

# JADAVPUR UNIVERSITY

## MACHINE LEARNING LABORATORY

### ASSIGNMENT #2 - A) IRIS PLANTS DATASET

NAME - RITIK BAID

DEPARTMENT - INFORMATION TECHNOLOGY

ROLL NO - 001811001035

```
In [ ]:
# IRIS PLANT DATASET
# Random Forest Classifier(Without Tuning)[70-30 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, test_size=0.3, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier(n_estimators=20, random_state=0)
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

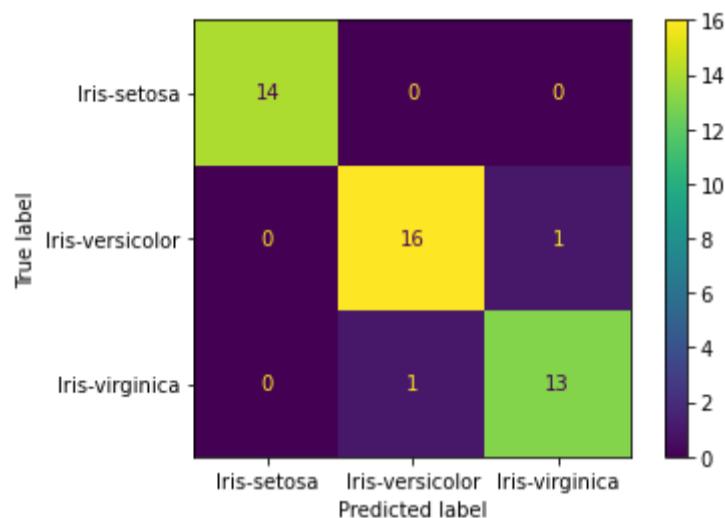
print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

Confusion Matrix:
[[14  0  0]
 [ 0 16  1]
 [ 0  1 13]]
-----
-----
Performance Evaluation
```

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	14
Iris-versicolor	0.94	0.94	0.94	17
Iris-virginica	0.93	0.93	0.93	14
accuracy			0.96	45
macro avg	0.96	0.96	0.96	45
weighted avg	0.96	0.96	0.96	45

-----  
Accuracy:  
0.9555555555555556



```
In [ ]:
# IRIS PLANT DATASET
# Random Forest Classifier(Without Tuning)[60-40 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.6, test_size=0.4, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier(n_estimators=20, random_state=0)
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

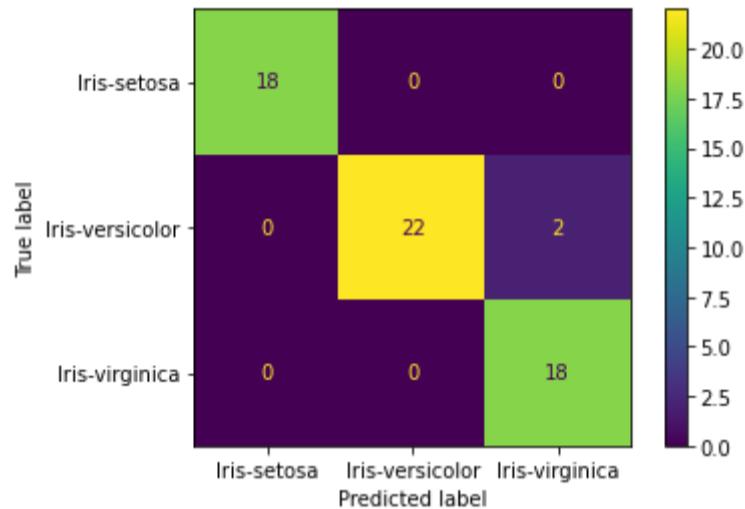
import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()
```

```

Confusion Matrix:
[[18  0  0]
 [ 0 22  2]
 [ 0  0 18]]
-----
Performance Evaluation
      precision    recall   f1-score   support
Iris-setosa       1.00     1.00     1.00      18
Iris-versicolor   1.00     0.92     0.96      24
Iris-virginica    0.90     1.00     0.95      18
accuracy          0.97
macro avg         0.97     0.97     0.97      60
weighted avg      0.97     0.97     0.97      60

```

Accuracy:  
0.9666666666666667



```

In [ ]:
# IRIS PLANT DATASET
# Random Forest Classifier(Without Tuning)[50-50 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.5, test_size=0.5, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier(n_estimators=20, random_state=0)
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

```

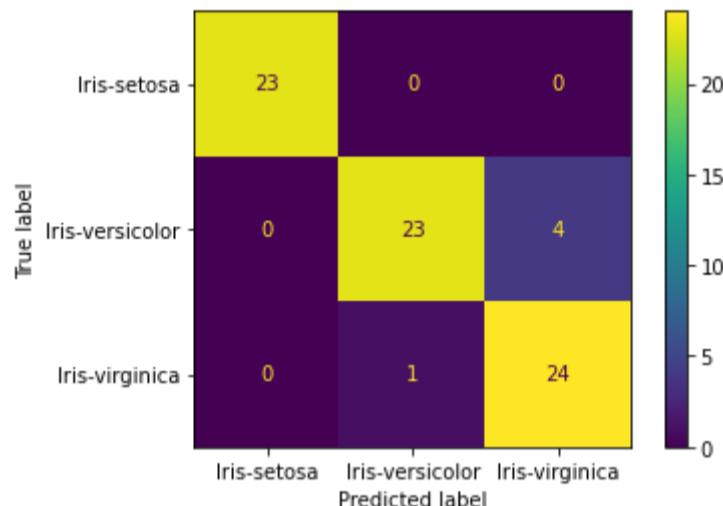
```

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

Confusion Matrix:
[[23  0  0]
 [ 0 23  4]
 [ 0  1 24]]
-----
-----
Performance Evaluation
      precision    recall   f1-score   support
Iris-setosa       1.00     1.00     1.00      23
Iris-versicolor   0.96     0.85     0.90      27
Iris-virginica    0.86     0.96     0.91      25
accuracy          0.93     0.93     0.93      75
macro avg         0.94     0.94     0.94      75
weighted avg      0.94     0.93     0.93      75
-----
-----

```

Accuracy:  
0.9333333333333333



```

In [ ]:
# IRIS PLANT DATASET
# Random Forest Classifier(Without Tuning)[40-60 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.4, test_size=0.6, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier(n_estimators=20, random_state=0)
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

```

```

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```

Confusion Matrix:

```

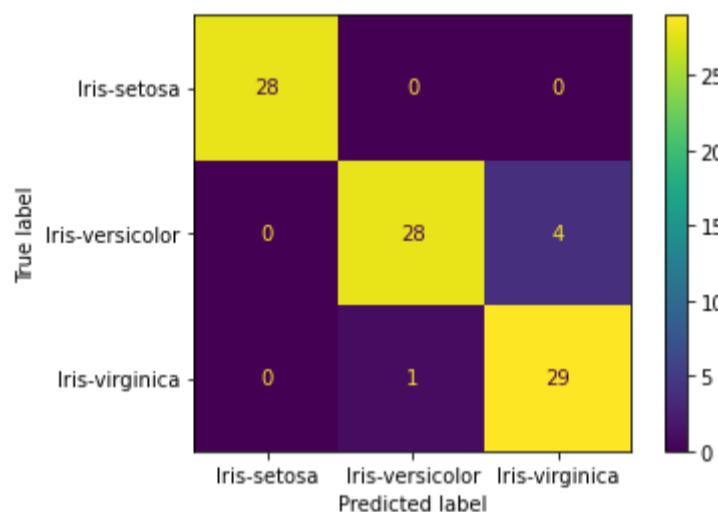
[[28  0  0]
 [ 0 28  4]
 [ 0  1 29]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	28
Iris-versicolor	0.97	0.88	0.92	32
Iris-virginica	0.88	0.97	0.92	30
accuracy			0.94	90
macro avg	0.95	0.95	0.95	90
weighted avg	0.95	0.94	0.94	90

Accuracy:

```
0.9444444444444444
```



In [ ]:

```

# IRIS PLANT DATASET
# Random Forest Classifier(Without Tuning)[30-70 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.3, test_size=0.7, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier(n_estimators=20, random_state=0)
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")

```

```

print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```

Confusion Matrix:

```

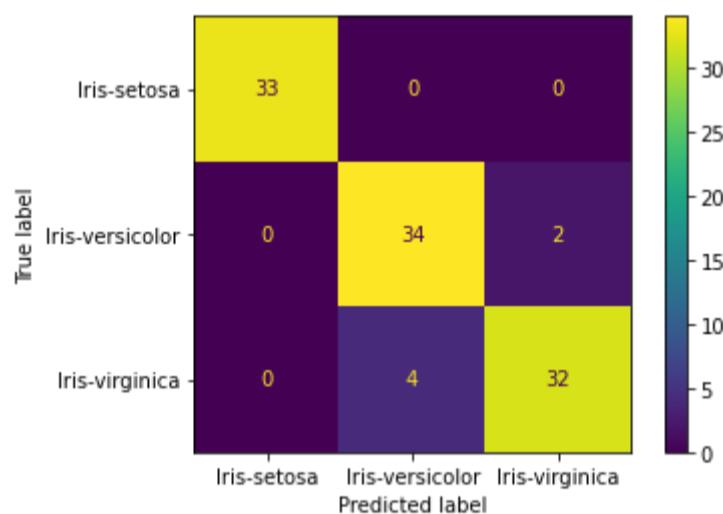
[[33  0  0]
 [ 0 34  2]
 [ 0  4 32]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	33
Iris-versicolor	0.89	0.94	0.92	36
Iris-virginica	0.94	0.89	0.91	36
accuracy			0.94	105
macro avg	0.95	0.94	0.94	105
weighted avg	0.94	0.94	0.94	105

Accuracy:

```
0.9428571428571428
```



In [ ]:

```
#####
#####
```

In [ ]:

```

# IRIS PLANT DATASET
# Random Forest Classifier(With Tuning)[70-30 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, test_size=0.30, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier()

