modèles fondationnels pour les séries temporelles.



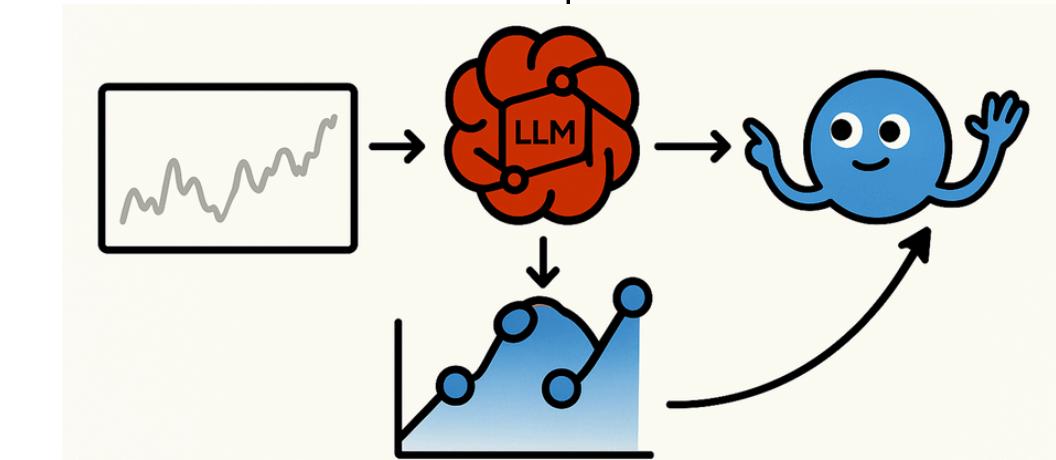
Follow-me

TIME SERIES FOUNDATION MODELS

UTILISATION DES LLM DANS L'ANALYSE DES TIME SERIES

Contexte et motivation

- **Les méthodes classiques (ARIMA, modèles d'état, ML peu profond) sont limitées :**
 - Faible capacité à modéliser les non-linéarités
 - Difficulté à capturer les dépendances à long terme
- **Les Transformers ont amélioré la modélisation temporelle, mais :**
 - Ils restent spécialisés par tâche
 - Leur capacité de généralisation inter-domaines est limitée
- **Apport des LLM: Les LLM possèdent :**
 - Une capacité de représentation générale à grande échelle
 - Des mécanismes d'attention globale
 - Une intégration de connaissances cross-domaines et cross-modales



» Principe fondamental : Les séries temporelles peuvent être transformées en tokens et traitées comme des séquences, permettant aux LLM d'apprendre des dynamiques temporelles complexes

UTILISATION DES LLM DANS L'ANALYSE DES TIME SERIES

Comment les LLM sont utilisés pour l'analyse des séries temporelles

Input-level techniques (adaptation des données)

- **Transformation des séries temporelles en tokens :**

- Quantification
- Mise à l'échelle
- Encodage séquentiel

- **Objectif :**

- Rendre les séries compatibles avec les LLM textuels
- Préserver l'ordre temporel et la dynamique

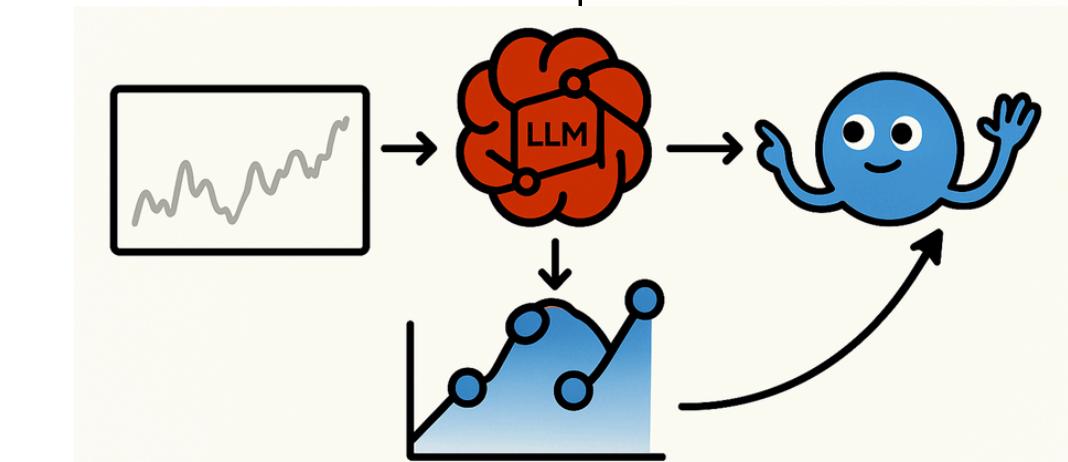
Optimization-level techniques (adaptation du modèle)

