

Chapter. 01

머신러닝의 개요

I 학습 데이터, 예측데이터

FAST CAMPUS
ONLINE

직장인을 위한 파이썬 데이터분석

강사. 이경록

Chapter. 01

학습 데이터, 예측데이터 (features, labels, train, validation, test)

I feature, label

```
from sklearn.linear_model import LinearRegression

model = LinearRegression()
model.fit(x, y)
prediction = model.predict(x2)
```

X:

- features라고 불리웁니다.
- **x_train, x_test**
- 학습을 위한 데이터 세트. 예측할 값은 빠져있습니다.
- 예시) 지역, 평형 정보, 층수 정보, 동네, 거주민 평균 나이...

Y:

- labels라고 불리웁니다.
- **y_train, y_test**
- 예측해야할 값입니다. 예측 값만 존재합니다.
- 예시) 집 값

I train, test

학습을 위한 데이터
= Training Set

예측을 위한 데이터
= Test Set

- 모델이 학습하기 위해 필요한 데이터
- **feature/label**이 모두 존재

- 모델이 예측하기 위한 데이터
- **feature** 만 존재

I train, test

학습을 위한 데이터
= Training Set

예측을 위한 데이터
= Test Set

- 모델이 학습하기 위해 필요한 데이터
- feature/label이 모두 존재

X, Y
x_train, y_train

- 모델이 예측하기 위한 데이터
- feature 만 존재

X, Y
x_test, y_test

| x_train, y_train, x_test

```
from sklearn.linear_model import LinearRegression  
  
model = LinearRegression()  
model.fit(x_train, y_train)  
prediction = model.predict(x_test)
```