



Université Chouaib Doukkali  
Ecole Nationale des Sciences Appliquées d'El Jadida  
Département Télécommunications, Réseaux et Informatique  
Module : **MLops**



## **TP**

Filière : **SDIA**  
Niveau : **2<sup>ème</sup> Année**

Lab 3 :

### **Versionnement des données et pipelines ML avec DVC**

Réalisé Par :  
**Wassima RAICHI**

Année Universitaire : 2025/2026

# I- Instalation DVC :

```
C:\Users\hp>python --version
Python 3.11.7

C:\Users\hp>pip install dvc
Collecting dvc
  Obtaining dependency information for dvc from https://files.pythonhosted.org/packages/ec/cf/da62f36c21fe7d6ebf3dd494cd3a72dc432b06ea4daf6afb5660f4ec7561/dvc-3.65.0-py3-none-any.whl.metadata
  Downloading dvc-3.65.0-py3-none-any.whl.metadata (17 kB)
Requirement already satisfied: attrs>=22.2.0 in c:\users\hp\appdata\local\programs\python\python311\lib\site-packages (from dvc) (25.4.0)
Collecting celery (from dvc)
  Obtaining dependency information for celery from https://files.pythonhosted.org/packages/01/4e/53a125038d6a814491a0ae3457435c13cf8821eb602292cf9db37ce35f62/celery-5.6.0-py3-none-any.whl.metadata
  Downloading celery-5.6.0-py3-none-any.whl.metadata (23 kB)
Requirement already satisfied: colorama>=0.3.9 in c:\users\hp\appdata\local\programs\python\python311\lib\site-packages (from dvc) (0.4.6)
Collecting configobj>=5.0.9 (from dvc)
  Obtaining dependency information for configobj>=5.0.9 from https://files.pythonhosted.org/packages/a6/c4/0679472c60052c27efa612b4cd3dd2a23e885dcdc73461781d2c802d39e/configobj-5.0.9-py2.py3-none-any.whl.metadata
  Downloading configobj-5.0.9-py2.py3-none-any.whl.metadata (3.2 kB)
Collecting distro>=1.3 (from dvc)
  Obtaining dependency information for distro>=1.3 from https://files.pythonhosted.org/packages/12/b3/231ffd4ab1fc9d679809f356cebee130ac7daa0d6d6f3206dd4fd137e9e/distro-1.9.0-py3-none-any.whl.metadata
  Downloading distro-1.9.0-py3-none-any.whl.metadata (6.8 kB)
Collecting dpath<3,>=2.1.0 (from dvc)
  Obtaining dependency information for dpath<3,>=2.1.0 from https://files.pythonhosted.org/packages/05/d1/8952806fbf9583004ab479d8f58a9496c3d35f6b6009ddd458bdd9978eaf/dpath-2.2.0-py3-none-any.whl.metadata
  Downloading dpath-2.2.0-py3-none-any.whl (11 kB)
Created wheel for antlr4-python3-runtime: filename=antlr4_python3_runtime-4.9.3-py3-none-any.whl size=144577 sha256=63d7ff4aca2013bf0ba063ac31b3732af5a6587a09d39b1fa09e6064a06c97df
Stored in directory: c:\users\hp\appdata\local\pip\cache\wheels\1a\97\32\461f837398029ad76911109f07047fde1d7b661a147c7c56d1
Successfully built antlr4-python3-runtime
Installing collected packages: pygtrie, funcy, dictdiffer, appdirs, antlr4-python3-runtime, zc.lockfile, wcwidth, voluptuous, vine, tomkit, shtab, shortuuid, shellingham, semver, ruamel.yaml.clib, python-dotenv, pydot, pycparser, propcache, platformdirs, pathspec, orjson, omegaconf, multidict, grandalf, frozendict, flatten_dict, exceptiongroup, entrypoints, dvc-render, dvc-objects, dulwich, dpath, distro, diskcache, configobj, billiard, atpublic, aiohappyeyeballs, yarl, sqlalchemy, ruamel.yaml, prompt-toolkit, iterative-telemetry, hydra-core, flupy, lock, dvc-studio-client, click-plugins, click-didyoumean, cffi, amqp, aiosignal, typer, pygit2, pydantic-settings, kombu, dvc-data, cryptography, click-repl, aiohttp, celery, asynssh, aiohttp-retry, scmrepo, dvc-task, dvc-http, gto, dvc
Successfully installed aiohappyeyeballs-2.6.1 aiohttp-3.13.2 aiohttp-retry-2.9.1 aiosignal-1.4.0 amqp-5.3.1 antlr4-python3-runtime-4.9.3 appdirs-1.4.4 asynssh-2.22.0 atpublic-7.0.0 billiard-4.2.4 celery-5.6.0 cffi-2.0.0 click-didyoumean-0.3.1 click-plugins-1.1.1.2 click-repl-0.3.0 configobj-5.0.9 cryptography-46.0.3 dictdiffer-0.9.0 diskcache-5.6.3 distro-1.9.0 dpath-2.2.0 dulwich-0.25.0 dvc-3.65.0 dvc-data-3.17.0 dvc-http-2.32.0 dvc-objects-5.2.0 dvc-render-1.0.2 dvc-studio-client-0.22.0 dvc-task-0.40.2 entrypoints-0.4 exceptiongroup-1.3.1 flatten_dict-0.4.2 flupy-0.4.2 frozendict-1.8.0 funcy-2.0 grandalf-0.8 gto-1.9.0 hydra-core-1.3.2 iterative-telemetry-0.0.10 kombu-5.6.1 multidict-6.7.0 omegaconf-2.3.0 orjson-3.11.5 pathspec-0.12.1 platformdirs-4.5.1 prompt-toolkit-3.0.52 propcache-0.4.1 pycparser-2.23 pydantic-settings-2.12.0 pydot-4.0.1 pygit2-1.19.0 pygtrie-2.5.0 python-dotenv-1.2.1 ruamel.yaml-0.18.17 ruamel.yaml.clib-0.2.15 scmrepo-3.6.0 semver-3.0.4 shellingham-1.5.4 shortuuid-1.0.13 shtab-1.8.0 sqlalchemy-0.11.2 tomkit-0.13.3 typer-0.21.0 vine-5.1.0 voluptuous-0.16.0 wcwidth-0.2.14 yarl-1.22.0 zc.lockfile-4.0

[notice] A new release of pip is available: 23.2.1 -> 25.3
[notice] To update, run: python.exe -m pip install --upgrade pip

C:\Users\hp>dvc --version
3.65.0

C:\Users\hp>
```

