MLflow 실습 4

1. Example code 살펴보기 (Automatic Logging)
2. Example code 살펴보기 (XGB Model)

1. Example code 살펴보기 (Automatic Logging)

- https://github.com/mlflow/mlflow/tree/master/examples/sklearn_autolog

```
wget https://raw.githubusercontent.com/mlflow/mlflow/master/examples/sklearn_autolog/utils.py
wget https://raw.githubusercontent.com/mlflow/mlflow/master/examples/sklearn_autolog/pipeline.py
```

- mlflow 에서 example 로 제공해주는 example 중 하나
  - 간단한 training data 를 가지고 **sklearn** 의 **Pipeline** 을 사용해, StandardScaler 전처리 이후 **LinearRegression** 을 수행하는 코드
- scikit-learn 과 같은 패키지는 mlflow 레벨에서 `autolog` 를 지원
  - model 의 parameters, metrics 와 model artifacts 를 사용자가 명시하지 않아도 **자동으로** mlflow 에 로깅

▼ 예시

Parameters

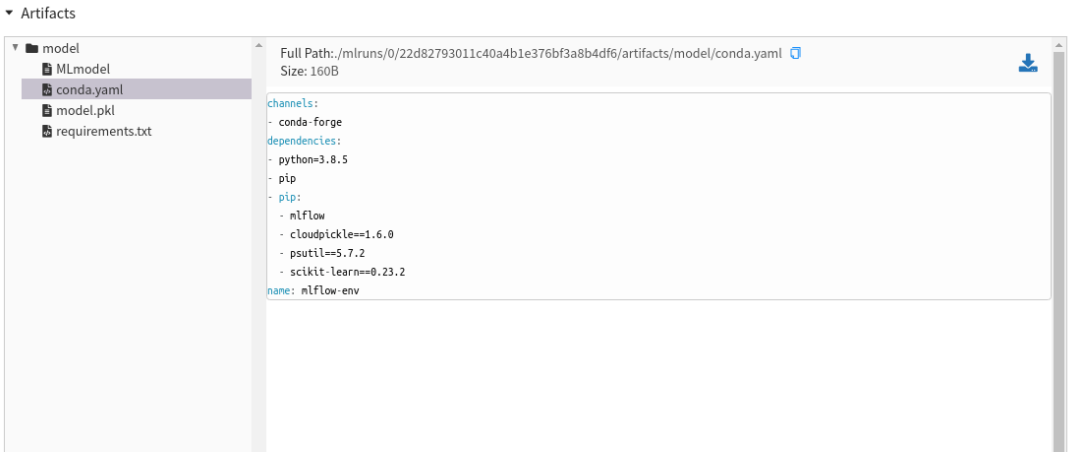
Name	Value
lr	LinearRegression()
lr__copy_X	True
lr__fit_intercept	True
lr__n_jobs	None
lr__normalize	False
memory	None
scaler	StandardScaler()
scaler__copy	True
scaler__with_mean	True
scaler__with_std	True
steps	[('scaler', StandardScaler()), ('lr', LinearRegression())]
verbose	False

Metrics

Name	Value
training_mae	2.220e-16
training_mse	1.972e-31
training_r2_score	1
training_rmse	4.441e-16
training_score	1

Tags

Name	Value	Actions
estimator_class	sklearn.pipeline.Pipeline	



---

2. Example code 살펴보기 (XGB Model)

- https://github.com/mlflow/mlflow/tree/master/examples/xgboost

```
wget https://raw.githubusercontent.com/mlflow/mlflow/master/examples/xgboost/train.py

# xgboost==1.4.2 설치 필요
```

- mlflow 에서 example 로 제공해주는 example 중 하나
  - **iris** data 를 가지고 **xgboost** 모델로 **classification** 을 수행하는 코드
- mlflow 에서 지원하는 xgboost 용 autolog 를 사용했고, 추가적인 custom metric 을 남기기 위해 `mlflow.log_metrics()` 사용

  ▼ 예시

▼ Parameters

Name	Value
colsample_bytree	1.0
early_stopping_rounds	None
eval_metric	mlogloss
learning_rate	0.3
maximize	None
num_boost_round	10
num_class	3
objective	multi:softprob
seed	42
subsample	1.0
verbose_eval	True

▼ Metrics

Name	Value
accuracy	1
log_loss	0.066
train-mlogloss	0.091

▸ Tags

▼ Artifacts