2025/2026

RAPPORT DE PROJET

Volatility Forecasting : GARCH vs LSTM

Élèves:
TAKATRI Marwa (DS)
EN-NASRY Salma (DS)
EL ALAMI Salma (SE)
REGRAGUI Basma (SE)

Professeur
MR.JANATI HICHAM

1. Problématique et motivation

La volatilité est une mesure fondamentale du risque en finance. Elle quantifie l'ampleur des variations de prix d'un actif et constitue un élément central dans de nombreuses applications : gestion de portefeuille, couverture, pricing d'options ou encore estimation du *Value-at-Risk (VaR)*. Prévoir correctement la volatilité future est donc un enjeu majeur, tant pour la gestion des risques que pour la prise de décision d'investissement.

Traditionnellement, les économètres ont recours à des modèles paramétriques tels que le **GARCH (Generalized AutoRegressive Conditional Heteroskedasticity)**, qui capturent la dépendance temporelle de la variance conditionnelle. Ces modèles reposent cependant sur des hypothèses fortes (normalité des résidus, symétrie des chocs) et peinent à modéliser des comportements non linéaires complexes.

L'avènement du **Deep Learning**, notamment des réseaux récurrents de type **LSTM (Long Short-Term Memory)**, permet de dépasser ces limites en apprenant directement les relations dynamiques dans les séries temporelles sans hypothèses de forme fonctionnelle explicite. Ce projet vise ainsi à **comparer empiriquement la performance des modèles GARCH et LSTM** dans la prévision de la volatilité conditionnelle sur des actifs financiers réels.

2. Objectifs du projet

Les objectifs sont les suivants :

— Comparer la performance prédictive des modèles **GARCH** et **LSTM** sur des séries temporelles financières (indices boursiers, cryptomonnaies) ;

— Évaluer les performances hors échantillon à l'aide de métriques quantitatives : RMSE, MAPE, log-vraisemblance ;

— Analyser le comportement des modèles en période de forte instabilité (crises, pics de volatilité) ;

— Identifier les forces et limites de chaque approche : robustesse, interprétabilité, sensibilité aux paramètres.

3. Données utilisées

Les données proviendront de prix journaliers historiques d'actifs financiers téléchargés via la librairie `yfinance`. Deux séries principales seront exploitées :

— **S&P 500 (SPX)** : indice large du marché actions américain ;

— **BTC/USD** : actif à forte volatilité, représentatif de marchés non stationnaires.

Les rendements logarithmiques seront calculés selon :

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

et serviront de base à la modélisation de la volatilité. Les données seront divisées en deux sous-échantillons : un jeu d'apprentissage (80%) et un jeu de test (20%) pour évaluer la capacité de prévision hors échantillon.

4. Méthodologie envisagée

1. **Prétraitement des données** : nettoyage, visualisation, tests de stationnarité (ADF), normalisation et transformation en rendements.
2. **Modélisation GARCH** : estimation des modèles GARCH(1,1), EGARCH et GJR-GARCH via la librairie `arch`. Validation par analyse des résidus et tests d'autocorrélation.
3. **Modélisation LSTM** : construction d'un réseau LSTM univarié avec fenêtre glissante, normalisation des séquences et optimisation par validation croisée (nombre de neurones, taux d'apprentissage, epochs).
4. **Évaluation comparative** : calcul des métriques (RMSE, MAE, MAPE), test de Diebold–Mariano pour la comparaison statistique des prévisions, et analyse graphique des volatilités prévues.
5. **Discussion et interprétation** : comparaison des résultats, analyse des régimes de volatilité et identification des contextes où l'un des modèles surpasse l'autre.

Test de Diebold–Mariano (DM). Afin de comparer la performance prédictive des modèles GARCH et LSTM de manière statistiquement rigoureuse, nous utilisons le **test de Diebold–Mariano (1995)**. Ce test permet d'évaluer si la différence entre les erreurs de prévision des deux modèles est significative ou simplement due au hasard. On définit la différence de pertes entre les modèles par :

$$d_t = L(e_{1,t}) - L(e_{2,t}),$$

où $L(e_t)$ est une fonction de perte (par exemple e_t^2 pour l'erreur quadratique). Sous l'hypothèse nulle $H_0 : \mathbb{E}[d_t] = 0$, les deux modèles ont une performance équivalente. La statistique du test est donnée par :

$$DM = \frac{\bar{d}}{\sqrt{\widehat{\mathrm{Var}}(\bar{d})}}, \qquad \text{où} \qquad \bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t.$$

Une valeur absolue élevée de DM (et une p-valeur faible) indique qu'un modèle fournit des prévisions significativement meilleures que l'autre.