Candra Kurniawan Mentoring 1 - Machine Learning Process
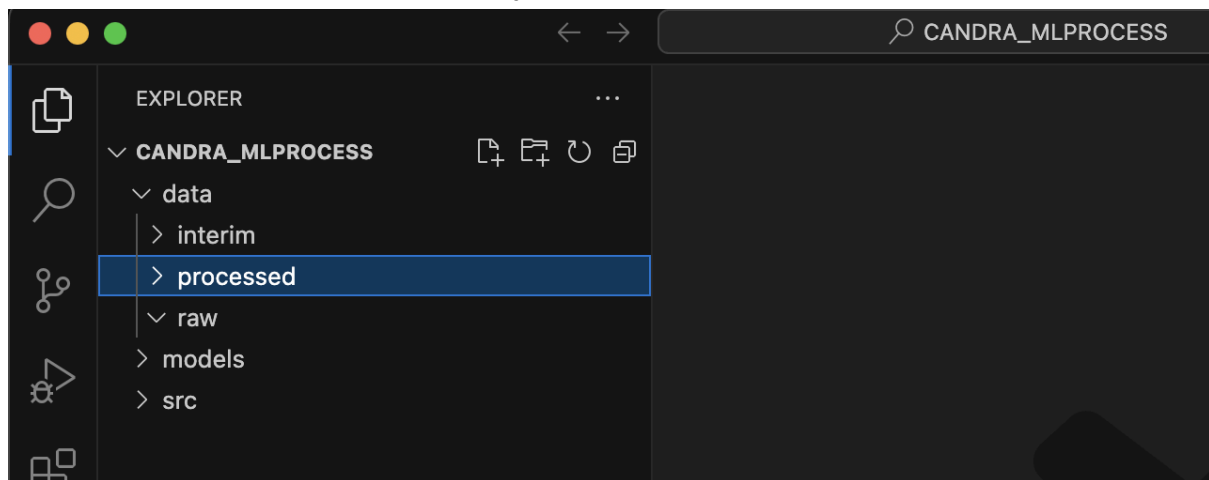
Job Preparation Program, Data Science, Pacmann AI
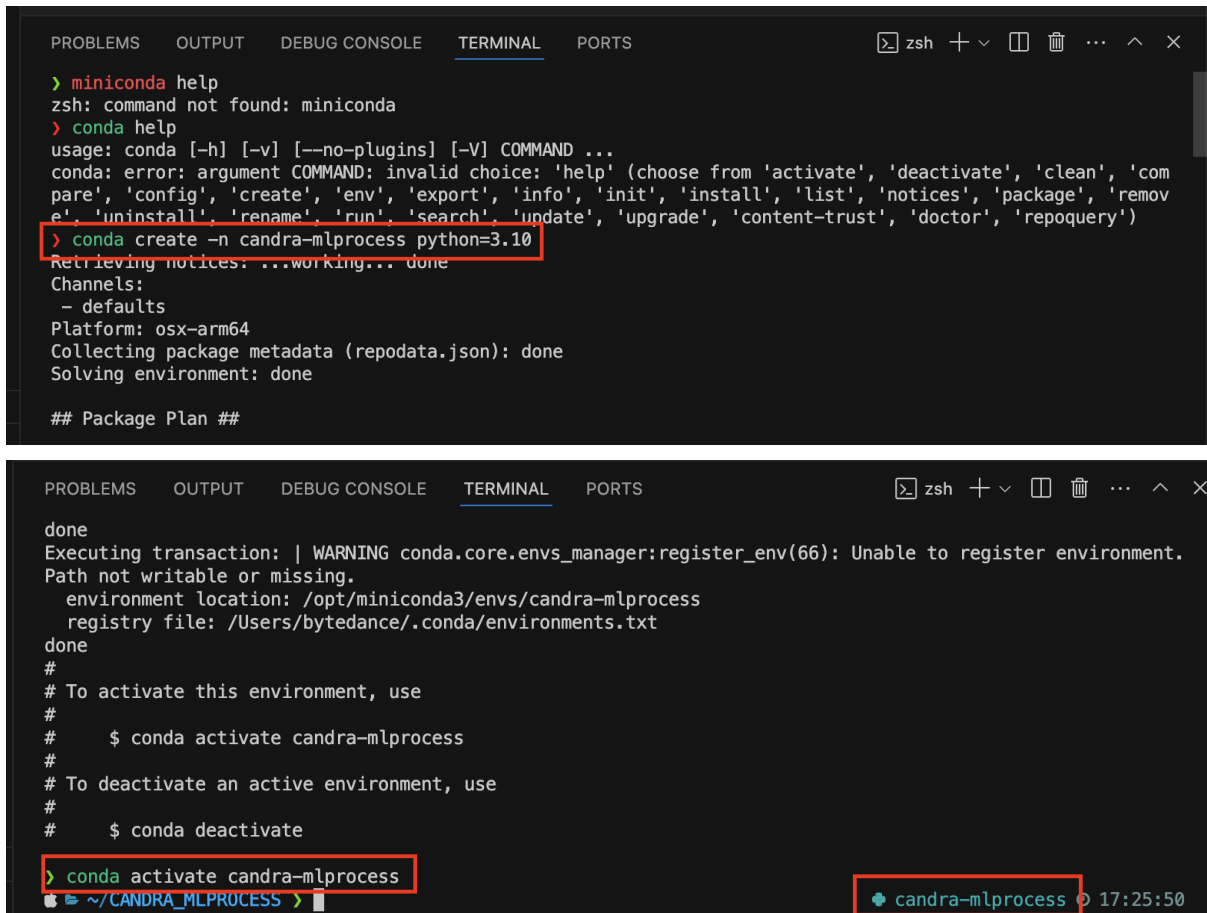
Catatan:
- Make a copy docs ini sebelum menjawab.
- Studi kasus pada soal ini akan **digunakan** untuk mentoring selanjutnya.
- Tulis jawaban Anda dalam word processor dan ekspor ke PDF untuk soal nomor 1 dan 2
- Tulis jawaban Anda dalam jupyter notebook dan python script untuk soal nomor 3 dan 4
- Archive file PDF dan keseluruhan root folder project Anda ke ZIP, **kecuali folder venv**
- Beri nama `[NAMA_LENGKAP]_MLPROCESS_1` file archive Anda
- Submit file archive Anda melalui link submission yang telah disediakan