```

```

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

from sklearn.model_selection import RandomizedSearchCV
# Number of trees in random forest
n_estimators = [int(x) for x in np.linspace(start = 200, stop = 2000, num = 10)]
# Number of features to consider at every split
max_features = ['auto', 'sqrt']
# Maximum number of levels in tree
max_depth = [int(x) for x in np.linspace(10, 110, num = 11)]
max_depth.append(None)
# Minimum number of samples required to split a node
min_samples_split = [2, 5, 10]
# Minimum number of samples required at each leaf node
min_samples_leaf = [1, 2, 4]
# Method of selecting samples for training each tree
bootstrap = [True, False]
# Create the random grid
random_grid = {'n_estimators': n_estimators,
               'max_features': max_features,
               'max_depth': max_depth,
               'min_samples_split': min_samples_split,
               'min_samples_leaf': min_samples_leaf,
               'bootstrap': bootstrap}
pprint(random_grid)

#####

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = RandomForestClassifier()
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores
rf_random = RandomizedSearchCV(estimator = classifier, param_distributions = random_grid, n_iter = 100, cv = 3, verbose=2, random_
# Fit the random search model
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```

Parameters currently in use:

```

{'bootstrap': True,
 'ccp_alpha': 0.0,
 'class_weight': None,
 'criterion': 'gini',
 'max_depth': None,
 'max_features': 'auto',
 'max_leaf_nodes': None,
 'max_samples': None,
 'min_impurity_decrease': 0.0,
 'min_impurity_split': None,
 'min_samples_leaf': 1,
 'min_samples_split': 2,
 'min_weight_fraction_leaf': 0.0,
 'n_estimators': 100,
 'n_jobs': None,
 'oob_score': False,
 'random_state': None,
 'verbose': 0,
 'warm_start': False}
{'bootstrap': [True, False],
 'max_depth': [10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, None],
 'max_features': ['auto', 'sqrt'],
 'max_leaf_nodes': None,
 'max_samples': None,
 'min_impurity_decrease': 0.0,
 'min_impurity_split': None,
 'min_samples_leaf': 1,
 'min_samples_split': 2,
 'min_weight_fraction_leaf': 0.0,
 'n_estimators': 100,
 'n_jobs': None,
 'oob_score': False,
 'random_state': None,
 'verbose': 0,
 'warm_start': False}

```

```
'max_features': ['auto', 'sqrt'],
'min_samples_leaf': [1, 2, 4],
'min_samples_split': [2, 5, 10],
'n_estimators': [200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000]}
Fitting 3 folds for each of 100 candidates, totalling 300 fits
```

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done 37 tasks | elapsed: 45.5s
[Parallel(n_jobs=-1)]: Done 158 tasks | elapsed: 3.1min
[Parallel(n_jobs=-1)]: Done 300 out of 300 | elapsed: 5.9min finished
```

Confusion Matrix:

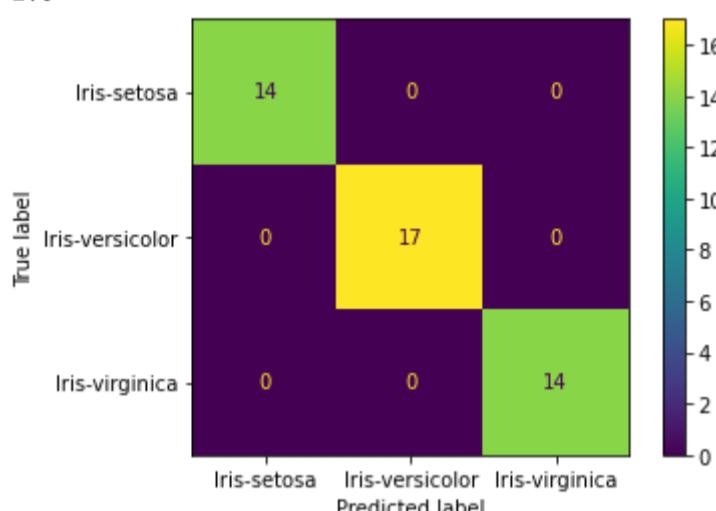
```
[[14  0  0]
 [ 0 17  0]
 [ 0  0 14]]
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	14
Iris-versicolor	1.00	1.00	1.00	17
Iris-virginica	1.00	1.00	1.00	14
accuracy			1.00	45
macro avg	1.00	1.00	1.00	45
weighted avg	1.00	1.00	1.00	45

Accuracy:

1.0



In [ ]:

```
# IRIS PLANT DATASET
# Random Forest Classifier(With Tuning)[60-40 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.6, test_size=0.4, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier()

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features
```

```

from sklearn.model_selection import RandomizedSearchCV
# Number of trees in random forest
n_estimators = [int(x) for x in np.linspace(start = 200, stop = 2000, num = 10)]
# Number of features to consider at every split
max_features = ['auto', 'sqrt']
# Maximum number of levels in tree
max_depth = [int(x) for x in np.linspace(10, 110, num = 11)]
max_depth.append(None)
# Minimum number of samples required to split a node
min_samples_split = [2, 5, 10]
# Minimum number of samples required at each leaf node
min_samples_leaf = [1, 2, 4]
# Method of selecting samples for training each tree
bootstrap = [True, False]
# Create the random grid
random_grid = {'n_estimators': n_estimators,
               'max_features': max_features,
               'max_depth': max_depth,
               'min_samples_split': min_samples_split,
               'min_samples_leaf': min_samples_leaf,
               'bootstrap': bootstrap}
pprint(random_grid)

#####
# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = RandomForestClassifier()
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores
rf_random = RandomizedSearchCV(estimator = classifier, param_distributions = random_grid, n_iter = 100, cv = 3, verbose=2, random_
# Fit the random search model
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```

Parameters currently in use:

```

{'bootstrap': True,
 'ccp_alpha': 0.0,
 'class_weight': None,
 'criterion': 'gini',
 'max_depth': None,
 'max_features': 'auto',
 'max_leaf_nodes': None,
 'max_samples': None,
 'min_impurity_decrease': 0.0,
 'min_impurity_split': None,
 'min_samples_leaf': 1,
 'min_samples_split': 2,
 'min_weight_fraction_leaf': 0.0,
 'n_estimators': 100,
 'n_jobs': None,
 'oob_score': False,
 'random_state': None,
 'verbose': 0,
 'warm_start': False}
{'bootstrap': [True, False],
 'max_depth': [10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, None],
 'max_features': ['auto', 'sqrt'],
 'min_samples_leaf': [1, 2, 4],
 'min_samples_split': [2, 5, 10],
 'n_estimators': [200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000]}
Fitting 3 folds for each of 100 candidates, totalling 300 fits

```

```

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done 37 tasks      | elapsed:  43.7s
[Parallel(n_jobs=-1)]: Done 158 tasks      | elapsed:  3.1min
[Parallel(n_jobs=-1)]: Done 300 out of 300 | elapsed:  5.9min finished

```

Confusion Matrix:

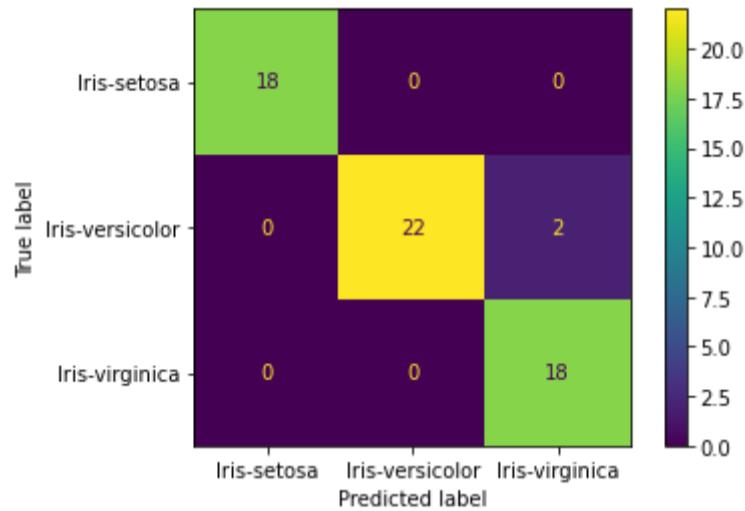
```

[[18  0  0]
 [ 0 22  2]
 [ 0  0 18]]

```

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	18
Iris-versicolor	1.00	0.92	0.96	24
Iris-virginica	0.90	1.00	0.95	18
accuracy			0.97	60
macro avg	0.97	0.97	0.97	60
weighted avg	0.97	0.97	0.97	60

Accuracy:  
0.9666666666666667



```
In [ ]:
# IRIS PLANT DATASET
# Random Forest Classifier(With Tuning)[50-50 split]
```

```
import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.5, test_size=0.5, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier()

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

from sklearn.model_selection import RandomizedSearchCV
# Number of trees in random forest
n_estimators = [int(x) for x in np.linspace(start = 200, stop = 2000, num = 10)]
# Number of features to consider at every split
max_features = ['auto', 'sqrt']
# Maximum number of levels in tree
max_depth = [int(x) for x in np.linspace(10, 110, num = 11)]
max_depth.append(None)
# Minimum number of samples required to split a node
min_samples_split = [2, 5, 10]
# Minimum number of samples required at each Leaf node
min_samples_leaf = [1, 2, 4]
# Method of selecting samples for training each tree
```

```

bootstrap = [True, False]
# Create the random grid
random_grid = {'n_estimators': n_estimators,
               'max_features': max_features,
               'max_depth': max_depth,
               'min_samples_split': min_samples_split,
               'min_samples_leaf': min_samples_leaf,
               'bootstrap': bootstrap}
pprint(random_grid)

#####
# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = RandomForestClassifier()
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores
rf_random = RandomizedSearchCV(estimator = classifier, param_distributions = random_grid, n_iter = 100, cv = 3, verbose=2, random_
# Fit the random search model
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```

Parameters currently in use:

```

{'bootstrap': True,
 'ccp_alpha': 0.0,
 'class_weight': None,
 'criterion': 'gini',
 'max_depth': None,
 'max_features': 'auto',
 'max_leaf_nodes': None,
 'max_samples': None,
 'min_impurity_decrease': 0.0,
 'min_impurity_split': None,
 'min_samples_leaf': 1,
 'min_samples_split': 2,
 'min_weight_fraction_leaf': 0.0,
 'n_estimators': 100,
 'n_jobs': None,
 'oob_score': False,
 'random_state': None,
 'verbose': 0,
 'warm_start': False}
{'bootstrap': [True, False],
 'max_depth': [10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, None],
 'max_features': ['auto', 'sqrt'],
 'min_samples_leaf': [1, 2, 4],
 'min_samples_split': [2, 5, 10],
 'n_estimators': [200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000]}
Fitting 3 folds for each of 100 candidates, totalling 300 fits

```

```

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done 37 tasks      | elapsed:  44.2s
[Parallel(n_jobs=-1)]: Done 158 tasks      | elapsed:  3.1min
[Parallel(n_jobs=-1)]: Done 300 out of 300 | elapsed:  5.9min finished

```

Confusion Matrix:

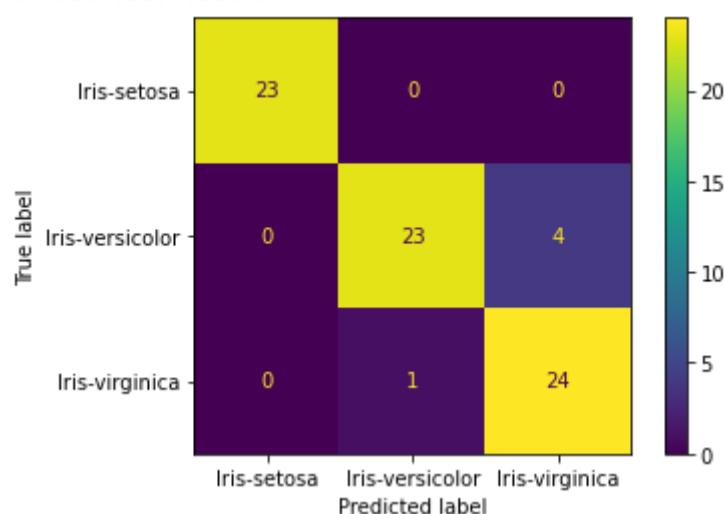
```

[[23  0  0]
 [ 0 23  4]
 [ 0  1 24]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	23
Iris-versicolor	0.96	0.85	0.90	27
Iris-virginica	0.86	0.96	0.91	25
accuracy			0.93	75
macro avg	0.94	0.94	0.94	75
weighted avg	0.94	0.93	0.93	75

-----  
Accuracy:  
0.9333333333333333



```
In [ ]:  
# IRIS PLANT DATASET  
# Random Forest Classifier(With Tuning)[40-60 split]
```

```
import pandas as pd  
import numpy as np  
  
# Dataset Preparation  
df = pd.read_csv("iris.data", header=None)  
  
col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']  
  
df.columns = col_name  
  
X = df.drop(['Class'], axis=1)  
y = df['Class']  
  
from sklearn.model_selection import train_test_split  
  
X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.4, test_size=0.6, random_state=10)  
  
# Feature Scaling  
from sklearn.preprocessing import StandardScaler  
  
sc = StandardScaler()  
X_train = sc.fit_transform(X_train)  
X_test = sc.transform(X_test)  
  
# Classification  
from sklearn.ensemble import RandomForestClassifier  
  
classifier = RandomForestClassifier()  
  
#####  
# Showing all the parameters  
  
from pprint import pprint  
# Look at parameters used by our current forest  
print('Parameters currently in use:\n')  
pprint(classifier.get_params())  
  
#####  
# Creating a set of important sample features  
  
from sklearn.model_selection import RandomizedSearchCV  
# Number of trees in random forest  
n_estimators = [int(x) for x in np.linspace(start = 200, stop = 2000, num = 10)]  
# Number of features to consider at every split  
max_features = ['auto', 'sqrt']  
# Maximum number of levels in tree  
max_depth = [int(x) for x in np.linspace(10, 110, num = 11)]  
max_depth.append(None)  
# Minimum number of samples required to split a node  
min_samples_split = [2, 5, 10]  
# Minimum number of samples required at each Leaf node  
min_samples_leaf = [1, 2, 4]  
# Method of selecting samples for training each tree  
bootstrap = [True, False]  
# Create the random grid  
random_grid = {'n_estimators': n_estimators,  
              'max_features': max_features,  
              'max_depth': max_depth,  
              'min_samples_split': min_samples_split,  
              'min_samples_leaf': min_samples_leaf,  
              'bootstrap': bootstrap}  
pprint(random_grid)  
  
#####  
# Use the random grid to search for best hyperparameters
```

```

# First create the base model to tune
classifier = RandomForestClassifier()
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores
rf_random = RandomizedSearchCV(estimator = classifier, param_distributions = random_grid, n_iter = 100, cv = 3, verbose=2, random_
# Fit the random search model
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```

Parameters currently in use:

```

{'bootstrap': True,
 'ccp_alpha': 0.0,
 'class_weight': None,
 'criterion': 'gini',
 'max_depth': None,
 'max_features': 'auto',
 'max_leaf_nodes': None,
 'max_samples': None,
 'min_impurity_decrease': 0.0,
 'min_impurity_split': None,
 'min_samples_leaf': 1,
 'min_samples_split': 2,
 'min_weight_fraction_leaf': 0.0,
 'n_estimators': 100,
 'n_jobs': None,
 'oob_score': False,
 'random_state': None,
 'verbose': 0,
 'warm_start': False}
{'bootstrap': [True, False],
 'max_depth': [10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, None],
 'max_features': ['auto', 'sqrt'],
 'min_samples_leaf': [1, 2, 4],
 'min_samples_split': [2, 5, 10],
 'n_estimators': [200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000]}
Fitting 3 folds for each of 100 candidates, totalling 300 fits

```

```

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done 37 tasks      | elapsed:  44.2s
[Parallel(n_jobs=-1)]: Done 158 tasks      | elapsed:  3.0min
[Parallel(n_jobs=-1)]: Done 300 out of 300 | elapsed:  5.9min finished

```

Confusion Matrix:

```

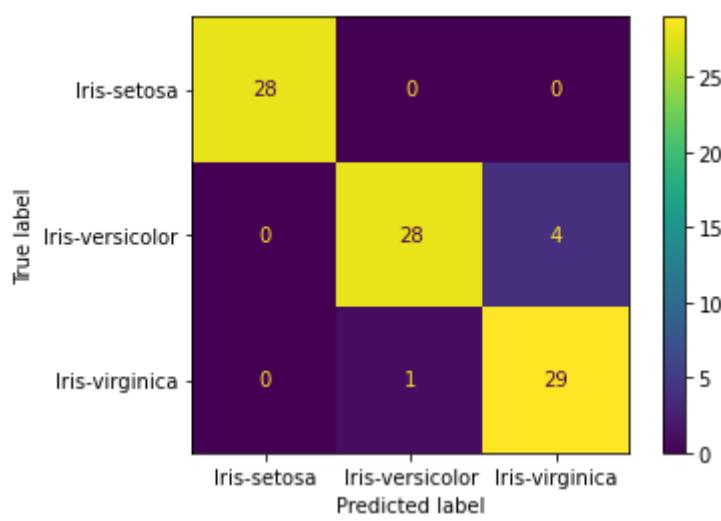
[[28  0  0]
 [ 0 28  4]
 [ 0  1 29]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	28
Iris-versicolor	0.97	0.88	0.92	32
Iris-virginica	0.88	0.97	0.92	30
accuracy			0.94	90
macro avg	0.95	0.95	0.95	90
weighted avg	0.95	0.94	0.94	90

Accuracy:

0.9444444444444444



```
In [ ]:
# IRIS PLANT DATASET
# Random Forest Classifier(With Tuning)[30-70 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.3, test_size=0.7, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier()

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

from sklearn.model_selection import RandomizedSearchCV
# Number of trees in random forest
n_estimators = [int(x) for x in np.linspace(start = 200, stop = 2000, num = 10)]
# Number of features to consider at every split
max_features = ['auto', 'sqrt']
# Maximum number of levels in tree
max_depth = [int(x) for x in np.linspace(10, 110, num = 11)]
max_depth.append(None)
# Minimum number of samples required to split a node
min_samples_split = [2, 5, 10]
# Minimum number of samples required at each leaf node
min_samples_leaf = [1, 2, 4]
# Method of selecting samples for training each tree
bootstrap = [True, False]
# Create the random grid
random_grid = {'n_estimators': n_estimators,
               'max_features': max_features,
               'max_depth': max_depth,
               'min_samples_split': min_samples_split,
               'min_samples_leaf': min_samples_leaf,
               'bootstrap': bootstrap}
pprint(random_grid)

#####
# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = RandomForestClassifier()
# Random search of parameters, using 3 fold cross validation,
```

```

# search across 100 different combinations, and use all available cores
rf_random = RandomizedSearchCV(estimator = classifier, param_distributions = random_grid, n_iter = 100, cv = 3, verbose=2, random_
# Fit the random search model
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```

Parameters currently in use:

```

{'bootstrap': True,
 'ccp_alpha': 0.0,
 'class_weight': None,
 'criterion': 'gini',
 'max_depth': None,
 'max_features': 'auto',
 'max_leaf_nodes': None,
 'max_samples': None,
 'min_impurity_decrease': 0.0,
 'min_impurity_split': None,
 'min_samples_leaf': 1,
 'min_samples_split': 2,
 'min_weight_fraction_leaf': 0.0,
 'n_estimators': 100,
 'n_jobs': None,
 'oob_score': False,
 'random_state': None,
 'verbose': 0,
 'warm_start': False}
{'bootstrap': [True, False],
 'max_depth': [10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, None],
 'max_features': ['auto', 'sqrt'],
 'min_samples_leaf': [1, 2, 4],
 'min_samples_split': [2, 5, 10],
 'n_estimators': [200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000]}

```

Fitting 3 folds for each of 100 candidates, totalling 300 fits

