

Mini Projet

Le site [Yahoo Finance](https://finance.yahoo.com/) contient de nombreuses séries temporelles multidimensionnelles (actions, indices, devises, cryptomonnaies).

Vous allez :

1. Importation et préparation des données

- Choisir une série multidimensionnelle (ex. prix, volume, indicateurs techniques).
- La transformer en série temporelle sous **R** (avec vecteur temps).
- Diviser l'échantillon : 90 % pour l'apprentissage, 10 % pour le test.

2. Analyse exploratoire

Visualiser la série (trends, saisonnalités, anomalies).

Produire et interpréter ACF, PACF, décomposition.

Détecter et commenter les valeurs aberrantes.

3. Modélisation classique

Estimer et comparer des modèles stochastiques : ARIMA, SARIMA, VAR, modèles de volatilité (ARCH/GARCH).

Vérifier les hypothèses et faire les diagnostics de résidus.

4. Modélisation avancée

Implémenter et comparer des réseaux de neurones récurrents (LSTM, GRU).

Essayer des architectures hybrides (ARIMA-LSTM, Prophet-RNN,...).

Évaluer la robustesse avec validation croisée.

5. IA générative et LLM (4 points)

Utiliser un LLM pour :

- Générer automatiquement des hypothèses de modèles à tester à partir des statistiques descriptives.
- Expliquer les résultats de manière vulgarisée (rapport automatisé).
- Générer des recommandations d'investissement simulées (explicitement les limites et risques).

Comparer la cohérence entre explications humaines et générées par IA.

6. Prévision et évaluation

Produire des prévisions à horizon $h=5 \dots 10$

Comparer avec les observations réelles (MSE, RMSE, MAE, MAPE).

Tester les performances relatives des modèles.

7. Extensions possibles

Introduire des variables exogènes (actualités économiques, tweets, indices de sentiment).

Appliquer une détection d'événements avec des embeddings de texte (transformers).

Étudier la corrélation dynamique entre deux séries (ex. action et indice sectoriel).

Livrables

1. Codes complets en R et Python (avec commentaires).

2. Rapport détaillé (Word/PDF) :

Introduction du problème.

Étapes d'analyse, graphiques et diagnostics.

Comparaison détaillée des modèles.

Interprétation des résultats et apport des LLM.

3. Notebook exécutable (RMarkdown ou Jupyter) permettant de reproduire l'étude.

4. Discussion critique :

Limites des modèles.

Avantages/inconvénients de l'utilisation d'IA générative pour l'analyse financière.

Des projets seront sélectionnés d'une manière aléatoire pour faire une soutenance des travaux.

Partie IA générative et LLM : 4 points

Le reste : 16 points

Dernier délai pour la remise est le **12 Novembre 2025** à 23 h 55mn par mail, les soutenances auront lieu ultérieurement :

ahmed.dhouibi@tek-up.de