

기계학습개론

- 머신러닝 개요 -

교수 이홍로

MP : 010-6611-3896

E-mail : hrlee@cnu.ac.kr

강의 홈페이지 : <https://cyber.hanbat.ac.kr/>



오늘의 강의 목표

- **Google Colab & Jupyter Notebook**
- **첫 머신러닝 프로그래밍**
 - K-평균 알고리즘의 간단한 이해
 - 실습 : 오류 제품 검출 또는 분류해 내기

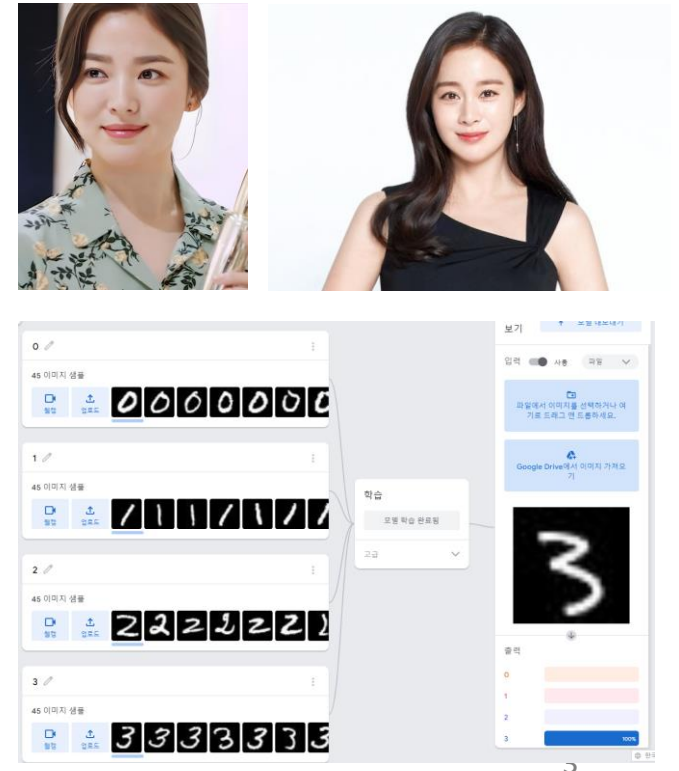
지난 주 강의 요약

■ 기계학습을 공부하기에 앞서

- 약인공지능, 강인공지능, 초인공지능
