

Grid Search CV

Imports

```
In [1]: import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

Loading Dataset

```
In [2]: from sklearn.datasets import load_iris
```

```
In [3]: iris = load_iris()
```

```
In [4]: type(iris)
```

```
Out[4]: sklearn.utils.Bunch
```

See all the keys

```
In [5]: iris.keys()
```

```
Out[5]: dict_keys(['data', 'target', 'frame', 'target_names', 'DESCR', 'feature_names', 'filename'])
```

Description

```
In [6]: print(iris["DESCR"])
```

```
.. _iris_dataset:
```

```
Iris plants dataset
```

```
-----
```

```
**Data Set Characteristics:**
```

```
:Number of Instances: 150 (50 in each of three classes)
:Number of Attributes: 4 numeric, predictive attributes and the class
:Attribute Information:
  - sepal length in cm
  - sepal width in cm
  - petal length in cm
  - petal width in cm
  - class:
    - Iris-Setosa
    - Iris-Versicolour
    - Iris-Virginica
```

```
:Summary Statistics:
```

```
=====  =====  =====  =====  =====
                Min   Max    Mean     SD    Class Correlation
=====  =====  =====  =====  =====
```


In [10]:

```
Out[10]: array(['setosa', 'versicolor', 'virginica'], dtype='<U10')
```

Create Dataframe

In [11]:

Out[11]:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	target
0	5.1	3.5	1.4	0.2	0
1	4.9	3.0	1.4	0.2	0
2	4.7	3.2	1.3	0.2	0
3	4.6	3.1	1.5	0.2	0
4	5.0	3.6	1.4	0.2	0

In [12]:

Out[12]:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	target
0	5.1	3.5	1.4	0.2	0
1	4.9	3.0	1.4	0.2	0
2	4.7	3.2	1.3	0.2	0
3	4.6	3.1	1.5	0.2	0
4	5.0	3.6	1.4	0.2	0

In [13]:

```
Out[13]: array([0, 1, 2])
```

In [14]:

In [15]:

Out[15]:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	target
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa

sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	target
4	5.0	3.6	1.4	0.2 setosa

```
In [16]: df["target"].unique()
```

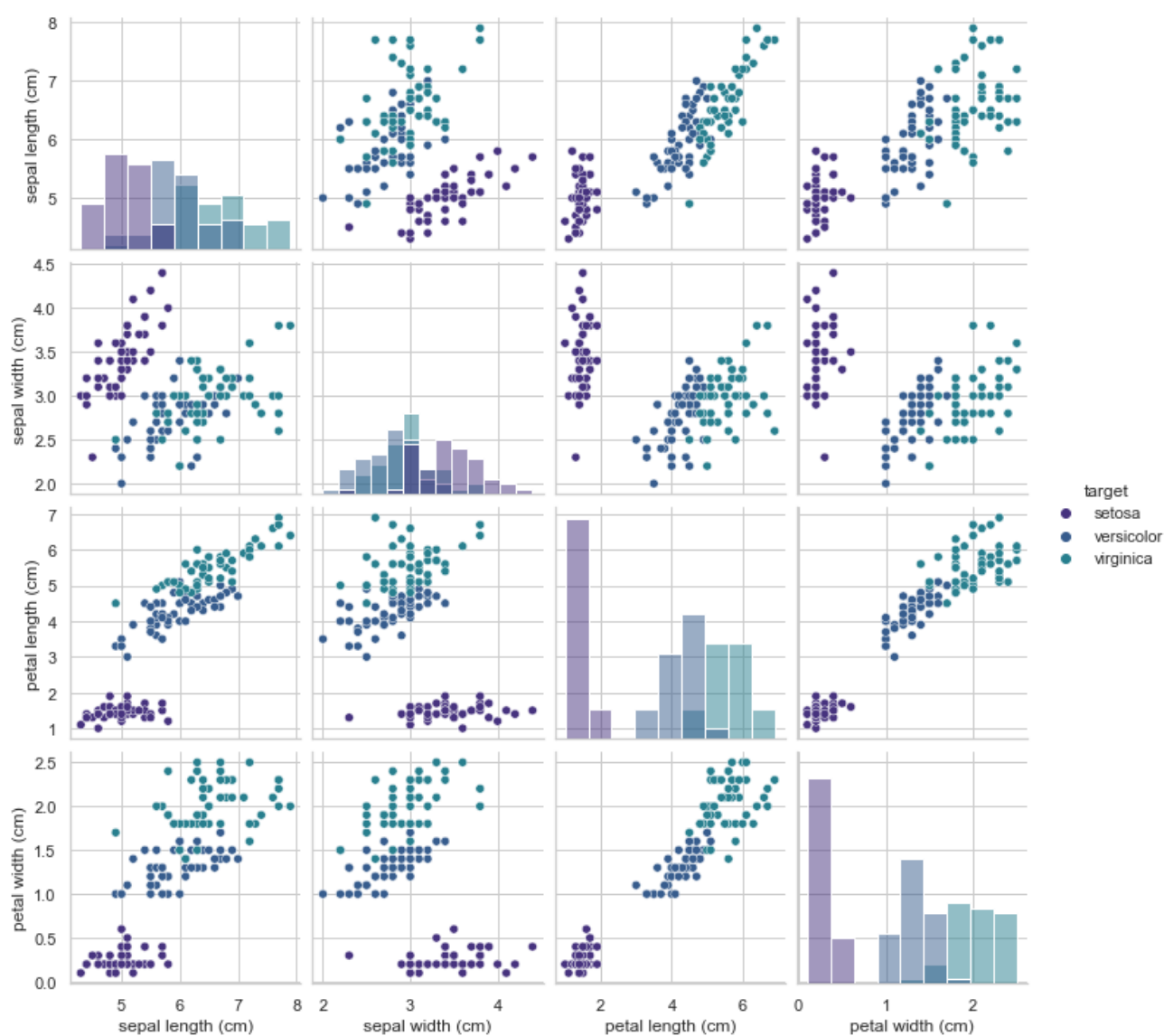
```
Out[16]: array(['setosa', 'versicolor', 'virginica'], dtype=object)
```

EDA

