**Person Re-Identification & Scene Classification — Take-Home**

**Task Description**

You are given 4 short MP4 clips (no audio). Your job is to write code that will create:

**A. Person Identity Catalogue (cross-clip).**

Create a global roster of unique people appearing anywhere in the dataset. The same individual must receive the same global ID across all clips where they appear.

**B. Scene Labelling (per-clip).**

For each clip, output a “normal” or “crime” label with a brief justification grounded in what is visible.

**Data**

- Folder: Videos

- 1.mp4, 2.mp4, 3.mp4, 4.mp4

**Requirements**

**A. Person Identity Catalogue**

Produce a machine-readable artifact that, for every global person ID, lists all their appearances with at least:

- clip\_id

- time span(s) or frame range(s)

- an internal reference to the person instance(s) (your choice of format)

**B. Scene Labelling**

Produce a dataset-level file containing one record per clip with:

- clip\_id

- label ∈ {normal, crime}

- a concise justification that references timestamps and, where applicable, the global person IDs from Part A.

If something is ambiguous, make a decision and state why.

**Deliverables**

Provide a public repository link in GitHub that contains the following:

1. Source code in python that reproduces your results.

- Include clear run instructions (README): environment setup, commands, and any external assets you rely on.

2. Outputs (checked into the repo or generated by one command and documented).

- Identity catalogue (Part A) in your chosen machine-readable format.

- Per-clip labels (Part B) in a single machine-readable file.

- Any optional visualizations you deem helpful (e.g., annotated frames or short previews) are welcome but not required.

3. Brief write-up (≤2 pages).

- Your approach, assumptions, limitations, and anything you would do next with more time.

**Rules & Logistics**

- Any libraries or pre-trained models may be used. If you rely on external services/APIs, document clearly.

- Optimize for clarity, determinism, and reproducibility. If randomness is involved, fix a seed.

- Keep the scope reasonable; this is a take-home, not a production system.