Exercise 4D: Time-Series Cross-Validation

BUSI 722: Data-Driven Finance II

Demonstrate time-series cross-validation for hyperparameter selection using the same ranked features and target from Exercise 3.

Submission

Submit a **Jupyter notebook** (`.ipynb`) containing all code, output, and charts. Use markdown cells for any written discussion.

Demonstrate time-series CV for hyperparameter selection on a single prediction date (January 2020):

1. Define the training window as all data before Jan 2023 through December 2024. Create 4 time-series folds within the training window, each of 6 months duration. For each fold, train LightGBM with three different max-depth values (3, 5, 7) and compute the Spearman correlation on the validation set.
2. Report the average Spearman across folds for each max-depth. In a markdown cell, state which depth performs best.
3. In a markdown cell, discuss whether the improvement from CV is large enough to justify the computational cost.