**Assignment 1.1 — Repository Essentials for a Time-Series ML Pipeline (NVDA Baseline)**

**Overview**

Create three repository documents, README.md, LICENSE.txt, and CITATION.cff, for a time-series project of your choosing. For Week 1, you will follow the provided notebooks to build an NVIDIA stock prediction baseline in Colab and place these documents at the project root in your Google Drive project directory (authored locally in VS Code (or IDE of your chooing), then uploaded).

**Learning goals**

* Apply compliance-first documentation for datasets and pretrained models.
* Produce clear, reproducible repo metadata recognized by colleagues and tools.
* Practice VS Code → Google Drive → Colab project organization.

**Required project context (Week 1 baseline)**

* **Dataset:** NVIDIA Stock Market History Dataset (Kaggle, Adil Shamim), **CC0 1.0**.
* **Model:** FinBERT (ProsusAI), **Apache-2.0**; not an official Prosus product; contacts: Dogu Araci, Zulkuf Genc.

**Tasks**

1. **Create README.md (in VS Code)**
   * Project title and 3–5 sentence overview (time series + sentiment).
   * “Project Structure” section that reflects your Drive/Colab layout (root files + notebooks/).
   * “Data” subsection citing the Kaggle NVDA dataset with source URL + CC0 license URL.
   * “Model” subsection citing FinBERT with repo URL + Apache-2.0 license URL + non-official note.
   * “License” summary pointing to LICENSE.txt.
   * “Citation” note pointing to CITATION.cff.
2. **Create LICENSE.txt (in** VS Code, chosen IDE**)**
   * Choose a license for your code (MIT or Apache-2.0 recommended).
   * Add a Third-Party Notices section listing:
     + NVIDIA Stock Dataset — CC0 1.0 (URL).
     + FinBERT — Apache-2.0 (URL) + non-official note.
   * (Optional) Add LICENSES/Apache-2.0.txt and LICENSES/CC0-1.0.txt with full texts.
3. **Create CITATION.cff (in VS Code)**
   * Include your name(s), project title, version, date, and your chosen license.
   * **References**:
     + Dataset: NVIDIA Stock Dataset (Adil Shamim), CC0 1.0, source URL.
     + Software: FinBERT (ProsusAI), Apache-2.0, repo URL.
4. **Upload to Google Drive project directory**
   * Place README.md, LICENSE.txt, and CITATION.cff at the root of your project folder.

**Deliverables**

* A short note with the Drive path to your project folder (e.g., My Drive/Nvidia\_Stock\_Market\_History/) and confirmation that files appear at the root.
* The three files (README.md, LICENSE.txt, CITATION.cff) uploaded to Colab project structure as attachments.

**Acceptance checklist**

* Filenames are exact: README.md, LICENSE.txt, CITATION.cff.
* README includes data/model sources, license links, and usage for Colab.
* LICENSE.txt includes your project license and third-party notices.
* CITATION.cff is valid YAML and contains references for the dataset and FinBERT.
* Notebooks are under notebooks/; docs are at the project root.

**Grading rubric (100 pts)**

* README.md completeness & clarity — 40 pts
* LICENSE.txt correctness & third-party notices — 40 pts
* CITATION.cff validity & references — 20 pts