Grid Search CV

Imports

```
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', 'filenam e'])
```

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
```

```
sepal length:
                      5.84
             4.3 7.9
                                   0.7826
                           0.83
sepal width:
             2.0 4.4
                            0.43
                                  -0.4194
                      3.05
                                   0.9490 (high!)
petal length:
             1.0 6.9
                      3.76
                            1.76
petal width:
                      1.20 0.76
             0.1 2.5
                                   0.9565 (high!)
:Missing Attribute Values: None
:Class Distribution: 33.3% for each of 3 classes.
:Creator: R.A. Fisher
:Donor: Michael Marshall (MARSHALL%PLU@io.arc.nasa.gov)
:Date: July, 1988
```

The famous Iris database, first used by Sir R.A. Fisher. The dataset is taken from Fisher's paper. Note that it's the same as in R, but not as in the UCI Machine Learning Repository, which has two wrong data points.

This is perhaps the best known database to be found in the pattern recognition literature. Fisher's paper is a classic in the field and is referenced frequently to this day. (See Duda & Hart, for example.) The data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant. One class is linearly separable from the other 2; the latter are NOT linearly separable from each other.

.. topic:: References

- Fisher, R.A. "The use of multiple measurements in taxonomic problems" Annual Eugenics, 7, Part II, 179-188 (1936); also in "Contributions to Mathematical Statistics" (John Wiley, NY, 1950).
- Duda, R.O., & Hart, P.E. (1973) Pattern Classification and Scene Analysis. (Q327.D83) John Wiley & Sons. ISBN 0-471-22361-1. See page 218.
- Dasarathy, B.V. (1980) "Nosing Around the Neighborhood: A New System Structure and Classification Rule for Recognition in Partially Exposed Environments". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No. 1, 67-71.
- Gates, G.W. (1972) "The Reduced Nearest Neighbor Rule". IEEE Transactions on Information Theory, May 1972, 431-433.
- See also: 1988 MLC Proceedings, 54-64. Cheeseman et al"s AUTOCLASS II conceptual clustering system finds 3 classes in the data.
- Many, many more ...

Data

```
In [7]:
      iris["data"][:5] # Independent columns
Out[7]: array([[5.1, 3.5, 1.4, 0.2],
          [4.9, 3., 1.4, 0.2],
          [4.7, 3.2, 1.3, 0.2],
          [4.6, 3.1, 1.5, 0.2],
          [5., 3.6, 1.4, 0.2]]
In [8]:
      iris["feature names"] # Column names
Out[8]: ['sepal length (cm)',
      sepal width (cm)',
      'petal length (cm)',
      'petal width (cm)']
In [9]:
      np.array(iris["target"]) # Dependent column
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
```

```
In [10]:
          iris["target names"]
          # So, 0 is setosa, 1 is versicolor and 2 is virginica
Out[10]: array(['setosa', 'versicolor', 'virginica'], dtype='<U10')
         Create Dataframe
In [11]:
          df = pd.DataFrame(data=iris["data"], columns=iris["feature_names"])
          df["target"] = iris["target"]
          df.head()
            sepal length (cm) sepal width (cm) petal length (cm) petal width (cm) target
Out[11]:
          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
                        5.0
                                      3.6
                                                      1.4
                                                                    0.2
                                                                            0
In [12]:
          df.head()
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
                                                                    0.2
                                                                            0
                                      3.0
                                                      1.4
          2
                        4.7
                                      3.2
                                                      1.3
                                                                    0.2
                                                                            0
          3
                        4.6
                                      3.1
                                                      1.5
                                                                    0.2
                                                                            0
                        5.0
                                      3.6
                                                      1.4
                                                                    0.2
In [13]:
          df["target"].unique()
         array([0, 1, 2])
Out[13]:
In [14]:
           df["target"].replace(range(3), iris["target_names"], inplace=True)
In [15]:
          df.head()
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
```

4.6

3.1

1.5

0.2 setosa

3