Soal

1. [25 points] Project Environment Setup
   a. [5 points] Buat folder project

b. **[5 points]** Buat python virtual environment (venv)

c. [5 points] Update pip

```
  --use-deprecated <feature>  Enable deprecated functionality, that will be removed in the future.
> python --version
Python 3.10.14
> python3 -m pip install --upgrade pip
Requirement already satisfied: pip in /Library/Frameworks/Python.framework/Versions/3.11/lib/python3.11/sit
e-packages (23.2.1)
Collecting pip
  Obtaining dependency information for pip from https://files.pythonhosted.org/packages/8a/6a/19e9fe04fca05
9ccf770861c7d5721ab4c2aebc539889e97c7977528a53b/pip-24.0-py3-none-any.whl.metadata
  Downloading pip-24.0-py3-none-any.whl.metadata (3.6 kB)
Downloading pip-24.0-py3-none-any.whl (2.1 MB)
                                              2.1/2.1 MB 6.1 MB/s eta 0:00:00
Installing collected packages: pip
  Attempting uninstall: pip
    Found existing installation: pip 23.2.1
    Uninstalling pip-23.2.1:
      Successfully uninstalled pip-23.2.1
```

```
  Attempting uninstall: pip
    Found existing installation: pip 23.2.1
    Uninstalling pip-23.2.1:
      Successfully uninstalled pip-23.2.1
  WARNING: The scripts pip, pip3 and pip3.11 are installed in '/Library/Frameworks/Python.framework/Version
s/3.11/bin' which is not on PATH.
  Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-l
ocation.
Successfully installed pip-24.0
> python --version
Python 3.10.14
> python3 --version
Python 3.11.6
> python -m pip install --upgrade pip
Requirement already satisfied: pip in /opt/miniconda3/envs/candra-mlprocess/lib/python3.10/site-packages (2
4.0)
  ~/CANDRA_MLPROCESS >                                              candra-mlprocess  17:30:56
```

