

```
In [ ]: from preamble import *
```

지도 학습

분류와 회귀

지도 학습 알고리즘

데이터셋 생성

```
In [ ]: # 데이터셋을 만듭니다
X, y = mglearn.datasets.make_forge()

# 산점도를 그립니다
mglearn.discrete_scatter(X[:, 0], X[:, 1], y)
plt.legend(["class 0", "class 1"], loc='best')
plt.xlabel("feature 1")
plt.ylabel("feature 2")
print("X.shape:", X.shape)
```

```
In [ ]: X, y = mglearn.datasets.make_wave(n_samples=40)
plt.plot(X, y, 'o')
plt.ylim(-3, 3)
plt.xlabel("feature")
plt.ylabel("target")
```

```
In [ ]: from sklearn.datasets import load_breast_cancer
cancer = load_breast_cancer()
print("cancer.keys():\n", cancer.keys())
```

```
In [ ]: print("유방암 데이터의 형태:", cancer['data'].shape)
```

```
In [ ]: print("종양 종류", cancer['target_names'])
```

```
In [ ]: print("종양", cancer['target'])
```

```
In [ ]: print("클래스별 샘플 갯수:\n",
      {n: v for n, v in zip(cancer['target_names'], np.bincount(cancer['target']))})
```

```
In [ ]: print("특성 이름:\n", cancer['feature_names'])
```

```
In [ ]: from sklearn.datasets import load_boston
boston = load_boston()
print("데이터의 형태:", boston.data.shape)
```

```
In [ ]: X, y = mglearn.datasets.load_extended_boston()
print("X.shape:", X.shape)
```

