

自主練習

- ❖ 請利用petal_length 和 petal_width 分類 Versicolour(1)和 Virginica(2)
- ❖ 請做標準化，並訓練 logistic regression 模型
- ❖ 請自行完成下列程式碼，以畫出資料散佈圖與決策邊界

```
t = np.linspace(-3,3,200)
plt.plot(X_std[y==1, 0], X_std[y==1, 1], 'o')
plt.plot(X_std[y==2, 0], X_std[y==2, 1], 'x')
plt.plot(t, -(theta_0+theta_1*t)/theta_2)
plt.xlabel('petal length [standardized]')
plt.ylabel('petal width [standardized]')
```

- ❖ p.s 不能使用 plot_decision_regions

自主練習

1. 導入 Iris 數據集
2. 只考慮兩個特徵 “Sepal length” (花萼長)和 “Sepal width” (花萼寬)
3. 分割 training set 及 test set (train:test = 7:3)，並標準化
4. 訓練模型

```
from sklearn.naive_bayes import GaussianNB  
  
gnb = GaussianNB()  
gnb.fit(X_train_std, y_train)
```

5. 算出正確率並畫出決策區域圖

Python

```
X_combined_std = np.vstack((X_train_std, X_test_std))
y_combined = np.hstack((y_train, y_test))

plot_decision_regions(X_combined_std, y_combined,
                      classifier=gnb, test_idx=range(105, 150))

plt.xlabel('sepal length [standardized]')
plt.ylabel('sepal width [standardized]')
plt.legend(loc='upper left')
plt.tight_layout()
#plt.savefig('images.png', dpi=300)
plt.show()
```