```

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done 37 tasks      | elapsed:  43.5s
[Parallel(n_jobs=-1)]: Done 158 tasks      | elapsed:  3.0min
[Parallel(n_jobs=-1)]: Done 300 out of 300 | elapsed:  5.9min finished

```

Confusion Matrix:

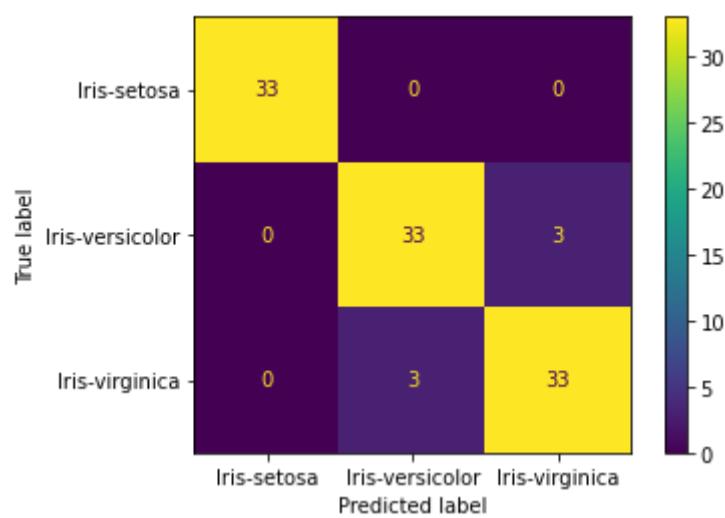
```

[[33  0  0]
 [ 0 33  3]
 [ 0  3 33]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	33
Iris-versicolor	0.92	0.92	0.92	36
Iris-virginica	0.92	0.92	0.92	36
accuracy			0.94	105
macro avg	0.94	0.94	0.94	105
weighted avg	0.94	0.94	0.94	105

Accuracy:  
0.9428571428571428



```
In [ ]: #####
```

```
In [ ]: # IRIS PLANT DATASET
# Multi Layer Perceptron(Without Tuning)[70-30 split]
```

```
import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, test_size=0.30, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification using MLP
from sklearn.neural_network import MLPClassifier

classifier = MLPClassifier()
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()
```

```
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200) reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
```

```
Confusion Matrix:
```

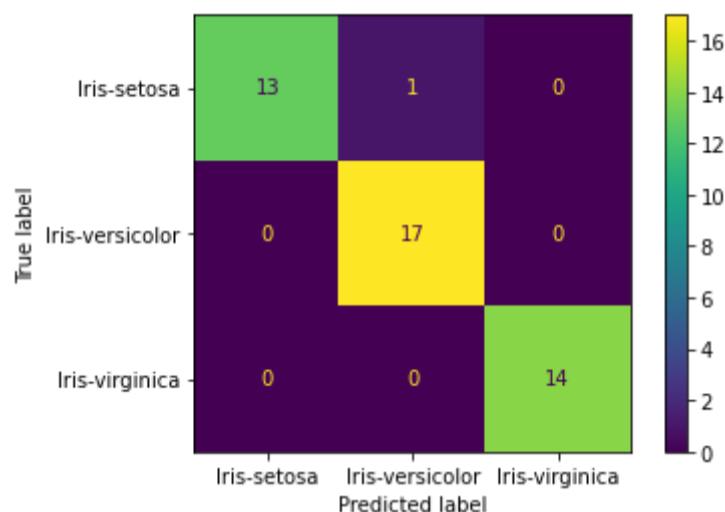
```
[[13  1  0]
 [ 0 17  0]
 [ 0  0 14]]
```

```
-----
```

```
Performance Evaluation
```

	precision	recall	f1-score	support
Iris-setosa	1.00	0.93	0.96	14
Iris-versicolor	0.94	1.00	0.97	17
Iris-virginica	1.00	1.00	1.00	14
accuracy			0.98	45
macro avg	0.98	0.98	0.98	45
weighted avg	0.98	0.98	0.98	45

-----  
Accuracy:  
0.9777777777777777



```
In [ ]:
# IRIS PLANT DATASET
# Multi Layer Perceptron(Without Tuning)[60-40 split]
```

```
import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.6, test_size=0.4, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification using MLP
from sklearn.neural_network import MLPClassifier

classifier = MLPClassifier()
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()
```

```
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200) reached and the optimization hasn't converged yet.
```

```
% self.max_iter, ConvergenceWarning)
```

```
Confusion Matrix:
```

```
[[17  1  0]
 [ 0 20  4]
 [ 0  0 18]]
```

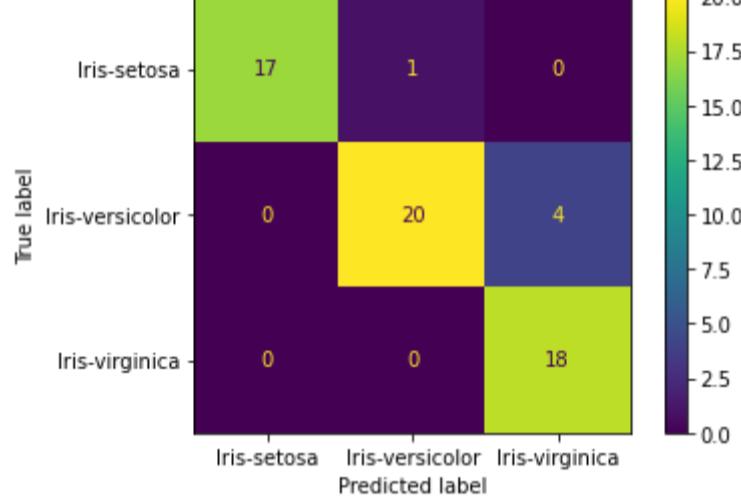
```
-----
```

Performance Evaluation				
	precision	recall	f1-score	support
Iris-setosa	1.00	0.94	0.97	18
Iris-versicolor	0.95	0.83	0.89	24
Iris-virginica	0.82	1.00	0.90	18
accuracy			0.92	60
macro avg	0.92	0.93	0.92	60
weighted avg	0.93	0.92	0.92	60

```
-----
```

```
Accuracy:
```

```
0.9166666666666666
```



```
In [ ]:
```

```
# IRIS PLANT DATASET
# Multi Layer Perceptron(Without Tuning)[50-50 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.5, test_size=0.5, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification using MLP
from sklearn.neural_network import MLPClassifier

classifier = MLPClassifier()
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
```

```

print("-----")
print("Accuracy:")
print(accuracy_score(y_test, y_pred))

```

```

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```

Confusion Matrix:

```

[[22  1  0]
 [ 0 23  4]
 [ 0  0 25]]
-----
```

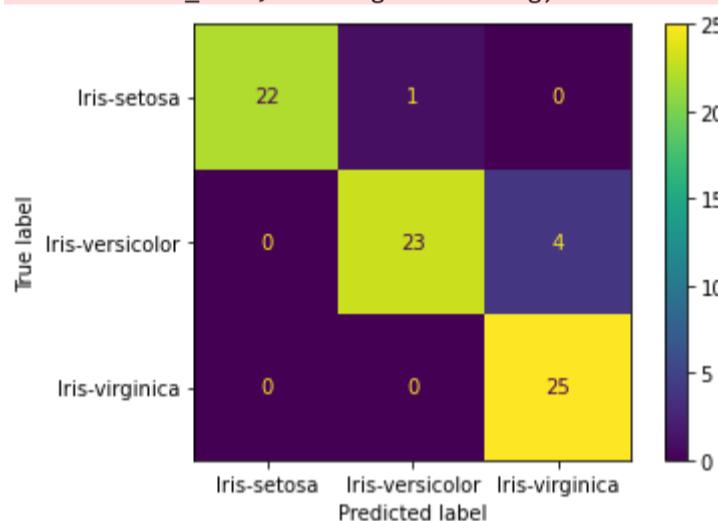
Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	0.96	0.98	23
Iris-versicolor	0.96	0.85	0.90	27
Iris-virginica	0.86	1.00	0.93	25
accuracy			0.93	75
macro avg	0.94	0.94	0.94	75
weighted avg	0.94	0.93	0.93	75

Accuracy:

```
0.9333333333333333
```

```
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200) reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
```



In [ ]:

```

# IRIS PLANT DATASET
# Multi Layer Perceptron(Without Tuning)[40-60 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.4, test_size=0.6, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification using MLP
from sklearn.neural_network import MLPClassifier

classifier = MLPClassifier()
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

```

```

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```

Confusion Matrix:

```

[[28  0  0]
 [ 0 28  4]
 [ 0  0 30]]
-----
```

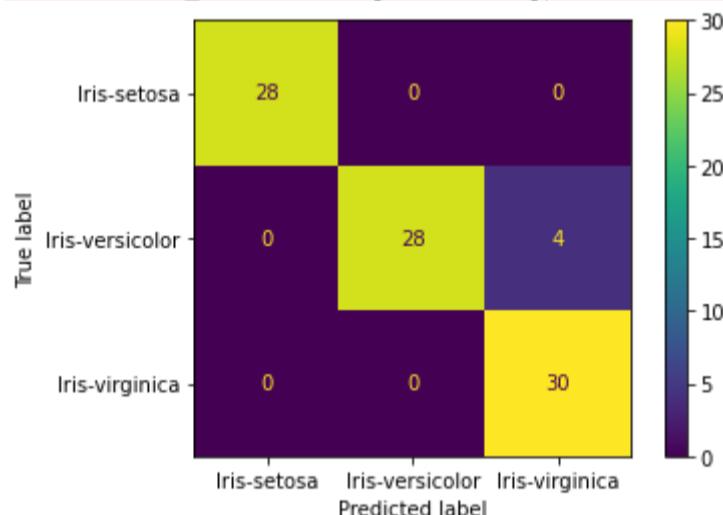
Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	28
Iris-versicolor	1.00	0.88	0.93	32
Iris-virginica	0.88	1.00	0.94	30
accuracy			0.96	90
macro avg	0.96	0.96	0.96	90
weighted avg	0.96	0.96	0.96	90

Accuracy:

```
0.9555555555555556
```

```
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200) reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
```