## Étape 1 : Initialisation de DVC dans le projet

### 1. Ouvre un terminal dans le dossier du projet :

Nous avons commencé par ouvrir un terminal et nous positionner dans le répertoire du projet *mlops-lab-01*, afin d'exécuter les commandes DVC directement à la racine du projet.

```
PS C:\Users\hp> cd "C:\Users\hp\OneDrive\Bureau\TP MLOps\mlops-lab-01"
PS C:\Users\hp\OneDrive\Bureau\TP MLOps\mlops-lab-01>
```

### 2. Lance l'initialisation de DVC :

L'outil DVC a été initialisé à l'aide de la commande `dvc init`, ce qui a permis de configurer le projet pour le versionnement des données et des artefacts de Machine Learning en complément de Git.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc init
Initialized DVC repository.

You can now commit the changes to git.

+-----+
|               DVC has enabled anonymous aggregate usage analytics.               |
| Read the analytics documentation (and how to opt-out) here:                     |
| <https://dvc.org/doc/user-guide/analytics>                                         |
+-----+

What's next?
-----
- Check out the documentation: <https://dvc.org/doc>
- Get help and share ideas: <https://dvc.org/chat>
- Star us on GitHub: <https://github.com/treeverse/dvc>
```

### 3. Vérifie les fichiers créés :

Après l'initialisation, le répertoire `.dvc/` ainsi que le fichier de configuration `.dvc/config` ont été automatiquement créés, confirmant que DVC est correctement intégré au projet.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dir

Répertoire : C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01

Mode                LastWriteTime         Length Name
----                -
d----- 27/12/2025 10:12                .dvc
da---l  14/12/2025 17:40                data
da---l  14/12/2025 17:56                logs
da---l  14/12/2025 18:08                models
da---l  14/12/2025 17:44                registry
da---l  14/12/2025 19:44                src
da---l  14/12/2025 16:40                venv_mlops
-a----  27/12/2025 10:12             142 .dvcignore
-a---l  14/12/2025 18:50              55 .gitignore
-a---l  14/12/2025 18:19          1883844 Lab 1 Du notebook au mini-système production-ready - Wassima RAICHI.pdf
-a---l  14/12/2025 19:47          1389609 Lab 2 Code Source management - Wassima RAICHI.pdf
-a---l  14/12/2025 18:38              659 README.md
```