```
In [17]: sns.set(style="whitegrid", palette="viridis")
```

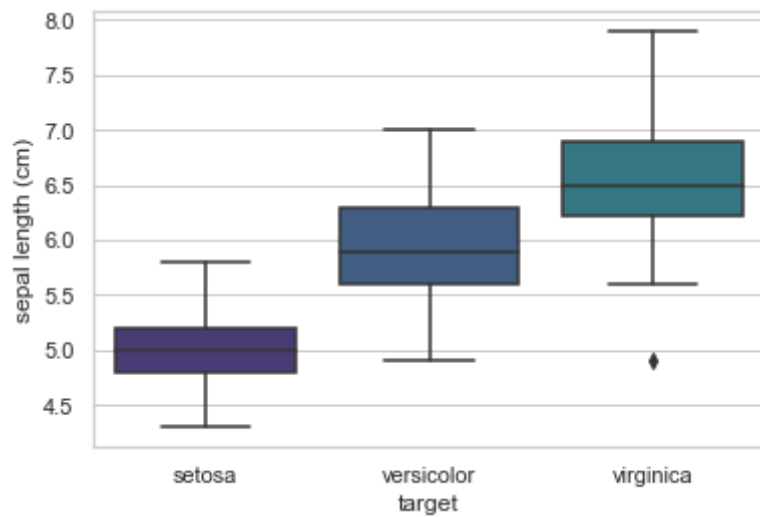
```
In [18]: sns.pairplot(df, hue="target", diag_kind='hist')
```

```
Out[18]: <seaborn.axisgrid.PairGrid at 0x234a9cdc070>
```



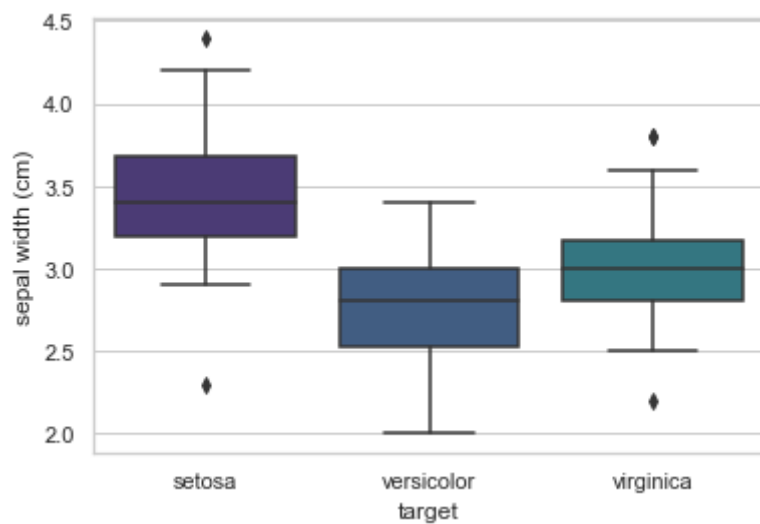
```
In [19]: sns.boxplot(x="target", y="sepal length (cm)", data=df)
```