In [ ]:

```

# IRIS PLANT DATASET
# Multi Layer Perceptron(Without Tuning)[30-70 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.3, test_size=0.7, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification using MLP
from sklearn.neural_network import MLPClassifier

```

```

classifier = MLPClassifier()
classifier.fit(X_train,y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```

Confusion Matrix:

```

[[33  0  0]
 [ 0 31  5]
 [ 0  1 35]]
-----
```

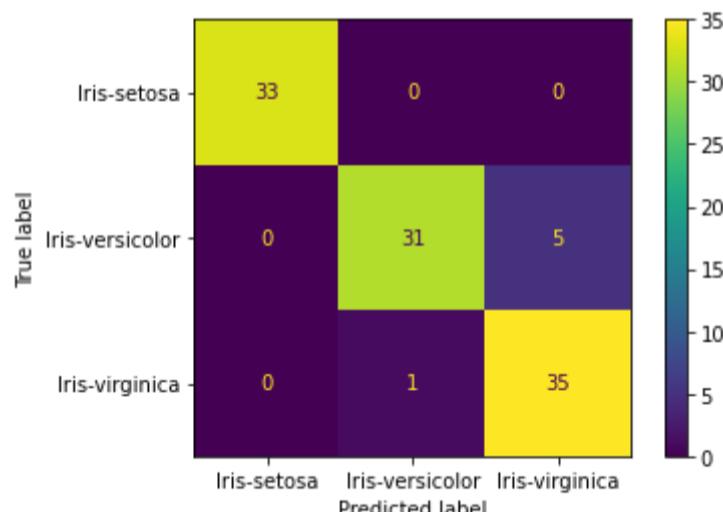
Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	33
Iris-versicolor	0.97	0.86	0.91	36
Iris-virginica	0.88	0.97	0.92	36
accuracy			0.94	105
macro avg	0.95	0.94	0.94	105
weighted avg	0.95	0.94	0.94	105

Accuracy:

0.9428571428571428

```
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200) reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
```



In [ ]:

```

# IRIS PLANT DATASET
# Multi Layer Perceptron(With Tuning)[70-30 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, test_size=0.3, random_state=10)

```

```

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.neural_network import MLPClassifier
classifier = MLPClassifier(max_iter=100)

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

parameter_space = {
    'hidden_layer_sizes': [(50,50,50), (50,100,50), (100,)],
    'activation': ['tanh', 'relu'],
    'solver': ['sgd', 'adam'],
    'alpha': [0.0001, 0.05],
    'learning_rate': ['constant','adaptive'],
}
pprint(parameter_space)

#####

from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = MLPClassifier(max_iter=100)
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(classifier, parameter_space, n_jobs=-1, cv=3)
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```

Parameters currently in use:

```
{
'activation': 'relu',
'alpha': 0.0001,
'batch_size': 'auto',
'beta_1': 0.9,
'beta_2': 0.999,
'early_stopping': False,
'epsilon': 1e-08,
'hidden_layer_sizes': (100,),
'learning_rate': 'constant',
'learning_rate_init': 0.001,
'max_fun': 15000,
'max_iter': 100,
'momentum': 0.9,
'n_iter_no_change': 10,
'nesterovs_momentum': True,
'power_t': 0.5,
'random_state': None,
'shuffle': True,
'solver': 'adam',
'tol': 0.0001,
'validation_fraction': 0.1,
```

```

'verbose': False,
'warm_start': False}
{'activation': ['tanh', 'relu'],
'alpha': [0.0001, 0.05],
'hidden_layer_sizes': [(50, 50, 50), (50, 100, 50), (100,)],
'learning_rate': ['constant', 'adaptive'],
'solver': ['sgd', 'adam']}
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (100) reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)

```

Confusion Matrix:

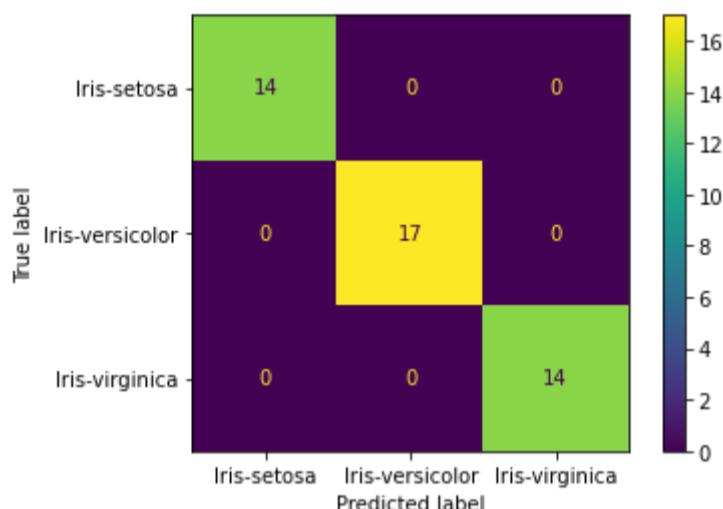
```
[[14  0  0]
 [ 0 17  0]
 [ 0  0 14]]
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	14
Iris-versicolor	1.00	1.00	1.00	17
Iris-virginica	1.00	1.00	1.00	14
accuracy			1.00	45
macro avg	1.00	1.00	1.00	45
weighted avg	1.00	1.00	1.00	45

Accuracy:

```
1.0
```



In [ ]:

```

# IRIS PLANT DATASET
# Multi Layer Perceptron(With Tuning)[60-40 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.6, test_size=0.4, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.neural_network import MLPClassifier
classifier = MLPClassifier(max_iter=100)

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

```

```

parameter_space = {
    'hidden_layer_sizes': [(50,50,50), (50,100,50), (100,)],
    'activation': ['tanh', 'relu'],
    'solver': ['sgd', 'adam'],
    'alpha': [0.0001, 0.05],
    'learning_rate': ['constant','adaptive'],
}
pprint(parameter_space)

#####
from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = MLPClassifier(max_iter=100)
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(classifier, parameter_space, n_jobs=-1, cv=3)
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```

Parameters currently in use:

```

{'activation': 'relu',
 'alpha': 0.0001,
 'batch_size': 'auto',
 'beta_1': 0.9,
 'beta_2': 0.999,
 'early_stopping': False,
 'epsilon': 1e-08,
 'hidden_layer_sizes': (100,),
 'learning_rate': 'constant',
 'learning_rate_init': 0.001,
 'max_fun': 15000,
 'max_iter': 100,
 'momentum': 0.9,
 'n_iter_no_change': 10,
 'nesterovs_momentum': True,
 'power_t': 0.5,
 'random_state': None,
 'shuffle': True,
 'solver': 'adam',
 'tol': 0.0001,
 'validation_fraction': 0.1,
 'verbose': False,
 'warm_start': False}
{'activation': ['tanh', 'relu'],
 'alpha': [0.0001, 0.05],
 'hidden_layer_sizes': [(50, 50, 50), (50, 100, 50), (100,)],
 'learning_rate': ['constant', 'adaptive'],
 'solver': ['sgd', 'adam']}

```

```

/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (100) reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)

```

Confusion Matrix:

```

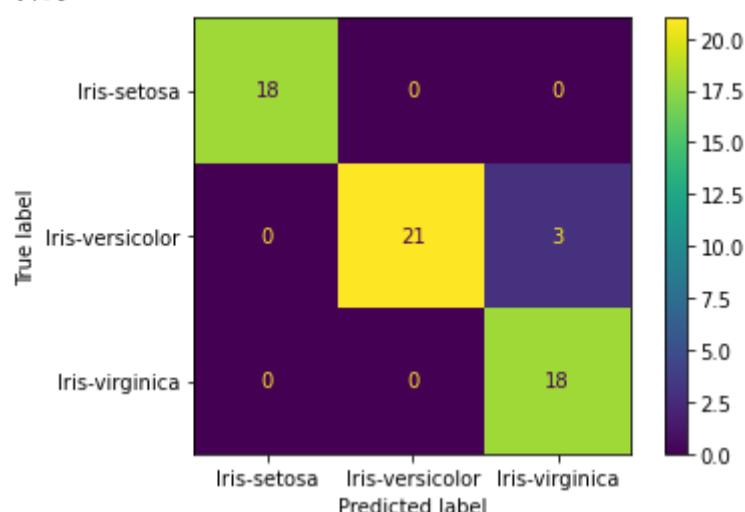
[[18  0  0]
 [ 0 21  3]
 [ 0  0 18]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	18
Iris-versicolor	1.00	0.88	0.93	24
Iris-virginica	0.86	1.00	0.92	18
accuracy			0.95	60

macro avg	0.95	0.96	0.95	60
weighted avg	0.96	0.95	0.95	60

-----  
Accuracy:  
0.95



```
In [ ]:
# IRIS PLANT DATASET
# Multi Layer Perceptron(With Tuning)[50-50 split]
```

```
import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.5, test_size=0.5, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.neural_network import MLPClassifier
classifier = MLPClassifier(max_iter=100)

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

parameter_space = {
    'hidden_layer_sizes': [(50,50,50), (50,100,50), (100,)],
    'activation': ['tanh', 'relu'],
    'solver': ['sgd', 'adam'],
    'alpha': [0.0001, 0.05],
    'learning_rate': ['constant', 'adaptive'],
}
pprint(parameter_space)

#####
from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = MLPClassifier(max_iter=100)
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(classifier, parameter_space, n_jobs=-1, cv=3)
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)
```

```

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```

Parameters currently in use:

```

{'activation': 'relu',
 'alpha': 0.0001,
 'batch_size': 'auto',
 'beta_1': 0.9,
 'beta_2': 0.999,
 'early_stopping': False,
 'epsilon': 1e-08,
 'hidden_layer_sizes': (100,),
 'learning_rate': 'constant',
 'learning_rate_init': 0.001,
 'max_fun': 15000,
 'max_iter': 100,
 'momentum': 0.9,
 'n_iter_no_change': 10,
 'nesterovs_momentum': True,
 'power_t': 0.5,
 'random_state': None,
 'shuffle': True,
 'solver': 'adam',
 'tol': 0.0001,
 'validation_fraction': 0.1,
 'verbose': False,
 'warm_start': False}
{'activation': ['tanh', 'relu'],
 'alpha': [0.0001, 0.05],
 'hidden_layer_sizes': [(50, 50, 50), (50, 100, 50), (100,)],
 'learning_rate': ['constant', 'adaptive'],
 'solver': ['sgd', 'adam']}