## Étape 2 — Versionner les données brutes avec DVC

### 1. Ajoute le dataset au suivi DVC :

Avant d'ajouter les données au suivi DVC, le répertoire data/ a été retiré du fichier .gitignore afin de permettre à DVC de gérer explicitement le versionnement des données brutes sans que celles-ci soient directement suivies par Git.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc add data/raw.csv  
ERROR: bad DVC file name 'data\raw.csv.dvc' is git-ignored.
```

```
venv_mlops/  
__pycache__/  
*.pyc  
logs/  
models/  
data/
```

```
venv_mlops/  
__pycache__/  
*.pyc  
logs/  
models/  
|
```

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc add data/raw.csv
100% Adding...|1/1 [00:00, 1.49file/s]

To track the changes with git, run:

    git add 'data/.gitignore' 'data/raw.csv.dvc'

To enable auto staging, run:

    dvc config core.autostage true
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01>
```

## 2. Cette commande crée automatiquement :

Le fichier de données brutes data/raw.csv, généré lors du lab précédent, a été ajouté au suivi de DVC à l'aide de la commande `dvc add data/raw.csv`, permettant à DVC de prendre en charge le versionnement de ce fichier volumineux.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dir data

Répertoire : C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01\data

Mode                LastWriteTime         Length Name
----                -
-a----           27/12/2025   10:33             10 .gitignore
-a---l           14/12/2025   17:40        28240 processed.csv
-a----           27/12/2025   10:33        27933 raw.csv
-a----           27/12/2025   10:33          93 raw.csv.dvc

PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

L'exécution de la commande `dvc add` a entraîné la création automatique du fichier de métadonnées data/raw.csv.dvc, ainsi que du fichier data/.gitignore, assurant que le fichier de données réel ne soit plus versionné directement par Git.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> type data\raw.csv.dvc
outs:
- md5: 7e72dd6dd1b8e9f60c239ca5079b4267
  size: 27933
  hash: md5
  path: raw.csv
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

Le contenu du fichier data/raw.csv.dvc a été inspecté afin de vérifier la présence des informations de traçabilité telles que le hash, la taille du fichier et le chemin des données, confirmant que le dataset est désormais correctement suivi par DVC.

### 3. Ajoute au versionnement Git :

Les fichiers de description générés par DVC (data/raw.csv.dvc, data/.gitignore et .gitignore) ont été ajoutés au versionnement Git et validés par un commit, garantissant une traçabilité complète du dataset brut sans alourdir le dépôt Git.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> git status
On branch main
Your branch is up to date with 'origin/main'.

Changes to be committed:
  (use "git restore --staged <file>..." to unstage)
    new file:   .dvc/.gitignore
    new file:   .dvc/config
    new file:   .dvcignore

Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git restore <file>..." to discard changes in working directory)
    modified:   .gitignore

Untracked files:
  (use "git add <file>..." to include in what will be committed)
    data/

PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> git add data/*.dvc .gitignore
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> git commit -m "data: suivi du dataset brut via DVC"
[main ecf2101] data: suivi du dataset brut via DVC
5 files changed, 11 insertions(+), 1 deletion(-)
create mode 100644 .dvc/.gitignore
create mode 100644 .dvc/config
create mode 100644 .dvcignore
create mode 100644 data/raw.csv.dvc
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

## Étape 3 : Configuration d'un remote DVC

### 1. Crée un dossier de remote :

Un dossier local nommé dvc\_storage a été créé à la racine du projet afin de servir de stockage distant (remote) pour les données et modèles versionnés par DVC