- 인공지능의 순기능 : 노동대체, 자동화, 개인 능력격차 완화
- 인공지능의 역기능 : Dumb AI, 법적 윤리적 책임소재, 빈부격차 심화, 예측불가
- 인공지능의 역사
- 프로그래밍 관점 : 기존 방식 vs 머신러닝
- 관련 용어 : 인공지능 > 머신러닝 > 딥러닝

■ 인공지능 관련 실습

- 분류 : 송혜교/김태희 분류, MNIST 이용 숫자분류



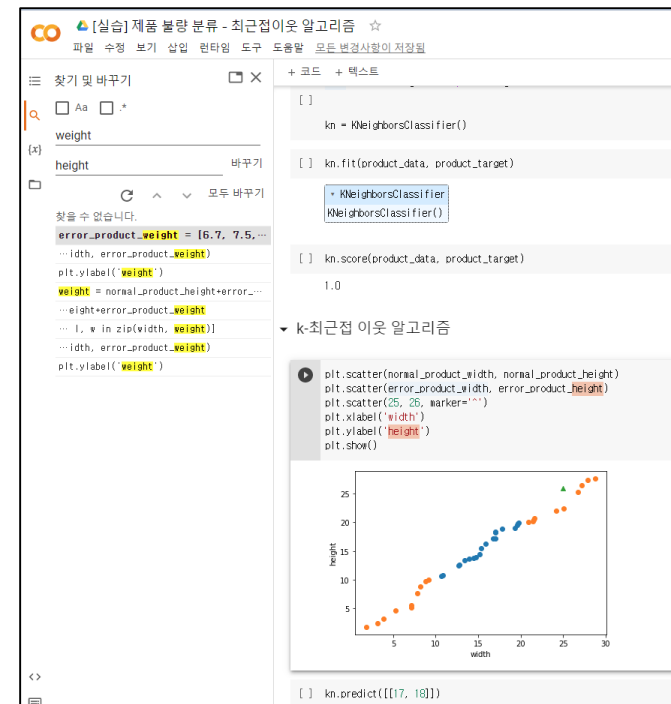
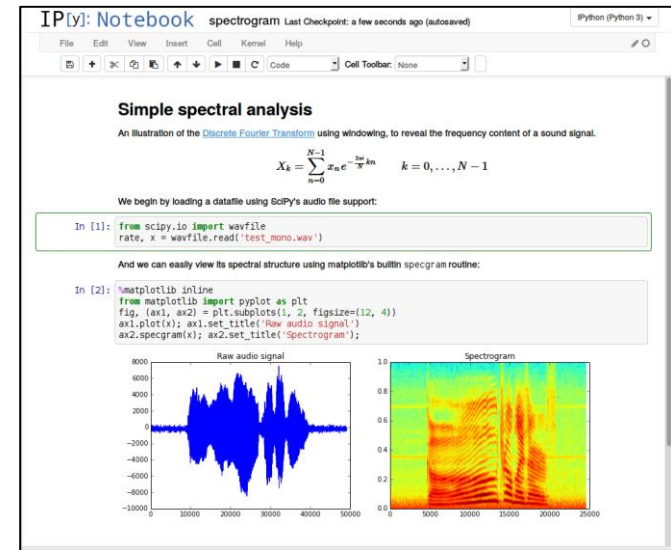
Google Colab & Jupyter Notebook