```

Confusion Matrix:

```

[[23  0  0]
 [ 0 23  4]
 [ 0  0 25]]
-----
```

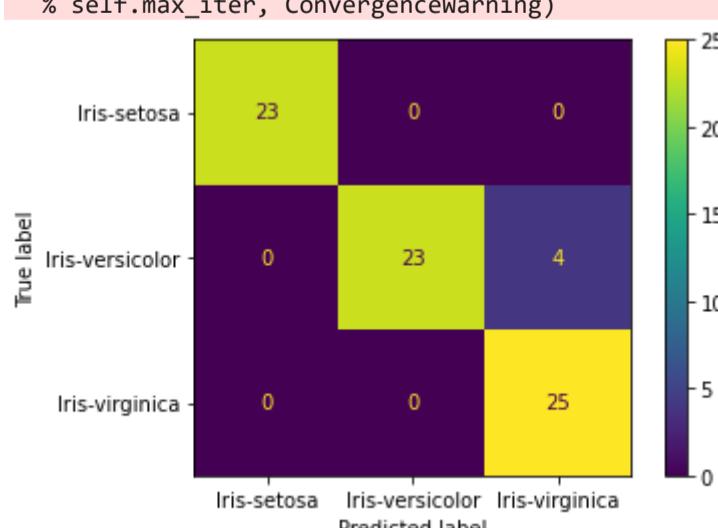
Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	23
Iris-versicolor	1.00	0.85	0.92	27
Iris-virginica	0.86	1.00	0.93	25
accuracy			0.95	75
macro avg	0.95	0.95	0.95	75
weighted avg	0.95	0.95	0.95	75

Accuracy:

0.9466666666666667

```
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (100) reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
```



In [ ]:

```
# IRIS PLANT DATASET
# Multi Layer Perceptron(With Tuning)[40-60 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data",header=None)

col_name = ['Sepal Length','Sepal Width','Petal Length','Petal Width','Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X,y,train_size=0.4,test_size=0.6,random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.neural_network import MLPClassifier
classifier = MLPClassifier(max_iter=100)

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

parameter_space = {
    'hidden_layer_sizes': [(50,50,50), (50,100,50), (100,)],
    'activation': ['tanh', 'relu'],
    'solver': ['sgd', 'adam'],
    'alpha': [0.0001, 0.05],
    'learning_rate': ['constant','adaptive'],
}
pprint(parameter_space)

#####

from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = MLPClassifier(max_iter=100)
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(classifier, parameter_space, n_jobs=-1, cv=3)
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
```

```

plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

Parameters currently in use:

{'activation': 'relu',
 'alpha': 0.0001,
 'batch_size': 'auto',
 'beta_1': 0.9,
 'beta_2': 0.999,
 'early_stopping': False,
 'epsilon': 1e-08,
 'hidden_layer_sizes': (100,),
 'learning_rate': 'constant',
 'learning_rate_init': 0.001,
 'max_fun': 15000,
 'max_iter': 100,
 'momentum': 0.9,
 'n_iter_no_change': 10,
 'nesterovs_momentum': True,
 'power_t': 0.5,
 'random_state': None,
 'shuffle': True,
 'solver': 'adam',
 'tol': 0.0001,
 'validation_fraction': 0.1,
 'verbose': False,
 'warm_start': False}
{'activation': ['tanh', 'relu'],
 'alpha': [0.0001, 0.05],
 'hidden_layer_sizes': [(50, 50, 50), (50, 100, 50), (100,)],
 'learning_rate': ['constant', 'adaptive'],
 'solver': ['sgd', 'adam']}
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (100) reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)

```

Confusion Matrix:

```

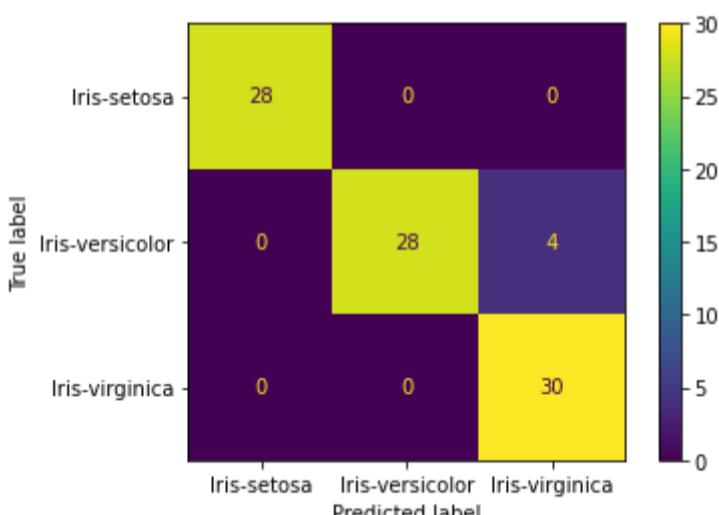
[[28  0  0]
 [ 0 28  4]
 [ 0  0 30]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	28
Iris-versicolor	1.00	0.88	0.93	32
Iris-virginica	0.88	1.00	0.94	30
accuracy			0.96	90
macro avg	0.96	0.96	0.96	90
weighted avg	0.96	0.96	0.96	90

Accuracy:

```
0.9555555555555556
```



In [ ]:

```

# IRIS PLANT DATASET
# Multi Layer Perceptron(With Tuning)[30-70 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

```

```

X_train, X_test, y_train, y_test = train_test_split(X,y,train_size=0.3,test_size=0.7,random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.neural_network import MLPClassifier
classifier = MLPClassifier(max_iter=100)

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

parameter_space = {
    'hidden_layer_sizes': [(50,50,50), (50,100,50), (100,)],
    'activation': ['tanh', 'relu'],
    'solver': ['sgd', 'adam'],
    'alpha': [0.0001, 0.05],
    'learning_rate': ['constant','adaptive'],
}
pprint(parameter_space)

#####

from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = MLPClassifier(max_iter=100)
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(classifier, parameter_space, n_jobs=-1, cv=3)
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```

Parameters currently in use:

```
{
'activation': 'relu',
'alpha': 0.0001,
'batch_size': 'auto',
'beta_1': 0.9,
'beta_2': 0.999,
'early_stopping': False,
'epsilon': 1e-08,
'hidden_layer_sizes': (100,),
'learning_rate': 'constant',
'learning_rate_init': 0.001,
'max_fun': 15000,
'max_iter': 100,
'momentum': 0.9,
'n_iter_no_change': 10,
'nesterovs_momentum': True,
'power_t': 0.5,
'random_state': None,
'shuffle': True,
'solver': 'adam',
}
```

```

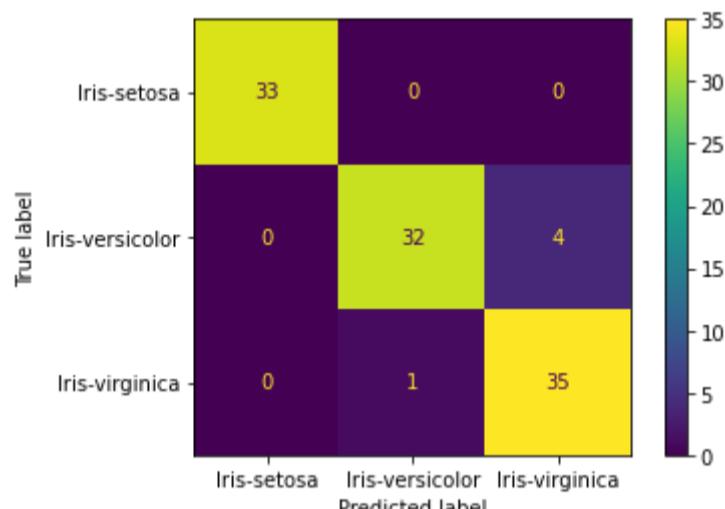
'tol': 0.0001,
'validation_fraction': 0.1,
'verbose': False,
'warm_start': False}
{'activation': ['tanh', 'relu'],
'alpha': [0.0001, 0.05],
'hidden_layer_sizes': [(50, 50, 50), (50, 100, 50), (100,)],
'learning_rate': ['constant', 'adaptive'],
'solver': ['sgd', 'adam']}
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (100) reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
Confusion Matrix:
[[33  0  0]
 [ 0 32  4]
 [ 0  1 35]]
-----
```

```

Performance Evaluation
      precision    recall   f1-score   support
Iris-setosa       1.00     1.00     1.00      33
Iris-versicolor   0.97     0.89     0.93      36
Iris-virginica    0.90     0.97     0.93      36

   accuracy          0.95
  macro avg       0.96     0.95     0.95      105
weighted avg     0.95     0.95     0.95      105
-----
```

Accuracy:  
0.9523809523809523



```

In [ ]:
# IRIS PLANT DATASET
# SVM(Without Tuning)[70-30 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, test_size=0.3, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.svm import SVC

classifier = SVC()
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))
```

```

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```

Confusion Matrix:

```

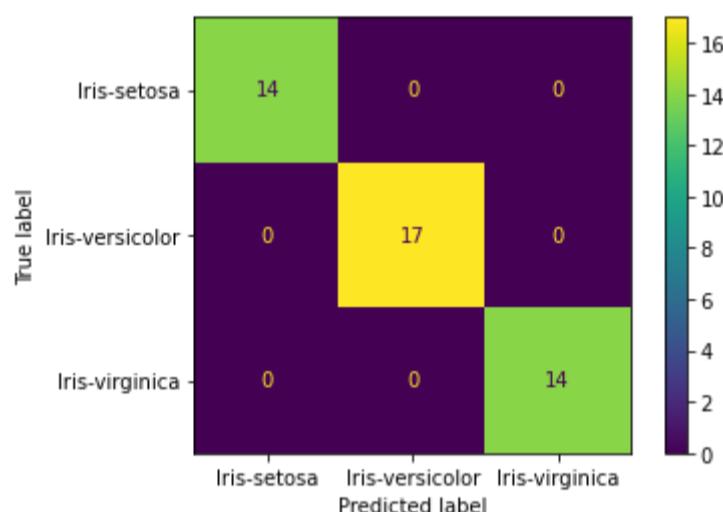
[[14  0  0]
 [ 0 17  0]
 [ 0  0 14]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	14
Iris-versicolor	1.00	1.00	1.00	17
Iris-virginica	1.00	1.00	1.00	14
accuracy			1.00	45
macro avg	1.00	1.00	1.00	45
weighted avg	1.00	1.00	1.00	45