d. [5 points] Install dependencies
- Install package berikut:
  - ○ pandas
  - ○ scikit-learn
  - ○ imblearn
  - ○ joblib
  - ○ numpy
  - ○ scipy
  - ○ seaborn
  - ○ fastapi
  - ○ jupyterlab
  - ○ Requests

e. [5 points] Non aktifkan venv Anda

2. [25 points] Rangkum permasalahan bisnis
Rangkum mockup interview antara user dan DS di bagian akhir dari dokumen ini
- **Rangkuman Anda harus bisa menjawab pertanyaan berikut:**

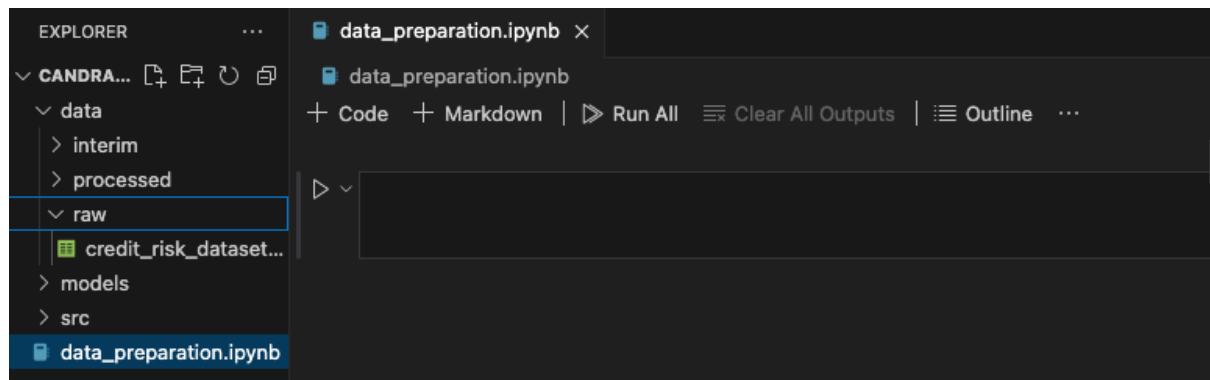| PERTANYAAN	JAWABAN
Latar belakang permasalahan bisnis	Terdapat masalah dalam peminjaman Uang, yaitu banyaknya nasabah yang menjadi peminjam menunggak-nunggak tidak membayar sehingga membuat masalah pada Perusahaan Loan ini
Objektif bisnis yang ingin dicapai	Ingin menurunkan Jumlah NPL (Non-Performing Loan) dan ingin mempunyai sistem yang mampu mendeteksi pinjaman yang beresiko sehingga dapat pihak Bank cegah agar tidak menunggak-nunggak pinjamannya.
Metrik pengukuran bisnis yang akan digunakan	*Turunnya jumlah Nasabah yang merupakan Non-Performing Loan dalam Percent(%)*
Kandidat solusi machine learning yang akan dibangun	*Membangun sistem Machine Learning Klasifikasi Binary*
Metrik pengukuran machine learning yang akan digunakan	*F1-score dengan tingkat prediksi > 95%*

3. [25 points] Persiapan Dataset
    a. [1 points] Buat satu file bernama `data_preparation.ipynb` dan letakkan pada root folder project Anda
    b. [1 points] Letakkan file CSV dataset yang bernama `credit_risk_dataset.csv` ke folder `raw`
        Sertakan screenshot hasil dari pemindahan file tersebut.

```
> conda deactivate
> cd Downloads
cd: no such file or directory: Downloads
> cd ../Downloads
> cp credit_risk_dataset.csv ../CANDRA_MLPROCESS
> cd ../CANDRA_MLPROCESS
> mv credit_risk_dataset.csv data/raw
  ~/CANDRA_MLPROCESS >                                          base  09:57:42
```

Note:

- **Jawaban selanjutnya berada di file Notebook ya kak, terimakasih..**