**Assignment 4.1 — Multi-Lookback Time-Series Training Pipeline for Meta-Model Ensembling**

**Objective**

Implement a reproducible training pipeline that trains separate models for multiple lookback windows (in days), saves outputs in a strict directory layout, and generates per-lookback prediction CSVs for later ensembling with a Ridge regressor.

The three shared example notebooks — Nvidia\_Stock\_Market\_History\_SimpleAttention\_365D\_Week3.ipynb, ...\_90D\_Week4.ipynb, and ...\_1D\_Week3.ipynb — serve as examples of all necessary components and configurations. Start from these and generalize.

**Data**

Use any suitable time-series dataset (e.g., NVIDIA stock history) from Kaggle.com (requires a Kaggle account). Document the dataset source in your notebook.

**Required Lookbacks**

Train one model per lookback for all of:  
[365D, 90D, 14D, 1D]

**Key Modeling Constraints**

* **No Conv1D for 1D lookback**: with only one timestep, a temporal kernel can’t convolve meaningfully (kernel >1 can’t slide; kernel=1 adds no temporal context). Use BiGRU layers instead.
* **Regularization**: when appropriate, apply kernel\_regularizer=regularizers.l2(1e-5) on the last BiGRU layer to reduce overfitting.
* Use attention or a comparable sequence modeling head consistent with the examples.

**Pipeline Requirements**

For **each lookback**:

1. Prepare sliding-window sequences for the specified lookback.
2. Define and train a model suited to that window (respecting constraints above).
3. Produce and save:
   * Prediction plot (\*\_predictions.png)
   * Training/validation loss plot (\*\_loss.png)
   * Predictions CSV (\*\_predictions.csv) with aligned dates and predicted close.
4. Write clean, parameterized code so only the lookback and model block need changing.

**Directory & Saving Conventions (strict)**

* Final save root must end with:  
  .../training/ensemble\_inputs/ + subfolder formatted with the number of days + D (e.g., 365D).
* For **each lookback**, save all artifacts in a timestamped subfolder named with the lookback (e.g., 365D/2025-08-13\_15-42-10/), where the folder name ends with YYYY-MM-DD\_HH-MM-SS.

**Workflow Hint (strongly recommended)**

1. Complete the longest lookback period notebook first.
2. Make copies for the remaining lookbacks, changing only the model’s layers and any sequence-length-dependent shapes. Keep data I/O, training loop, plotting, and saving logic as shown in the shared notebooks.

**Deliverables**

* One notebook per lookback (4 total) that saves artifacts to the required structure.
* In each run, produce and save:
  + \*\_predictions.png, \*\_loss.png
  + \*\_predictions.csv
* A short README cell at the top describing the dataset, features used, train/val/test split, and any regularization you applied to the final BiGRU.

**Grading Rubric**

|  |  |
| --- | --- |
| Criteria Points |  |
| Correctness & Reproducibility: runs end-to-end for all lookbacks; artifacts saved with exact structure/names. | 40 |
| Model Design: appropriate architectures per lookback; no Conv1D for 1D; BiGRU regularizer applied when appropriate. | 25 |
| Code Quality: commented, parameterized, minimal duplication, clear functions. | 20 |
| Analysis & Plots: clean prediction/loss plots; brief rationale/notes. | 15 |