```
Out[19]: <AxesSubplot:xlabel='target', ylabel='sepal length (cm)'>
```



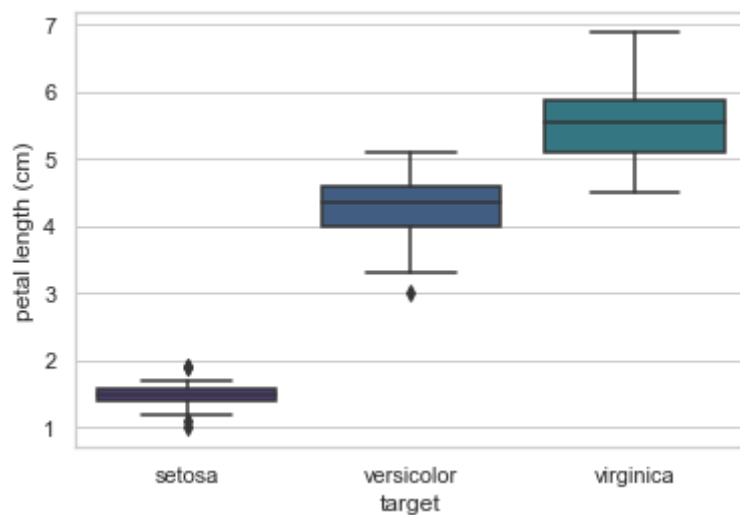
```
In [20]: sns.boxplot(x="target", y="sepal width (cm)", data=df)
```

```
Out[20]: <AxesSubplot:xlabel='target', ylabel='sepal width (cm)'\>
```



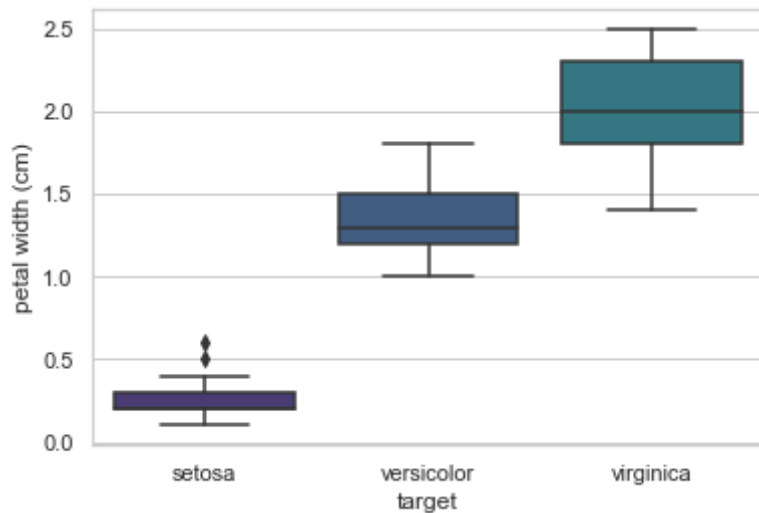
```
In [21]: sns.boxplot(x="target", y="petal length (cm)", data=df)
```

```
Out[21]: <AxesSubplot:xlabel='target', ylabel='petal length (cm)'\>
```



```
In [22]: sns.boxplot(x="target", y="petal width (cm)", data=df)
```

Out[22]: <AxesSubplot:xlabel='target', ylabel='petal width (cm)'\>



Feature Selection

In [23]: `df.head()`

Out[23]:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	target
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

In [24]:

```
X = df.drop("target", axis=1) # Independent
y = df["target"] # Dependent
```

Train Test Split

In [25]: `from sklearn.model_selection import train_test_split`

In [26]: `X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=202)`

In [27]: `X_train[:5]`

Out[27]:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
120	6.9	3.2	5.7	2.3
117	7.7	3.8	6.7	2.2
76	6.8	2.8	4.8	1.4
86	6.7	3.1	4.7	1.5
57	4.9	2.4	3.3	1.0