- **Fine-tuning partiel**

- Ajustement des embeddings positionnels
- Ajustement des couches de normalisation

- **Alignement temporel**

- Adapter les LLM aux propriétés spécifiques des séries temporelles



UTILISATION DES LLM DANS L'ANALYSE DES TIME SERIES

Comment les LLM sont utilisés pour l'analyse des séries temporelles

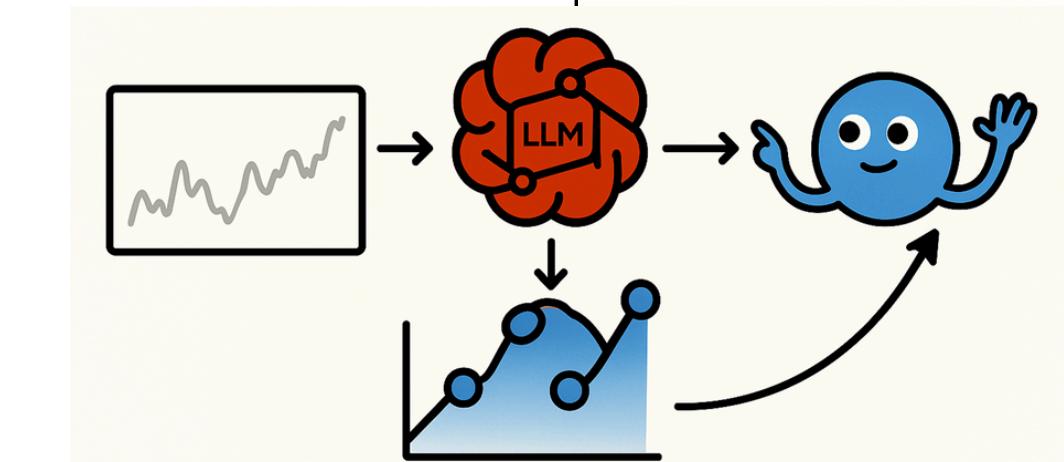
Lightweight-level techniques (déploiement efficace)

- **Réduction des coûts via :**
 - Prompting
 - In-context learning
 - Soft prompts
- **Avantage :**
 - Peu ou pas de ré-entraînement
 - Adaptation rapide à de nouvelles tâches

