

Zaawansowane Programowanie Obiektowe - MLPipeFlow

Szymon Kołodziejski 84647

1 Opis

MLPipeFlow to framework do budowy pipeline'ów uczenia maszynowego, który upraszcza proces przetwarzania danych, inżynierii cech, trenowania modeli, ewaluacji, wdrażania i monitorowania. Jest zaprojektowany jako modułowy i rozszerzalny, co pozwala użytkownikom na łatwą integrację własnych komponentów.

2 Instalacja

Aby zainstalować wymagane zależności, uruchom:

```
pip install -r requirements.txt
```

3 Użycie

Aby uruchomić pipeline, utwórz instancję MLPipeFlow z wybranymi komponentami i wywołaj metodę run:

```
from src.preprocessing import DataPreprocessing
from src.feature_engineering import FeatureEngineering
from src.evaluation import ModelEvaluation
from src.training import ModelTraining
from src.deployment import ModelDeployment
from src.monitoring import Monitoring
from src.mlpipeline import MLPipeFlow

components = [
    DataPreprocessing(data_path='test_data/exampleData_TCGA_LUAD_2000.csv'),
    FeatureEngineering(target='class'),
    ModelTraining(model_name='svm'),
    ModelEvaluation(),
    ModelDeployment(model_path='model_example/model.pkl'),
    Monitoring()
]
```

```
pipeline = MLPipeFlow(components)
results = pipeline.run()
print(results)
```

4 Komponenty

4.1 DataPreprocessing

Obsługuje ładowanie i przetwarzanie danych.

- **Metody:**
 - `load_data()`: Ładuje dane z określonej ścieżki.
 - `execute()`: Wykonuje kroki przetwarzania danych.

4.2 FeatureEngineering

Obsługuje zadania inżynierii cech.

- **Metody:**
 - `execute(data)`: Wykonuje inżynierię cech na dostarczonych danych.

4.3 ModelTraining

Obsługuje trenowanie modelu.

- **Metody:**
 - `execute(data)`: Trenuje model używając dostarczonych danych.

4.4 ModelEvaluation

Obsługuje ewaluację modelu.

- **Metody:**
 - `execute(model, data)`: Ewaluacja modelu używając dostarczonych danych.

4.5 ModelDeployment

Obsługuje wdrażanie modelu.

- **Metody:**
 - `execute(model)`: Wdraża wytrenowany model.

4.6 Monitoring

Obsługuje monitorowanie wdrożonego modelu.

- **Metody:**
 - `execute(model_path, data)`: Monitoruje wdrożony model.

5 Technologie użyte

- Python
- scikit-learn
- pandas
- NumPy

6 Licencja

Ten projekt jest licencjonowany na warunkach licencji MIT.