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> mkdir dvc_storage
mkdir : Il existe déjà un élément avec le nom spécifié C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01\dvc_storage.
Au caractère Ligne:1 : 1
+ mkdir dvc_storage
+ ~~~~~
+ CategoryInfo          : ResourceExists: (C:\Users\hp\One...-01\dvc_storage:String) [New-Item], IOException
+ FullyQualifiedErrorId : DirectoryExist,Microsoft.PowerShell.Commands.NewItemCommand

PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dir

Répertoire : C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01

Mode                LastWriteTime         Length Name
----                -
da---l          27/12/2025    10:33             .dvc
da---l          27/12/2025    12:03          .github
da---l          04/01/2026    23:33          data
da---l          27/12/2025    10:45       dvc_storage
da---l          14/12/2025    17:56          logs
da---l          27/12/2025    11:42          models
da---l          27/12/2025    11:42       registry
da---l          27/12/2025    11:42       reports
da---l          14/12/2025    19:44          src
da---l          14/12/2025    16:40       venv_mlops
-a---l          27/12/2025    10:12         142 .dvcignore
-a---l          27/12/2025    11:25         59 .gitignore
-a---l          27/12/2025    14:46         464 Dockerfile.txt.txt
-a---l          27/12/2025    11:42        1464 dvc.lock
-a---l          27/12/2025    11:26         480 dvc.yaml
-a---l          14/12/2025    18:19    1883844 Lab 1 Du notebook au mini-système production-ready - Wassima RAICHI.pdf
-a---l          14/12/2025    19:47    1389609 Lab 2 Code Source management - Wassima RAICHI.pdf
-a---l          14/12/2025    18:38         659 README.md
-a---l          27/12/2025    14:36         98 requirements.txt
```

## 2. Déclare-le comme remote principal :

Le dossier `dvc_storage` a été déclaré comme remote principal de DVC à l'aide de la commande `dvc remote add -d localremote dvc_storage`, permettant à DVC de stocker les artefacts volumineux en dehors du dépôt Git.

La commande `dvc remote list` a permis de vérifier que le remote `localremote` est bien configuré et défini comme remote par défaut du projet.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc remote add -d localremote dvc_storage
Setting 'localremote' as a default remote.
ERROR: configuration error - config file error: remote 'localremote' already exists. Use '-f|--force' to overwrite it.
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc remote list
localremote      C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01\dvc_storage (default)
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc remote default
localremote
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

## 3. Versionne la config

Le fichier de configuration `.dvc/config`, contenant la définition du remote DVC, a été ajouté au versionnement Git afin d'assurer la reproductibilité de la configuration DVC sur d'autres environnements.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> git add .dvc/config
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> type .dvc\config
[core]
    remote = localremote
['remote "localremote"']
    url = ../dvc_storage
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```



La configuration du remote DVC a été validée par un commit Git dédié, garantissant que l'ensemble de l'équipe peut retrouver et réutiliser la même configuration de stockage des données.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> git commit -m "dvc: configuration du remote local"
On branch main
Your branch is up to date with 'origin/main'.

Untracked files:
  (use "git add <file>..." to include in what will be committed)
        Dockerfile.txt.txt
        requirements.txt

nothing added to commit but untracked files present (use "git add" to track)
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

## Étape 4 : Push des données dans le remote DVC

Cette étape a pour objectif de transférer les données versionnées vers le remote DVC afin de les rendre persistantes, partageables et récupérables par tous les utilisateurs du projet, indépendamment du dépôt Git.

### 1. Envoie les données

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc push
Collecting |7.00 [00:00, 131entry/s]
Pushing
3 files pushed
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

### 2. Vérifie que le dossier dvc\_storage/ contient des fichiers de hash.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dir dvc_storage

Répertoire : C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01\dvc_storage

Mode                LastWriteTime         Length Name
----                -
da---l            27/12/2025      10:45         files

PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

## Étape 5 : imulation d'une collaboration : supprimer localement et récupérer depuis DVC

Cette étape a pour objectif de démontrer le rôle de DVC dans un contexte collaboratif en permettant de supprimer localement un dataset puis de le restaurer à l'identique depuis le remote, garantissant ainsi la reproductibilité et le partage des données.