```
sepal length (cm) sepal width (cm) petal length (cm) petal width (cm) target

4 5.0 3.6 1.4 0.2 setosa
```

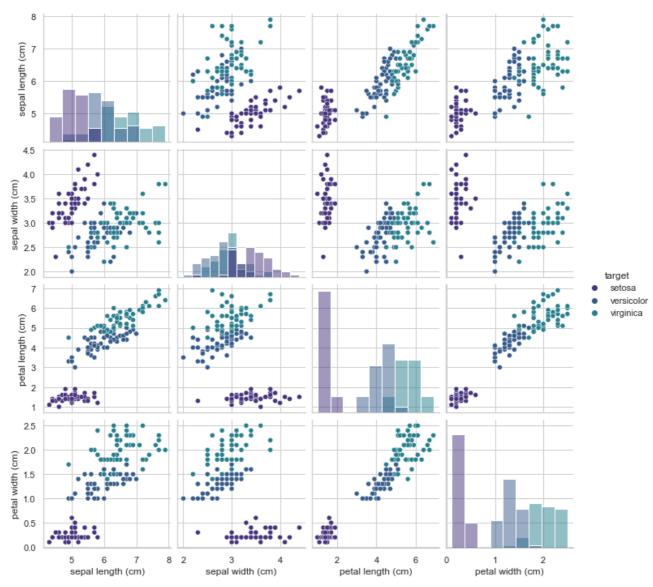
```
In [16]: df["target"].unique()
```

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

EDA

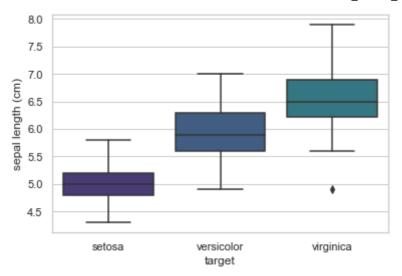
```
In [17]: sns.set(style="whitegrid", palette="viridis")
In [18]: sns.pairplot(df, hue="target", diag_kind='hist')
```

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



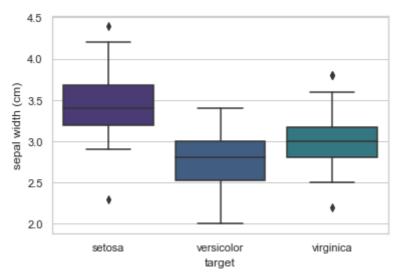
```
In [19]: sns.boxplot(x="target", y="sepal length (cm)", data=df)
```

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



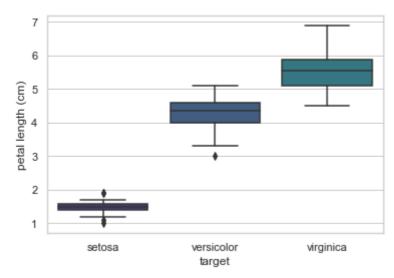
```
In [20]: sns.boxplot(x="target", y="sepal width (cm)", data=df)
```

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



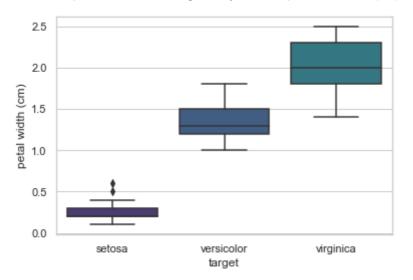
```
In [21]: sns.boxplot(x="target", y="petal length (cm)", data=df)
```

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

ıt[23]:	sep	al 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

Train Test Split

	•					
In [25]:	<pre>from sklearn.model_selection import train_test_split</pre>					
In [26]:	X_train, X_test,	y_train, y_test	= train_test_sp	lit(X, y, test_si	ze=0.3, random_s	
In [27]:	X_train[:5]					
Out[27]:	sepal length (cr	n) sepal width (cm)	petal length (cm)	petal width (cm)		
	120	5.9 3.2	5.7	2.3		
	117	7.7 3.8	6.7	2.2		
	76	5.8 2.8	4.8	1.4		
	86	5.7 3.1	4.7	1.5		

3.3

1.0

2.4

4.9

57

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_s hape', 'degree', 'gamma', 'kernel', 'max_iter', 'probability', 'random_state', 'shrinkin g', '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
0.0s
[CV] END ......C=0.01, gamma=0.01, 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=1, kernel=linear; total time=
            0.0s
[CV] END ......C=0.01, gamma=10, kernel=poly; total time=
            0.0s
0.0s
0.0s
0.0s
```

```
[CV] END ......C=0.01, gamma=10, kernel=poly; total time=
                    0.0s
[CV] END ......C=0.01, gamma=100, kernel=linear; total time=
                    0.0s
0.0s
0.05
[CV] END .................C=0.01, gamma=100, kernel=linear; total time=
                    0.05
[CV] END ......C=0.01, gamma=100, kernel=linear; total time=
                    0.05
[CV] END ......C=0.01, gamma=100, kernel=rbf; total time=
                    0.05
[CV] END ......C=0.01, gamma=100, kernel=poly; total time=
                    0.05
[CV] END ......C=0.01, gamma=100, kernel=poly; total time=
                    0.05
[CV] END ......C=0.01, gamma=100, kernel=poly; total time=
                    0.05
0.0s
[CV] END ......C=0.1, gamma=0.1, kernel=rbf; total time=
                    0.0s
[CV] END ......C=0.1, gamma=10, kernel=linear; total time=
                    0.0s
0.0s
0.0s
[CV] END ......C=0.1, gamma=10, kernel=linear; total time=
                    0.0s
[CV] END ......C=0.1, gamma=10, kernel=linear; total time=
                    0.0s
0.0s
0.0s
0.0s
```

```
[CV] END ......C=0.1, gamma=10, kernel=rbf; total time=
                              0.0s
[CV] END ......C=0.1, gamma=10, kernel=rbf; total time=
                              0.0s
0.0s
0.05
0.05
0.05
0.05
[CV] END ......C=0.1, gamma=100, kernel=linear; total time=
                              0.05
0.05
0.05
0.05
0.0s
[CV] END ......C=1, gamma=0.01, kernel=rbf; total time=
                              0.0s
[CV] END ......C=1, gamma=0.01, kernel=rbf; total time=
                              0.0s
[CV] END ......C=1, gamma=0.01, kernel=rbf; total time=
                              0.0s
[CV] END .....C=1, gamma=0.01, kernel=rbf; total time=
                              0.0s
[CV] END ......C=1, gamma=0.01, kernel=rbf; total time=
                              0.0s
[CV] END ......C=1, gamma=0.01, kernel=poly; total time=
                              0.0s
0.0s
[CV] END ......C=1, gamma=0.1, kernel=linear; total time=