Applications couvertes

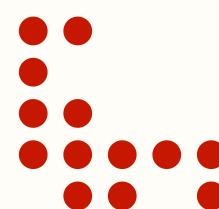
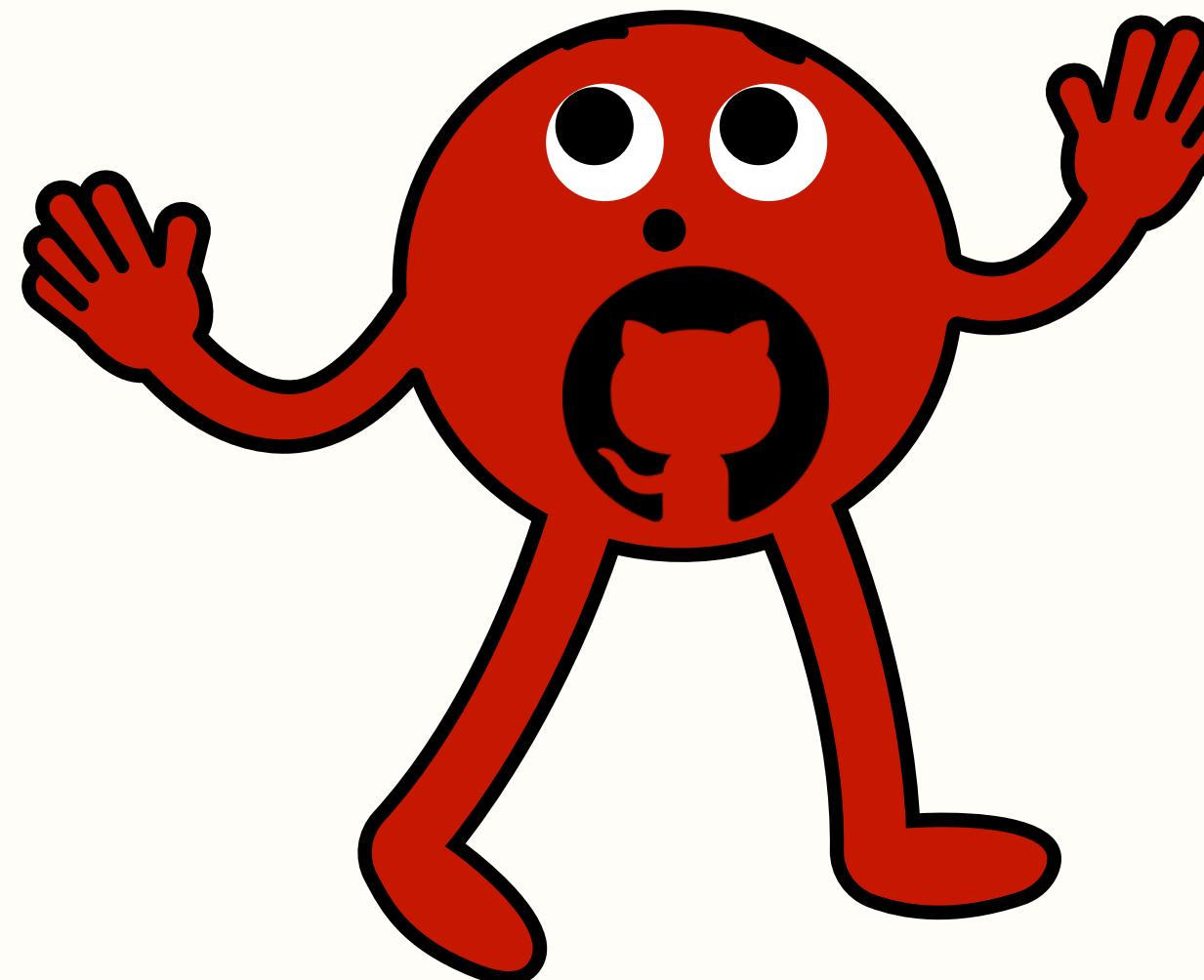
- Prévision (majoritaire)
- Classification
- Détection d'anomalies
- Analyse multimodale (TS + texte, métadonnées)

»» un modèle généraliste, pré-entraîné, capable de transférer des connaissances entre tâches et domaines temporels.