```
In [28]: y_train[:5]
```

```
Out[28]: 120    virginica
117    virginica
76     versicolor
86     versicolor
57     versicolor
Name: target, dtype: object
```

Grid Search CV

Grid Search

Grid-search is used to find the optimal hyperparameters of a model which results in the most accurate predictions. Grid search builds a model for every combination of hyperparameters specified and evaluates each model.

A model hyperparameter is a characteristic of a model that is external to the model and whose value cannot be estimated from data. The value of the hyperparameter has to be set before the learning process begins. For example, c in Support Vector Machines, k in k-Nearest Neighbors, the number of hidden layers in Neural Networks.

In contrast, a parameter is an internal characteristic of the model and its value can be estimated from data. Example, beta coefficients of linear/logistic regression or support vectors in Support Vector Machines.

Cross Validation

In K Fold cross validation, the data is divided into k subsets. Now the holdout method is repeated k times, such that each time, one of the k subsets is used as the test set/ validation set and the other $k-1$ subsets are put together to form a training set. The error estimation is averaged over all k trials to get total effectiveness of our model. As can be seen, every data point gets to be in a validation set exactly once, and gets to be in a training set $k-1$ times. This significantly reduces bias as we are using most of the data for fitting, and also significantly reduces variance as most of the data is also being used in validation set. Interchanging the training and test sets also adds to the effectiveness of this method. As a general rule and empirical evidence, $K = 5$ or 10 is generally preferred, but nothing's fixed and it can take any value.

```
In [29]: from sklearn.model_selection import GridSearchCV
from sklearn.svm import SVC
```

```
In [30]: print(SVC().get_params().keys())
```

```
dict_keys(['C', 'break_ties', 'cache_size', 'class_weight', 'coef0', 'decision_function_shape', 'degree', 'gamma', 'kernel', 'max_iter', 'probability', 'random_state', 'shrinking', 'tol', 'verbose'])
```

```
In [31]: params = {
    'C': [0.01, 0.1, 1, 10, 100],
    'gamma': [0.01, 0.1, 1, 10, 100],
    'kernel': ['linear', 'rbf', 'poly']
}
```

```
In [32]: grid = GridSearchCV(SVC(), params, verbose=2)
# 1st param: Sklearn model
# 2nd param: dictionary of model hyperparameter values
# verbose -> Higher verbose will give more information in output
```

```
In [33]: grid.fit(X_train, y_train)
```

Fitting 5 folds for each of 75 candidates, totalling 375 fits

```
[CV] END .....C=0.01, gamma=0.01, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=0.01, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=0.1, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=1, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=linear; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=rbf; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=poly; total time= 0.0s
[CV] END .....C=0.01, gamma=10, kernel=poly; total time= 0.0s
```


[illegible]

10/14

[illegible]

[illegible]

```

[CV] END .....C=100, gamma=1, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=1, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=1, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=1, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=1, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=1, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=1, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=1, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=1, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=1, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=linear; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=linear; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=linear; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=linear; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=linear; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=10, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=linear; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=linear; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=linear; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=linear; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=linear; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=rbf; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=poly; total time= 0.0s
[CV] END .....C=100, gamma=100, kernel=poly; total time= 0.0s

```

```

Out[33]: GridSearchCV(estimator=SVC(),
                      param_grid={'C': [0.01, 0.1, 1, 10, 100],
                                   'gamma': [0.01, 0.1, 1, 10, 100],
                                   'kernel': ['linear', 'rbf', 'poly']},
                      verbose=2)

```

```

In [34]: grid.best_params_

```

```

Out[34]: {'C': 0.1, 'gamma': 1, 'kernel': 'rbf'}

```

```

In [35]: grid_pred = grid.predict(X_test)
          # The grid model has been fitted with the best combination as seen above
          # We can directly predict without fitting again as refit=True by default in GridSearchCV

```

```

In [36]: from sklearn.metrics import classification_report, confusion_matrix

```

```

In [37]: print(confusion_matrix(y_test, grid_pred))
          print(classification_report(y_test, grid_pred))

```

```

[[18  0  0]
 [ 0 11  2]
 [ 0  3 11]]
precision    recall  f1-score   support

```

setosa	1.00	1.00	1.00	18
versicolor	0.79	0.85	0.81	13
virginica	0.85	0.79	0.81	14
accuracy			0.89	45
macro avg	0.88	0.88	0.88	45
weighted avg	0.89	0.89	0.89	45