                              0.0s
0.0s
0.0s
0.0s
0.0s
[CV] END ......C=1, gamma=0.1, kernel=rbf; total time=
                              0.0s
0.0s
0.0s
0.0s
0.0s
0.0s
[CV] END ......C=1, gamma=1, kernel=linear; total time=
                              0.0s
[CV] END ......C=1, gamma=1, kernel=linear; total time=
                              0.0s
0.0s
[CV] END ......C=1, gamma=1, kernel=linear; total time=
                              0.0s
[CV] END ......C=1, gamma=1, kernel=linear; total time=
                              0.0s
0.0s
0.0s
0.0s
0.0s
0.0s
[CV] END ......C=1, gamma=1, kernel=poly; total time=
                              0.0s
[CV] END ......C=1, gamma=10, kernel=linear; total time=
                              0.0s
[CV] END ......C=1, gamma=10, kernel=linear; total time=
                              0.0s
```

```
[CV] END ......C=1, gamma=10, kernel=linear; total time=
                                                 0.0s
[CV] END ......C=1, gamma=10, kernel=linear; total time=
                                                 0.0s
[CV] END ......C=1, gamma=10, kernel=linear; total time=
                                                 0.0s
[CV] END ......C=1, gamma=10, kernel=rbf; total time=
                                                 0.05
[CV] END ......C=1, gamma=10, kernel=poly; total time=
                                                 0.05
[CV] END ......C=1, gamma=10, kernel=poly; total time=
                                                 0.05
[CV] END ......C=1, gamma=10, kernel=poly; total time=
                                                 0.05
0.05
[CV] END ......C=1, gamma=10, kernel=poly; total time=
                                                 0.05
[CV] END ......C=1, gamma=100, kernel=linear; total time=
                                                 0.05
[CV] END ......C=1, gamma=100, kernel=linear; total time=
                                                 0.05
[CV] END ......C=1, gamma=100, kernel=linear; total time=
                                                 0.0s
0.0s
[CV] END ......C=1, gamma=100, kernel=linear; total time=
                                                 0.0s
[CV] END ......C=1, gamma=100, kernel=rbf; total time=
                                                 0.0s
0.0s
0.0s
0.0s
0.0s
[CV] END ......C=1, gamma=100, kernel=poly; total time=
                                                 0.0s
[CV] END ......C=1, gamma=100, kernel=poly; total time=
                                                 0.0s
0.0s
[CV] END ......C=1, gamma=100, kernel=poly; total time=
                                                 0.0s
[CV] END ......C=1, gamma=100, kernel=poly; total time=
                                                 0.0s
[CV] END ......C=10, gamma=0.01, kernel=linear; total time=
                                                 0.0s
0.0s
[CV] END ......C=10, gamma=0.01, kernel=rbf; total time=
                                                 0.0s
[CV] END ......C=10, gamma=0.01, kernel=poly; total time=
                                                 0.0s
0.0s
[CV] END ......C=10, gamma=0.1, kernel=linear; total time=
                                                 0.0s
[CV] END ..................C=10, gamma=0.1, kernel=linear; total time=
                                                 0.0s
[CV] END ......C=10, gamma=0.1, kernel=rbf; total time=
                                                 0.0s
0.0s
0.0s
0.0s
0.0s
[CV] END ......C=10, gamma=0.1, kernel=poly; total time=
                                                 0.0s
[CV] END ......C=10, gamma=0.1, kernel=poly; total time=
                                                 0.0s
[CV] END ......C=10, gamma=1, kernel=linear; total time=
                                                 0.0s
[CV] END ......C=10, gamma=1, kernel=rbf; total time=
                                                 0.0s
0.0s
```

