Exercise 3D: Prediction Quality Over Time

BUSI 722: Data-Driven Finance II

Using the train/test split in Exercise 3B and the four models trained in Exercises 3B and 3C (Ridge, Lasso, Random Forest, LightGBM), analyze how prediction quality varies over time.

Submission

Submit a **zip file** containing your **Jupyter notebook** (`.ipynb`) with all code, output, and charts, and the **`predictions.parquet`** file produced in this exercise. Use markdown cells for any written discussion.

---

1. For each of the four models, plot the **monthly Spearman correlation** over time on a single chart. In a markdown cell, discuss whether there are periods where all models perform poorly.
2. Compute the fraction of months with positive Spearman correlation for each model.
3. Create a summary table with: model name, mean Spearman correlation, median Spearman correlation, fraction of months positive, and standard deviation.
4. Save the predictions from the best-performing model as `predictions.parquet` (columns: `ticker`, `month`, `return`, `pred`) for use in later exercises.