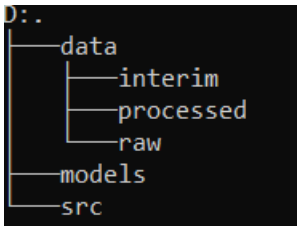
Screenshot folder



screenshot terminal yang menampilkan command saat Anda membuat venv tersebut

```
PS D:\pacmann_AIML\ml_process\task_1\Wahyunan_Andika_MLPROCESS\data> python -m venv WahyunanAndika_VENV
```

screenshot terminal yang menampilkan command saat Anda mengaktifkan venv tersebut

```
PS D:\pacmann_AIML\ml_process\task_1\Wahyunan_Andika_MLPROCESS\data> .\WahyunanAndika_VENV\Scripts\Activate
>>
(WahyunanAndika_VENV) PS D:\pacmann_AIML\ml_process\task_1\Wahyunan_Andika_MLPROCESS\data>
```

screenshot terminal yang menampilkan command saat Anda melakukan update PIP

```
(WahyunanAndika_VENV) PS D:\pacmann_AIML\ml_process\task_1\Wahyunan_Andika_MLPROCESS\data> python -m pip install --upgrade pip
Requirement already satisfied: pip in d:\pacmann_aiml\ml_process\task_1\wahyunan_andika_mlprocess\data\wahyunanandika_venv\lib\site-packages (24.2)
(WahyunanAndika_VENV) PS D:\pacmann_AIML\ml_process\task_1\Wahyunan_Andika_MLPROCESS\data>
```

screenshot terminal yang menampilkan command saat Anda melakukan pemasangan packages tersebut

```
PS D:\pacmann_AIML\ml_process\task_1\Wahyunan_Andika_MLPROCESS\data> .\WahyunanAndika_VENV\Scripts\Activate
>>
(WahyunanAndika_VENV) PS D:\pacmann_AIML\ml_process\task_1\Wahyunan_Andika_MLPROCESS\data> python -m pip install --upgrade pip
Requirement already satisfied: pip in d:\pacmann_aiml\ml_process\task_1\wahyunan_andika_mlprocess\data\wahyunanandika_venv\lib\site-packages (24.2)
(WahyunanAndika_VENV) PS D:\pacmann_AIML\ml_process\task_1\Wahyunan_Andika_MLPROCESS\data> pip install pandas scikit-learn imblearn joblib numpy scipy seaborn
 fastapi jupyterlab requests
>>
Collecting pandas
  Downloading pandas-2.2.3-cp312-cp312-win_amd64.whl.metadata (19 kB)
Collecting scikit-learn
  Downloading scikit_learn-1.5.2-cp312-cp312-win_amd64.whl.metadata (13 kB)
Collecting imblearn
  Downloading imblearn-0.0-py2.py3-none-any.whl.metadata (355 bytes)
Collecting joblib
  Using cached joblib-1.4.2-py3-none-any.whl.metadata (5.4 kB)
```