52 PM	Grid_Search_CV			
[CV]	ENDC=10, gamma=1, kernel=poly;	total	time=	0.0s
	END			0.0s
	END			0.0s
	END			0.0s
	ENDC=10, gamma=10, kernel=linear;			0.0s
	ENDC=10, gamma=10, kernel=linear;			0.0s
	ENDC=10, gamma=10, kernel=linear;			0.0s
[CV]	ENDC=10, gamma=10, kernel=linear;	total	time=	0.0s
[CV]	ENDC=10, gamma=10, kernel=linear;	total	time=	0.0s
	END			0.0s
	ENDC=10, gamma=10, kernel=rbf;			0.0s
[CV]	END	total	time=	0.0s
	END			0.0s
	END			0.0s
	ENDC=10, gamma=10, kernel=poly;			0.0s
	ENDC=10, gamma=10, kernel=poly;			0.0s
	ENDC=10, gamma=10, kernel=poly;			0.0s
[CV]	ENDC=10, gamma=10, kernel=poly;	total	time=	0.0s
[CV]	ENDC=10, gamma=10, kernel=poly;	total	time=	0.0s
[CV]	ENDC=10, gamma=100, kernel=linear;	total	time=	0.0s
	ENDC=10, gamma=100, kernel=linear;			0.0s
	ENDC=10, gamma=100, kernel=linear;			0.0s
	ENDC=10, gamma=100, kernel=linear;			0.0s
	END			0.0s
	ENDC=10, gamma=100, kernel=rbf;			0.0s
	ENDC=10, gamma=100, kernel=rbf;			0.0s
	ENDC=10, gamma=100, kernel=rbf;			0.0s
[CV]	ENDC=10, gamma=100, kernel=rbf;	total	time=	0.0s
[CV]	ENDC=10, gamma=100, kernel=rbf;	total	time=	0.0s
	ENDC=10, gamma=100, kernel=poly;			0.0s
	END			0.0s
	END			0.0s
	END			0.0s
	END			0.0s
	ENDC=100, gamma=0.01, kernel=linear;			0.0s
	ENDC=100, gamma=0.01, kernel=linear;			0.0s
[CV]	ENDC=100, gamma=0.01, kernel=linear;	total	time=	0.0s
[CV]	ENDC=100, gamma=0.01, kernel=linear;	total	time=	0.0s
[CV]	ENDC=100, gamma=0.01, kernel=linear;	total	time=	0.0s
[CV]	ENDC=100, gamma=0.01, kernel=rbf;	total	time=	0.0s
	ENDC=100, gamma=0.01, kernel=rbf;			0.0s
	ENDC=100, gamma=0.01, kernel=rbf;			0.0s
	END			0.0s
	END			0.0s
	END			0.0s
	ENDC=100, gamma=0.01, kernel=poly;			0.0s
	ENDC=100, gamma=0.01, kernel=poly;			0.0s
	ENDC=100, gamma=0.01, kernel=poly;			0.0s
	ENDC=100, gamma=0.01, kernel=poly;			0.0s
	ENDC=100, gamma=0.1, kernel=linear;			0.0s
	ENDC=100, gamma=0.1, kernel=linear;			0.0s
	ENDC=100, gamma=0.1, kernel=linear;			0.0s
	ENDC=100, gamma=0.1, kernel=linear;			0.0s
	END			0.0s
	END			0.0s
	END			0.0s
	ENDC=100, gamma=0.1, kernel=rbf;			0.0s
	ENDC=100, gamma=0.1, kernel=rbf;			0.0s
	ENDC=100, gamma=0.1, kernel=rbf;			0.0s
[CV]	ENDC=100, gamma=0.1, kernel=poly;	total	time=	0.0s
	ENDC=100, gamma=0.1, kernel=poly;			0.0s
	ENDC=100, gamma=0.1, kernel=poly;			0.0s
	END			0.0s
	END			0.0s
	END			0.0s
	ENDC=100, gamma=1, kernel=linear;			0.0s
	ENDC=100, gamma=1, kernel=linear;			0.0s
	ENDC=100, gamma=1, kernel=linear;			0.0s
[CV]	ENDC=100, gamma=1, kernel=linear;	total	time=	0.0s
_				

```
[CV] END ......C=100, gamma=1, kernel=rbf; total time=
                                                                            0.05
         [CV] END ......C=100, gamma=1, kernel=rbf; total time=
                                                                            0.05
         [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=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=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=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 macro avg weighted avg	0.88 0.89	0.88 0.89	0.89 0.88 0.89	45 45 45