■ Jupyter Notebook

- Python을 대화식으로 편리하게 실행할 수 있게 만든 IDE
- Jupyter라는 회사에서 제작, 페르난도 페레즈 개발
- 목성의 위성의 발견이 기록된 갈릴레오 갈릴레이의 공책에 대한 존경의 의미

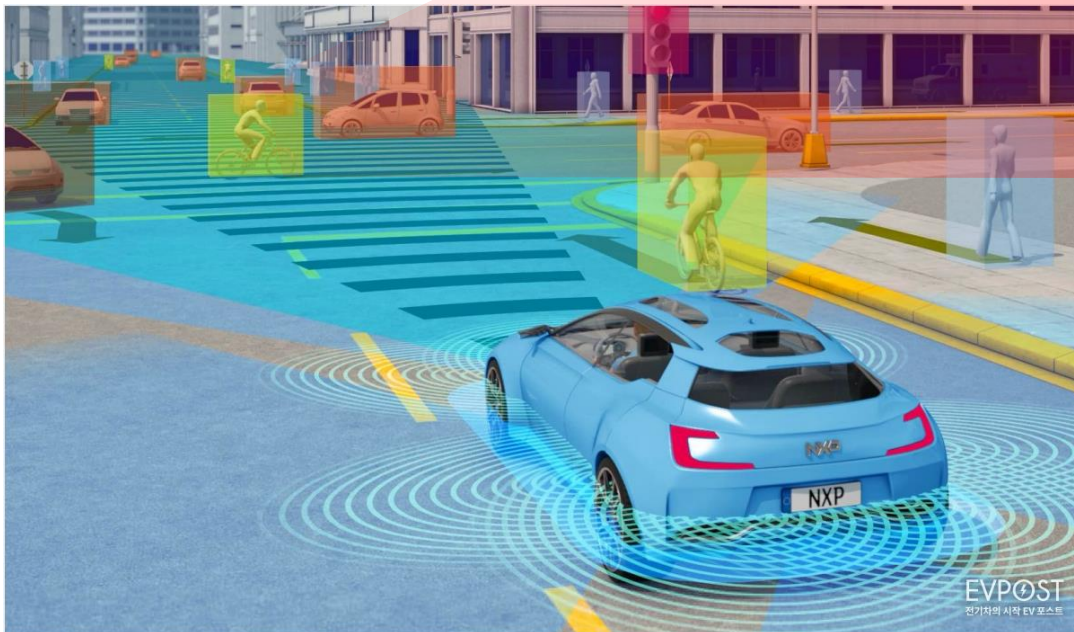
■ Google Colab

- Jupyter Notebook cloud version
- Google에서 제공
- 무료 GPU 사용, 간편한 공유



실습

- 자율주행에서 사람 인식하여 분류해보기

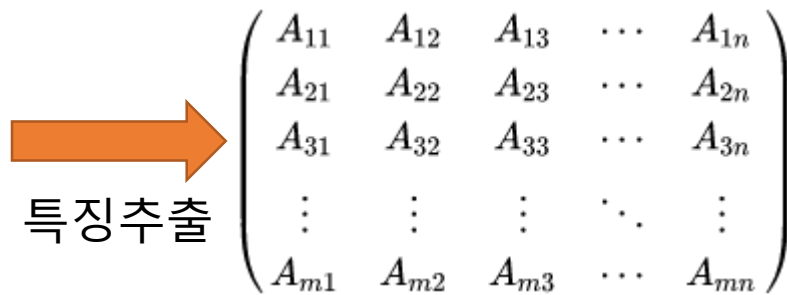
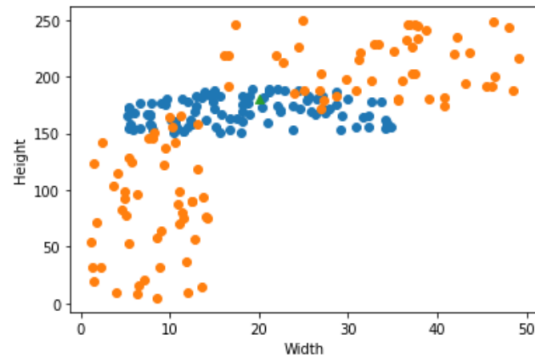
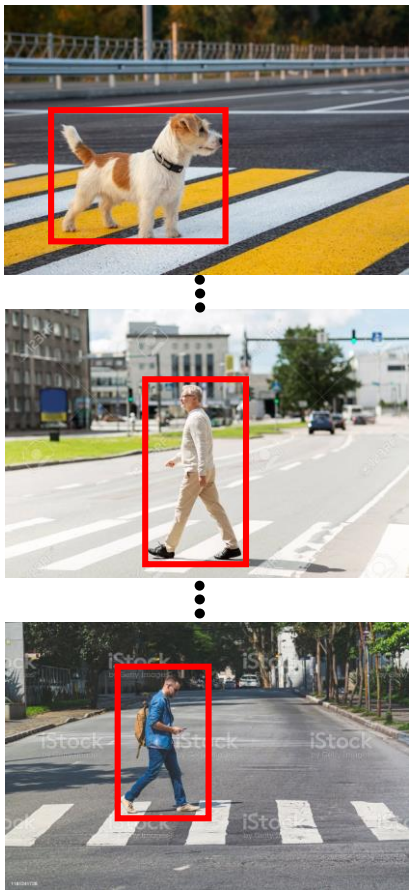


Height

Width

실습

■ 과정



인식된 사람 특징
데이터 생성하기

학습

K-NN
알고리즘

생성

학습모델

예측하기

모델평가
(정확도 측정)