screenshot terminal yang menampilkan command saat Anda menonaktifkan venv

```
Successfully installed MarkupSafe-2.1.5 annotated-types-0.7.0 anyio-4.6.0 argon2-cffi-23.1.0 argon2-cffi-bindings-21.2.0 arrow-1.3.0 asttokens-2.4.1 async-lru
-2.0.4 attrs-24.2.0 babel-2.16.0 beautifulsoup4-4.12.3 bleach-6.1.0 certifi-2024.8.30 cffi-1.17.1 charset-normalizer-3.3.2 colorama-0.4.6 comm-0.2.2 contourpy
-1.3.0 cycler-0.12.1 debugpy-1.8.6 decorator-5.1.1 defusedxml-0.7.1 executing-2.1.0 fastapi-0.115.0 fastjsonschema-2.20.0 fonttools-4.54.1 fqdn-1.5.1 h11-0.14
.0 httpcore-1.0.5 httpx-0.27.2 idna-3.10 imbalanced-learn-0.12.3 imblearn-0.0 ipykernel-6.29.5 ipython-8.27.0 isoduration-20.11.0 jedi-0.19.1 jinja2-3.1.4 job
lib-1.4.2 json5-0.9.25 jsonpointer-3.0.0 jsonschema-4.23.0 jsonschema-specifications-2023.12.1 jupyter-client-8.6.3 jupyter-core-5.7.2 jupyter-events-0.10.0 j
upyter-lsp-2.2.5 jupyter-server-2.14.2 jupyter-server-terminals-0.5.3 jupyterlab-4.2.5 jupyterlab-pygments-0.3.0 jupyterlab-server-2.27.3 kiwisolver-1.4.7 mat
plotlib-3.9.2 matplotlib-inline-0.1.7 mistune-3.0.2 nbclient-0.10.0 nbconvert-7.16.4 nbformat-5.10.4 nest-asyncio-1.6.0 notebook-shim-0.2.4 numpy-2.1.1 overri
des-7.7.0 packaging-24.1 pandas-2.2.3 pandocfilters-1.5.1 parso-0.8.4 pillow-10.4.0 platformdirs-4.3.6 prometheus-client-0.21.0 prompt-toolkit-3.0.48 psutil-6
.0.0 pure-eval-0.2.3 pycparser-2.22 pydantic-2.9.2 pydantic-core-2.23.4 pygments-2.18.0 pyparsing-3.1.4 python-dateutil-2.9.0.post0 python-json-logger-2.0.7 p
ytz-2024.2 pywin32-306 pywinpty-2.0.13 pyyaml-6.0.2 pyzmq-26.2.0 referencing-0.35.1 requests-2.32.3 rfc3339-validator-0.1.4 rfc3986-validator-0.1.1 rpds-py-0.
20.0 scikit-learn-1.5.2 scipy-1.14.1 seaborn-0.13.2 send2trash-1.8.3 setuptools-75.1.0 six-1.16.0 sniffio-1.3.1 soupsieve-2.6 stack-data-0.6.3 starlette-0.38.
6 terminado-0.18.1 threadpoolctl-3.5.0 tinycss2-1.3.0 tornado-6.4.1 traitlets-5.14.3 types-python-dateutil-2.9.0.20240906 typing-extensions-4.12.2 tzdata-2024
.2 uri-template-1.3.0 urllib3-2.2.3 wcwidth-0.2.13 webcolors-24.8.0 webencodings-0.5.1 websocket-client-1.8.0
(WahyunanAndika_VENV) PS D:\pacmann_AIML\ml_process\task_1\Wahyunan_Andika_MLPROCESS\data> deactivate
>>
PS D:\pacmann_AIML\ml_process\task_1\Wahyunan_Andika_MLPROCESS\data>
```

Summarize

Business problem background:

The bank is facing challenges with credit risk and wants to predict customers who might default on their loans.

Business objective:

Reduce Non-Performing Loans (NPL) and detect risky loans early to mitigate potential financial losses.

Business metric to measure success:

Key metrics include reducing the NPL ratio, minimizing defaults, and increasing the accuracy of loan default predictions.

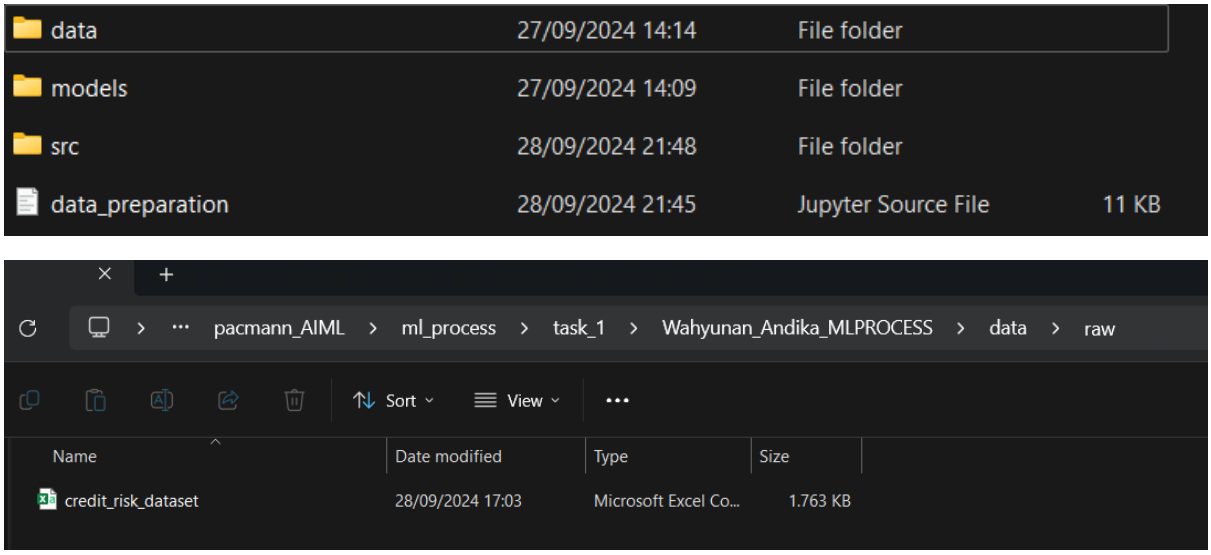
Candidate Machine Learning solution:

A machine learning model that predicts whether a customer will default or not, using historical loan data.

Machine Learning metric to measure success:

Evaluation metrics such as Accuracy, Precision, Recall, and F1-score will be used to assess the model's performance.

screenshot hasil dari pemindahan file tersebut



screenshot hasil dari fungsi yang dijalankan

```
print(data.head())
```

```
   person_age  person_income person_home_ownership  person_emp_length  \
0          22          59000                  RENT              123.0
1          21           9600                   OWN                5.0
2          25           9600              MORTGAGE                1.0
3          23          65500                  RENT                4.0
4          24          54400                  RENT                8.0

  loan_intent loan_grade  loan_amnt  loan_int_rate  loan_status  \
0    PERSONAL          D      35000          16.02            1
1   EDUCATION          B       1000          11.14            0
2     MEDICAL          C       5500          12.87            1
3     MEDICAL          C      35000          15.23            1
4     MEDICAL          C      35000          14.27            1

   loan_percent_income cb_person_default_on_file  cb_person_cred_hist_length
0                 0.59                       Y                           3
1                 0.10                       N                           2
2                 0.57                       N                           3
3                 0.53                       N                           2
4                 0.55                       Y                           4
```

screenshot hasil dari fungsi yang dijalankan

```
Original data shape: (32581, 12)
X data shape: (32581, 11)
y data shape: (32581,)
   person_age  person_income person_home_ownership  person_emp_length  \
0          22          59000                  RENT              123.0
1          21           9600                   OWN                5.0
2          25           9600              MORTGAGE                1.0
3          23          65500                  RENT                4.0
4          24          54400                  RENT                8.0

  loan_intent loan_grade  loan_amnt  loan_int_rate  loan_percent_income  \
0    PERSONAL          D      35000          16.02                 0.59
1   EDUCATION          B       1000          11.14                 0.10
2     MEDICAL          C       5500          12.87                 0.57
3     MEDICAL          C      35000          15.23                 0.53
4     MEDICAL          C      35000          14.27                 0.55

  cb_person_default_on_file  cb_person_cred_hist_length
0                         Y                           3
1                         N                           2
2                         N                           3
3                         N                           2
4                         Y                           4
0    1
1    0
2    1
3    1
4    1
Name: loan_status, dtype: int64
```

screenshot hasil dari fungsi yang dijalankan

```
X_train, X_non_train, y_train, y_non_train = split_train_test(X, y, test_size=0.2, random_state=42)

X_valid, X_test, y_valid, y_test = split_train_test(X_non_train, y_non_train, test_size=0.5, random_state=42)
```

```
X train shape: (26064, 11)
X test shape: (6517, 11)
y train shape: (26064,)
y test shape: (6517,)
X train shape: (3258, 11)
X test shape: (3259, 11)
y train shape: (3258,)
y test shape: (3259,)
```

screenshot hasil dari fungsi yang dijalankan

```
serialize_data(X_train, "data/interim/X_train.pkl")
serialize_data(y_train, "data/interim/y_train.pkl")
serialize_data(X_test, "data/interim/X_test.pkl")
serialize_data(y_test, "data/interim/y_test.pkl")
serialize_data(X_valid, "data/interim/X_valid.pkl")
serialize_data(y_valid, "data/interim/y_valid.pkl")
```

✓ 0.0s

screenshot hasil dari fungsi yang dijalankan (sertakan screenshot terminal)