

# SVM

- Using the SVC algorithm implemented by the Python Scikit-learn, classify the three types of flowers (Setosa, Versicolor, Virgin) in Iris dataset according to the **Petal length and width**.
- Change a sample into an outlier.
- Then use different values of inverse regularization parameter  $C$  for separate classifications and observe how different values of  $C$  affect this outlier.
- Plot the scatterplot with decision region for each different value of  $C$ . Also explain why a large value of  $C$  will make the outlier fall in the right decision region?

- Related parameter settings:
  - The ratio of training and test set is 7:3 using `random_state=1` for dataset splitting
  - Use Standardized features for both training and testing dataset
  - The way to change a sample into an outlier: after feature standardization, modify the features of the first sample in the Iris dataset to `[1, -1]` to make it an outlier
  - Parameter setting for the Scikit-learn SVC:
    - The values of `C` are: 0.1, 1, 10, 100
    - `random_state=0`
    - `kernel=linear`

# Results