### 1. Supprime le dataset local :

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> del data\raw.csv
```

### 2. Vérifie qu'il n'y a plus rien :

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dir data

Répertoire : C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01\data

Mode                LastWriteTime         Length Name
----                -
-a---l            27/12/2025      11:23             26 .gitignore
-a---l            27/12/2025      11:42          28240 processed.csv
-a----            04/01/2026      23:33             93 raw.csv.dvc

PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc status
prepare:
  changed deps:
    deleted:          data\raw.csv
data\raw.csv.dvc:
  changed outs:
    deleted:          data\raw.csv
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

### 3. Récupère le dataset via DVC :

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc pull
Collecting                                     |7.00 [00:00, 389entry/s]
Fetching
Building workspace index                       |10.0 [00:00, 1.11kentry/s]
Comparing indexes                             |11.0 [00:00, 1.99kentry/s]
Applying changes                              |1.00 [00:00, 249file/s]
A      data\raw.csv
1 file added
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dir data

Répertoire : C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01\data

Mode                LastWriteTime         Length Name
----                -
-a---l            27/12/2025      11:23             26 .gitignore
-a---l            27/12/2025      11:42          28240 processed.csv
-a----            04/01/2026      23:51          27933 raw.csv
-a----            04/01/2026      23:33             93 raw.csv.dvc

PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

## Étape 6 : Création d'un pipeline reproductible dvc.yaml

Cette étape a pour objectif de définir un pipeline de Machine Learning entièrement reproductible avec DVC, en formalisant les étapes de préparation, d'entraînement et d'évaluation, ce qui permet d'enregistrer automatiquement les dépendances, les sorties et les commandes dans le fichier dvc.yaml et d'assurer la reproductibilité complète du processus.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc stage add --force -n prepare -d src/prepare_data.py -d data/raw.csv -o data/processed.csv -o registry/train_stats.json python src/prepare_data.py
Added stage 'prepare' in 'dvc.yaml'
```

To track the changes with git, run:

```
git add dvc.yaml
```

To enable auto staging, run:

```
dvc config core.autostage true
```

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> type dvc.yaml
stages:
```

```
  prepare:
```

```
    cmd: python src/prepare_data.py
```

```
    deps:
```

```
      - data/raw.csv
```

```
      - src/prepare_data.py
```

```
    outs:
```

```
      - data/processed.csv
```

```
      - registry/train_stats.json
```

```
  train:
```

```
    cmd: python src/train.py
```

```
    deps:
```

```
      - data/processed.csv
```

```
      - src/train.py
```

```
    outs:
```

```
      - models
```

```
  evaluate:
```

```
    cmd: python src/evaluate.py
```

```
    deps:
```

```
      - data/processed.csv
```

```
      - models/model.joblib
```

```
      - src/evaluate.py
```

```
    outs:
```

```
      - reports/metrics.json
```

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc stage add --force -n train -d src/train.py -d data/processed.csv -o models/model.joblib python src/train.py
Added stage 'train' in 'dvc.yaml'
```

To track the changes with git, run:

```
git add dvc.yaml
```

To enable auto staging, run:

```
dvc config core.autostage true
```

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> dvc stage add --force -n evaluate -d src/evaluate.py -d models/model.joblib -d data/processed.csv -o reports/metrics.json python src/evaluate.py
Added stage 'evaluate' in 'dvc.yaml'
```

To track the changes with git, run:

```
git add dvc.yaml
```

To enable auto staging, run:

```
dvc config core.autostage true
```

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> git add dvc.yaml registry/.gitignore reports/.gitignore
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> git commit -m "pipeline: ajout des étapes prepare/train/evaluate"
[main fc547d3] pipeline: ajout des étapes prepare/train/evaluate
 1 file changed, 1 insertion(+), 1 deletion(-)
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```

## Étape 7 : Reproduire automatiquement tout le pipeline

Cette étape a pour objectif de démontrer la reproductibilité du pipeline DVC en relançant automatiquement uniquement les étapes impactées par une modification, tout en générant le fichier dvc.lock qui fige l'état exact du pipeline et garantit sa traçabilité et sa reproductibilité dans le temps.

```
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> git add dvc.lock
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> git commit -m "pipeline: lock after repro"
[main fe5cfa4] pipeline: lock after repro
1 file changed, 7 insertions(+), 8 deletions(-)
PS C:\Users\hp\OneDrive\Bureau\TP MLops\mlops-lab-01> |
```