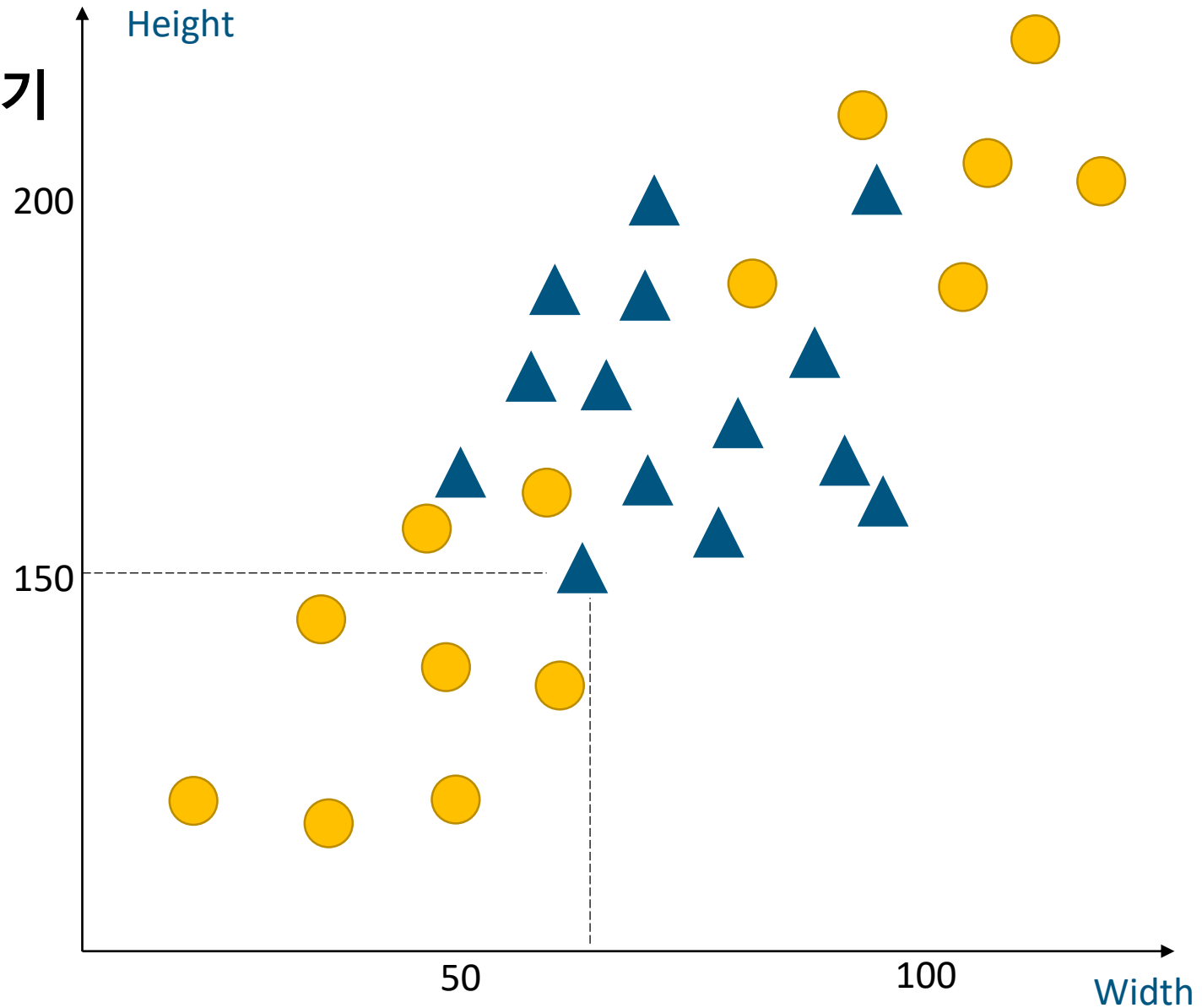
실습

인식된 사람 특징 데이터 생성하기

Height	Width	Person?
220	50	0
200	55	0
190	54	0
175	52	1
160	100	1

●
●
●

155	45	1
150	50	1
50	10	0
45	44	0
32	31	0
20	53	0



실습

▪ Numpy 및 Matplotlib 활용하여 생성하기

```
import numpy as np
import matplotlib.pyplot as plt

normal_person_width = np.random.uniform(low=20, high=100, size=100)
normal_person_height = np.random.uniform(low=150, high=190, size=100)

error_person_width = np.append(np.random.uniform(low= 1, high=19, size=50),
                                np.random.uniform(low=101, high=200, size=50))
error_person_height= np.append(np.random.uniform(low= 1, high=149, size=50),
                                np.random.uniform(low=191, high=250, size=50))

print(normal_person_width)

plt.scatter(normal_person_width, normal_person_height)
plt.scatter(error_person_width, error_person_height)
plt.xlabel('Width')
plt.ylabel('Height')
plt.show()
```


실습

- 학습을 위한 입력 데이터와 정답 준비하기

```
width = np.append(normal_person_width, error_person_width)
```

```
height = np.append(normal_person_height, error_person_height)
```

```
answer = [1]*100 + [0]*100
```

```
person_data = [[1, w] for l, w in zip(width, height)]
```

실습

- K-NN 알고리즘을 적용하여 학습하기(모델생성 및 평가)

```
from sklearn.neighbors import KNeighborsClassifier
```