RESSOURCES GITHUB



Du code pour aller plus vite

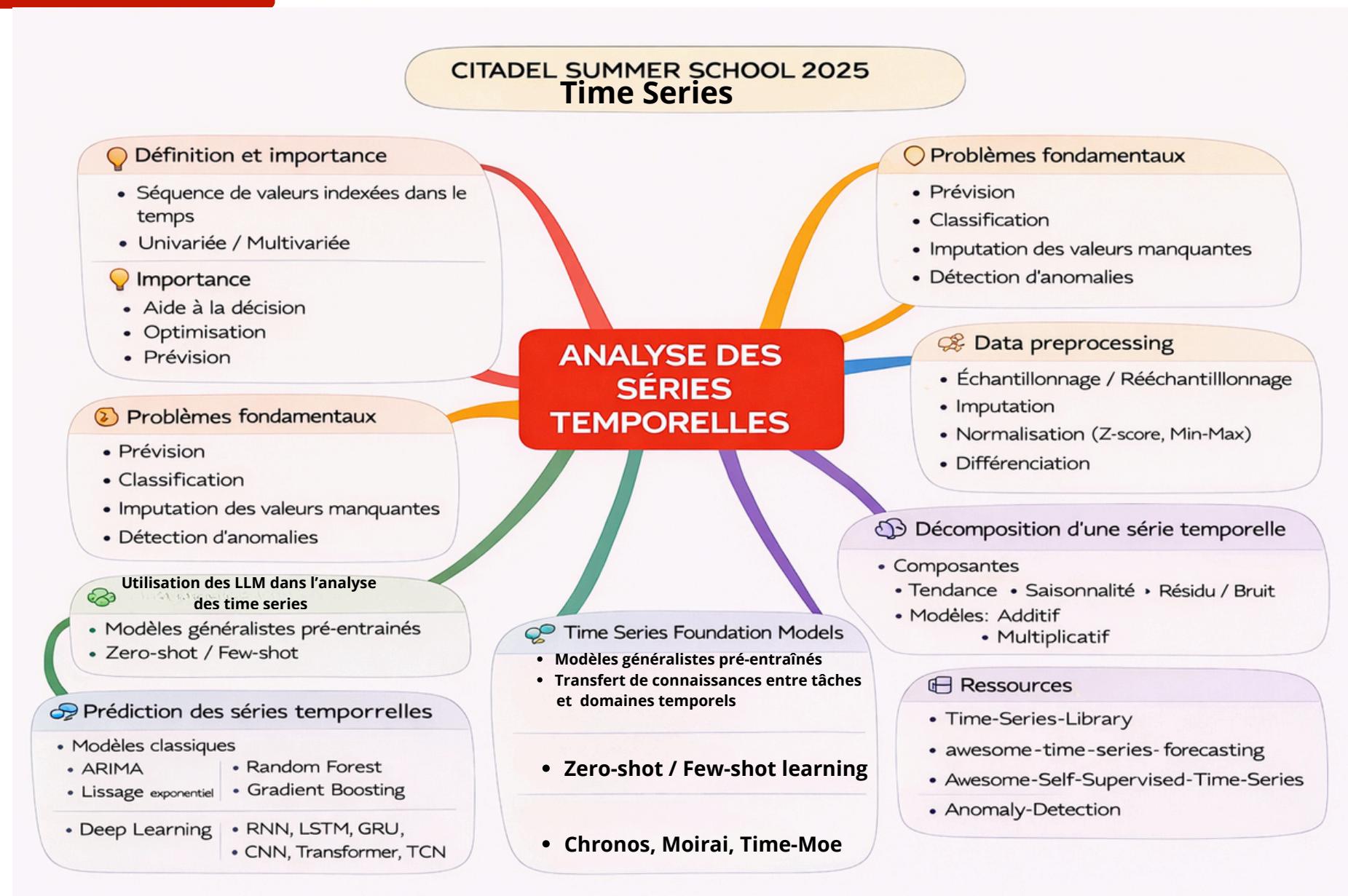
Bibliothèques et utilitaires généraux

- [Time-Series-Library](#)
- [awesome-time-series-forecasting](#)
- [Awesome-Self-Supervised-Time-Series-Anomaly-Detection](#)



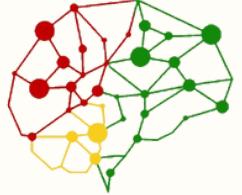
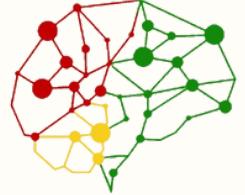
CONCLUSION

Récapitulatif



>>> Nous espérons que cette introduction vous donnera envie d'aller plus loin !

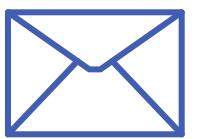




MERCI

de votre attention et de votre participation !

Continuons la discussion



abbonkoungou@gmail.com