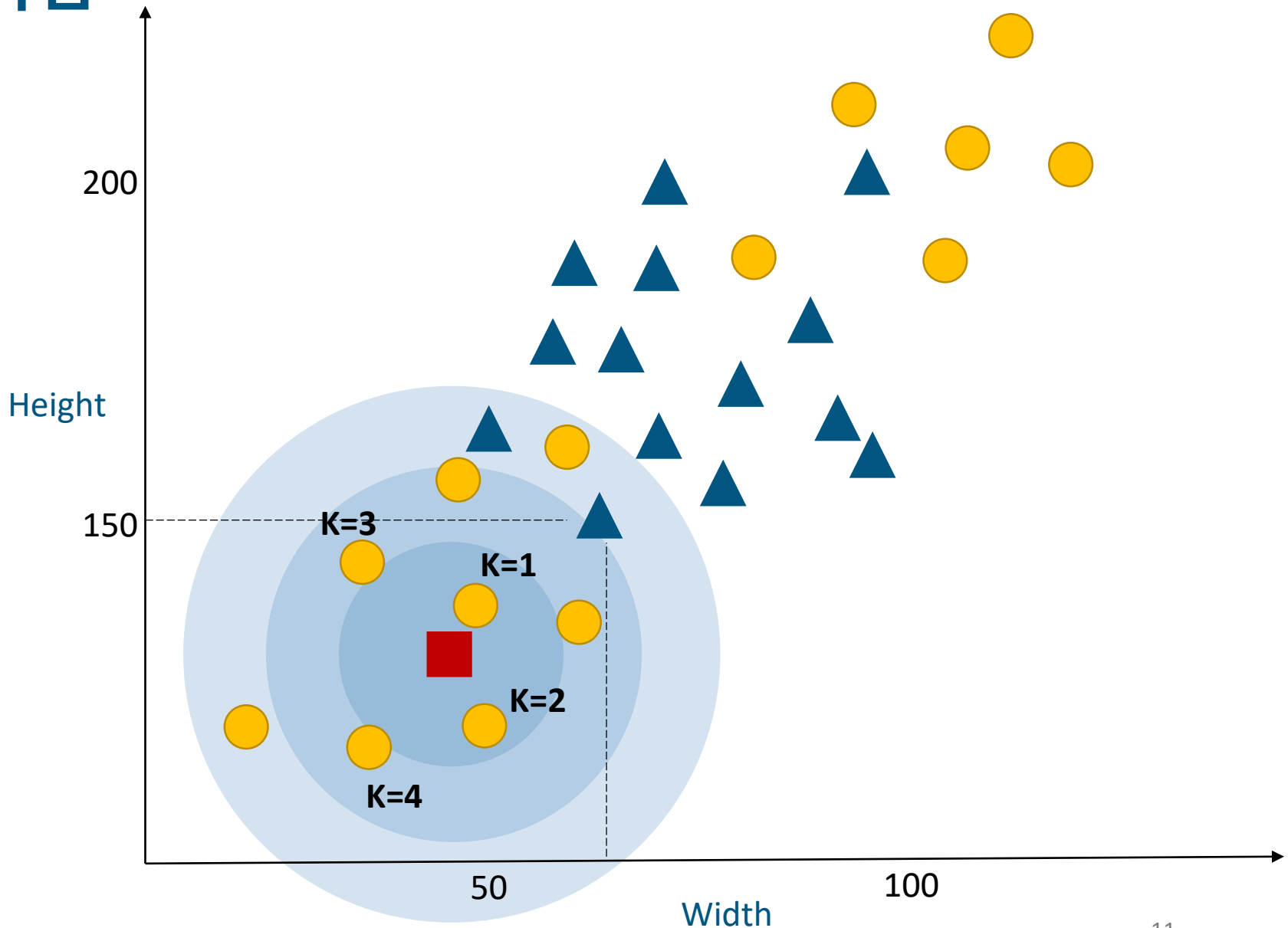
```
kn = KNeighborsClassifier()  
Kn.fit(person_data, answer)  
kn.score(person_data, answer)
```

실습 : K-NN 알고리즘

■ K-최근접 이웃(K-NN, Nearest Neighbor)

알고리즘

- 분류(Classification)을 위한 머신러닝 알고리즘
- 이미지 처리, 글자/얼굴 인식, 추천 알고리즘, 의료 분야 등에서 많이 사용



실습

- 예측하기

```
kn.predict([[20, 180]])  
print(kn._fit_X)  
print(kn._y)
```

- 그래프로 보기(Matplotlib 활용)

```
plt.scatter(normal_person_width, normal_person_height)  
plt.scatter(error_person_width, error_person_height)  
plt.scatter(20, 180, marker='^')  
plt.xlabel('Width')  
plt.ylabel('Height')  
plt.show()
```

실습

■ 최적의 파라미터 찾기

```
kn = KNeighborsClassifier()  
kn.fit(person_data, answer)  
  
for n in range(1, 100):  
    kn.n_neighbors = n  
    score = kn.score(person_data, answer)  
    print(str(n)+":"+str(score))
```

“하루에 3시간을 걸으면 7년 후에
지구를 한바퀴 돌 수 있다”
-사무엘존슨-