Accuracy:

```
1.0
```



In [ ]:

```

# IRIS PLANT DATASET
# SVM(Without Tuning)[60-40 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.6, test_size=0.4, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.svm import SVC

classifier = SVC()
classifier.fit(X_train, y_train)

```

```

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```

Confusion Matrix:

```

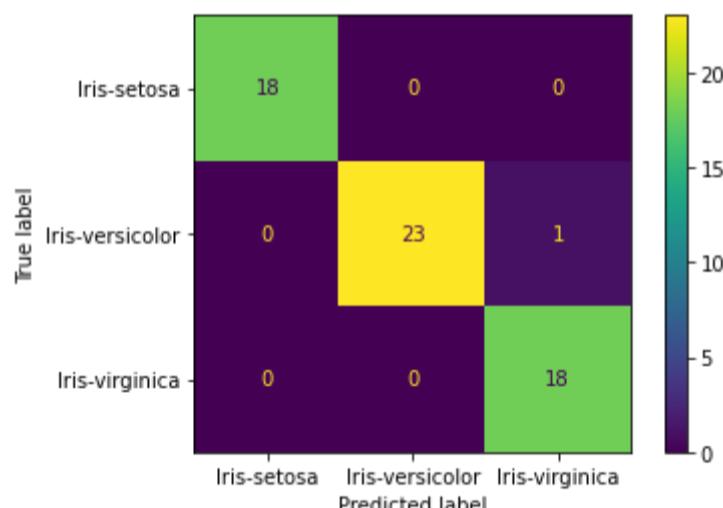
[[18  0  0]
 [ 0 23  1]
 [ 0  0 18]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	18
Iris-versicolor	1.00	0.96	0.98	24
Iris-virginica	0.95	1.00	0.97	18
accuracy			0.98	60
macro avg	0.98	0.99	0.98	60
weighted avg	0.98	0.98	0.98	60

Accuracy:

```
0.9833333333333333
```



In [ ]:

```

# IRIS PLANT DATASET
# SVM(Without Tuning)[50-50 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.5, test_size=0.5, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

```

```

# Classification
from sklearn.svm import SVC

classifier = SVC()
classifier.fit(X_train,y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```

Confusion Matrix:

```

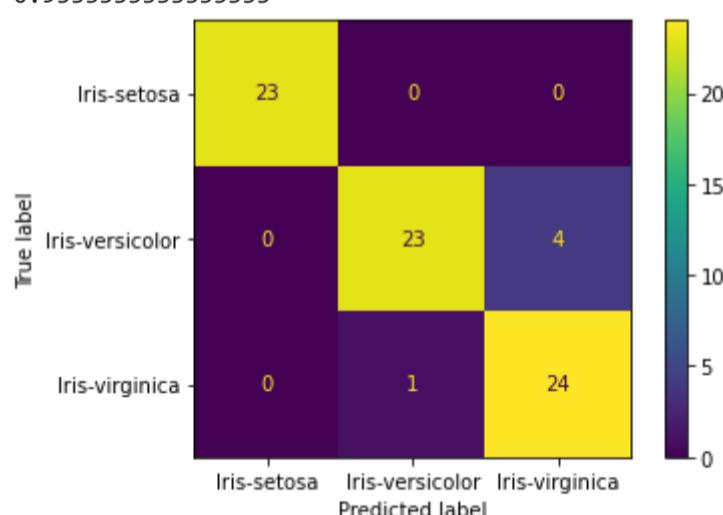
[[23  0  0]
 [ 0 23  4]
 [ 0  1 24]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	23
Iris-versicolor	0.96	0.85	0.90	27
Iris-virginica	0.86	0.96	0.91	25
accuracy			0.93	75
macro avg	0.94	0.94	0.94	75
weighted avg	0.94	0.93	0.93	75

Accuracy:

```
0.9333333333333333
```



In [ ]:

```

# IRIS PLANT DATASET
# SVM(Without Tuning)[40-60 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.4, test_size=0.6, random_state=10)

```

```

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.svm import SVC

classifier = SVC()
classifier.fit(X_train,y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```

Confusion Matrix:

```

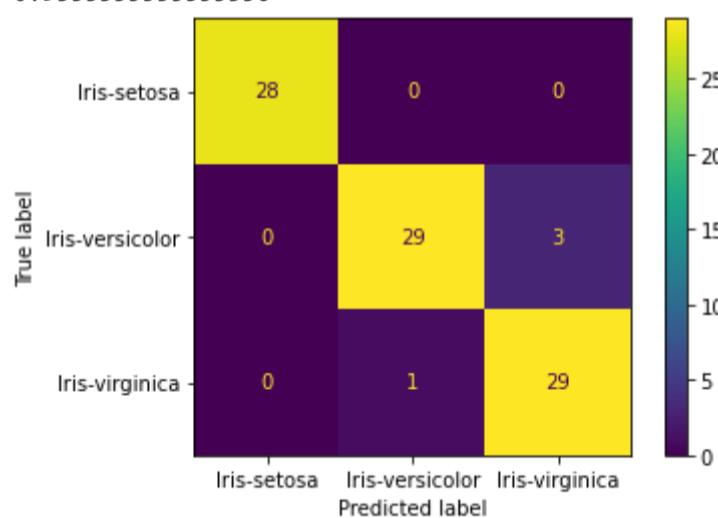
[[28  0  0]
 [ 0 29  3]
 [ 0  1 29]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	28
Iris-versicolor	0.97	0.91	0.94	32
Iris-virginica	0.91	0.97	0.94	30
accuracy			0.96	90
macro avg	0.96	0.96	0.96	90
weighted avg	0.96	0.96	0.96	90

Accuracy:

0.9555555555555556



In [ ]:

```

# IRIS PLANT DATASET
# SVM(Without Tuning)[30-70 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data",header=None)

col_name = ['Sepal Length','Sepal Width','Petal Length','Petal Width','Class']

df.columns = col_name

```

```

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X,y,train_size=0.3,test_size=0.7,random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.svm import SVC

classifier = SVC()
classifier.fit(X_train,y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```

Confusion Matrix:

```

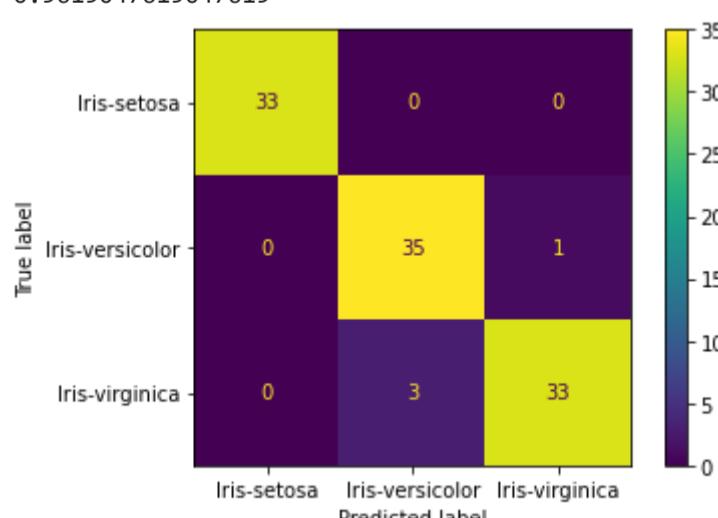
[[33  0  0]
 [ 0 35  1]
 [ 0  3 33]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	33
Iris-versicolor	0.92	0.97	0.95	36
Iris-virginica	0.97	0.92	0.94	36
accuracy			0.96	105
macro avg	0.96	0.96	0.96	105
weighted avg	0.96	0.96	0.96	105

Accuracy:

0.9619047619047619



In [ ]:

```

# IRIS PLANT DATASET
# SVM(With Tuning)[70-30 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

```

```

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X,y,train_size=0.7,test_size=0.3,random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.svm import SVC

classifier = SVC()

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

param_grid = {'C': [0.1, 1, 10, 100], 'gamma': [1, 0.1, 0.01, 0.001], 'kernel': ['rbf', 'poly', 'sigmoid']}
pprint(param_grid)

#####

from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = SVC()
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(SVC(), param_grid, refit=True, verbose=2)
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```

Parameters currently in use:

```

{'C': 1.0,
 'break_ties': False,
 'cache_size': 200,
 'class_weight': None,
 'coef0': 0.0,
 'decision_function_shape': 'ovr',
 'degree': 3,
 'gamma': 'scale',
 'kernel': 'rbf',
 'max_iter': -1,
 'probability': False,
 'random_state': null,
 'tol': 0.001,
 'verbose': False}

```







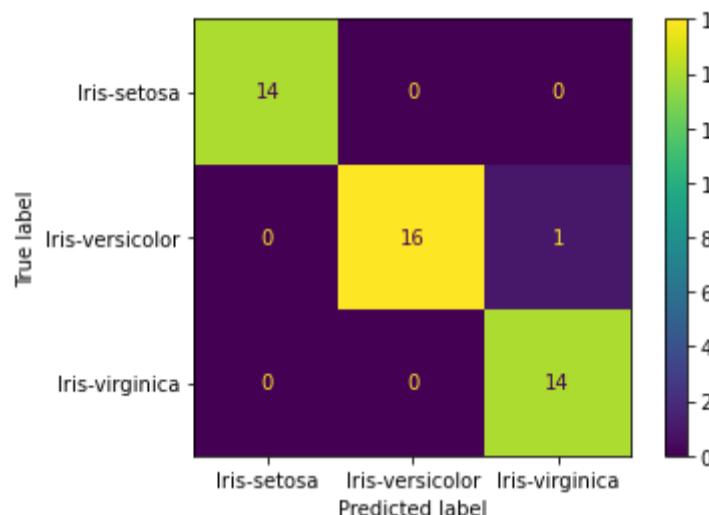


## Performance Evaluation precision

Iris-setosa	1.00	1.00	1.00	14
Iris-versicolor	1.00	0.94	0.97	17
Iris-virginica	0.93	1.00	0.97	14
accuracy			0.98	45
macro avg	0.98	0.98	0.98	45
weighted avg	0.98	0.98	0.98	45

Accuracy:  
0.9777777777777777

[Parallel(n\_jobs=1)]: Done 240 out of 240 | elapsed: 0.9s finished



In [ ]:

```
# IRIS PLANT DATASET
# SVM(With Tuning)[60-40 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.6, test_size=0.4, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.svm import SVC

classifier = SVC()

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

param_grid = {'C': [0.1, 1, 10, 100], 'gamma': [1, 0.1, 0.01, 0.001], 'kernel': ['rbf', 'poly', 'sigmoid']}
pprint(param_grid)

#####
from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = SVC()
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(SVC(), param_grid, refit=True, verbose=2)
```

```
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()
```

Parameters currently in use:









Confusion Matrix:  
[[18 0 0]

$$\begin{bmatrix} 18 & 0 & 0 \\ 0 & 21 & 3 \\ 0 & 0 & 18 \end{bmatrix}$$

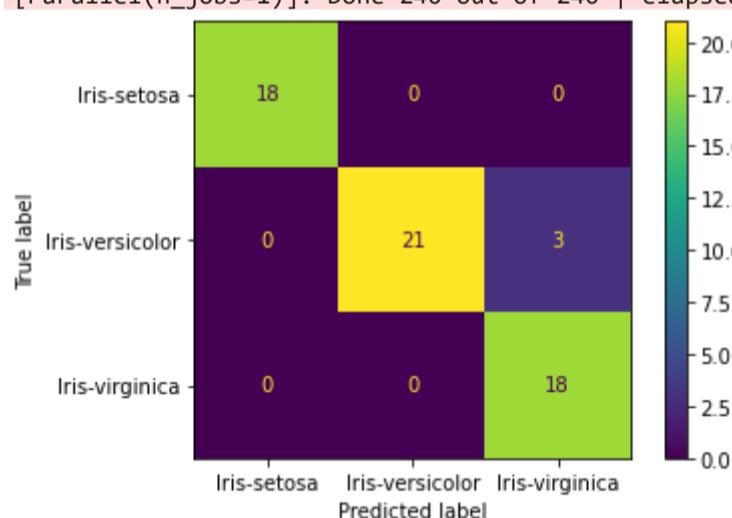
Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	18
Iris-versicolor	1.00	0.88	0.93	24
Iris-virginica	0.86	1.00	0.92	18
accuracy			0.95	60
macro avg	0.95	0.96	0.95	60
weighted avg	0.96	0.95	0.95	60

Accuracy:

Accur.  
0.95

[Parallel(n\_jobs=1)]: Done 240 out of 240 | elapsed: 1.0s finished



In [ ]:

```
# IRIS PLANT DATASET  
# SVM(with Tuning)[50-50 split]
```

```
import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.5, test_size=0.5)
```

```

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.svm import SVC

classifier = SVC()

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

param_grid = {'C': [0.1, 1, 10, 100], 'gamma': [1, 0.1, 0.01, 0.001], 'kernel': ['rbf', 'poly', 'sigmoid']}
pprint(param_grid)

#####

from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = SVC()
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(SVC(), param_grid, refit=True, verbose=2)
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```

Parameters currently in use:

```

{'C': 1.0,
 'break_ties': False,
 'cache_size': 200,
 'class_weight': None,
 'coef0': 0.0,
 'decision_function_shape': 'ovr',
 'degree': 3,
 'gamma': 'scale',
 'kernel': 'rbf',
 'max_iter': -1,
 'probability': False,
 'random_state': None,
 'shrinking': True,
 'tol': 0.001,
 'verbose': False}
{'C': [0.1, 1, 10, 100],
 'gamma': [1, 0.1, 0.01, 0.001],
 'kernel': ['rbf', 'poly', 'sigmoid']}
Fitting 5 folds for each of 48 candidates, totalling 240 fits
[CV] C=0.1, gamma=1, kernel=rbf .....
[CV] ..... C=0.1, gamma=1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=1, kernel=rbf .....
[CV] ..... C=0.1, gamma=1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=1, kernel=rbf .....

```

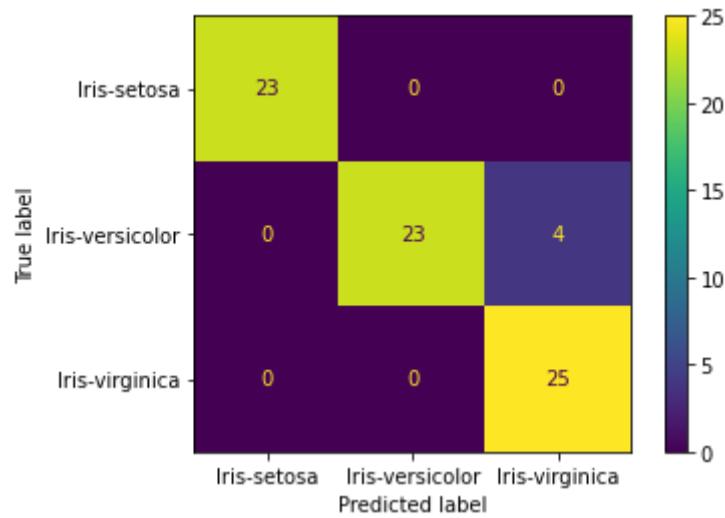












```
In [ ]:
# IRIS PLANT DATASET
# SVM(With Tuning)[40-60 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.4, test_size=0.6, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.svm import SVC

classifier = SVC()

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

param_grid = {'C': [0.1, 1, 10, 100], 'gamma': [1, 0.1, 0.01, 0.001], 'kernel': ['rbf', 'poly', 'sigmoid']}
pprint(param_grid)

#####

from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = SVC()
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(SVC(), param_grid, refit=True, verbose=2)
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")
```

```
print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-" * 50)
print("-" * 50)

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()
```

Parameters currently in use:

```
{'C': 1.0,
 'break_ties': False,
 'cache_size': 200,
 'class_weight': None,
 'coef0': 0.0,
 'decision_function_shape': 'ovr',
 'degree': 3,
 'gamma': 'scale',
 'kernel': 'rbf',
 'max_iter': -1,
 'probability': False,
 'random_state': None,
 'shrinking': True,
 'tol': 0.001,
 'verbose': False}
{'C': [0.1, 1, 10, 100],
 'gamma': [1, 0.1, 0.01, 0.001],
 'kernel': ['rbf', 'poly', 'sigmoid']}
Fitting 5 folds for each of 48 candidates, totalling 240 fits
[CV] C=0.1, gamma=1, kernel=rbf .....
[CV] ..... C=0.1, gamma=1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=1, kernel=rbf .....
[CV] ..... C=0.1, gamma=1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=1, kernel=rbf .....
[CV] ..... C=0.1, gamma=1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=1, kernel=rbf .....
[CV] ..... C=0.1, gamma=1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=1, kernel=rbf .....
[CV] ..... C=0.1, gamma=1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=1, kernel=rbf .....
[CV] ..... C=0.1, gamma=1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=1, kernel=rbf .....
[CV] ..... C=0.1, gamma=1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=1, kernel=rbf .....
[CV] ..... C=0.1, gamma=1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=1, kernel=poly .....
[CV] ..... C=0.1, gamma=1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=1, kernel=poly .....
[CV] ..... C=0.1, gamma=1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=1, kernel=poly .....
[CV] ..... C=0.1, gamma=1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=1, kernel=poly .....
[CV] ..... C=0.1, gamma=1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=1, kernel=poly .....
[CV] ..... C=0.1, gamma=1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=1, kernel=poly .....
[CV] ..... C=0.1, gamma=1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=1, kernel=poly .....
[CV] ..... C=0.1, gamma=1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=1, kernel=poly .....
[CV] ..... C=0.1, gamma=1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.1, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=poly .....
[CV] ..... C=0.1, gamma=0.1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=poly .....
[CV] ..... C=0.1, gamma=0.1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=poly .....
[CV] ..... C=0.1, gamma=0.1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=poly .....
[CV] ..... C=0.1, gamma=0.1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=poly .....
[CV] ..... C=0.1, gamma=0.1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=poly .....
[CV] ..... C=0.1, gamma=0.1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=poly .....
[CV] ..... C=0.1, gamma=0.1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=poly .....
[CV] ..... C=0.1, gamma=0.1, kernel=poly, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=0.1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=0.1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=0.1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=0.1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=0.1, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=0.1, kernel=sigmoid, total= 0.0s
[CV] C=0.1, gamma=0.01, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.01, kernel=rbf, total= 0.0s
[CV] C=0.1, gamma=0.01, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.01, kernel=rbf, total= 0.0s
```









```
[CV] C=100, gamma=0.001, kernel=poly .....  

[CV] ..... C=100, gamma=0.001, kernel=poly, total= 0.0s  

[CV] C=100, gamma=0.001, kernel=poly .....  

[CV] ..... C=100, gamma=0.001, kernel=poly, total= 0.0s  

[CV] C=100, gamma=0.001, kernel=poly .....  

[CV] ..... C=100, gamma=0.001, kernel=poly, total= 0.0s  

[CV] C=100, gamma=0.001, kernel=poly .....  

[CV] ..... C=100, gamma=0.001, kernel=poly, total= 0.0s  

[CV] C=100, gamma=0.001, kernel=poly .....  

[CV] ..... C=100, gamma=0.001, kernel=poly, total= 0.0s  

[CV] C=100, gamma=0.001, kernel=sigmoid .....  

[CV] ..... C=100, gamma=0.001, kernel=sigmoid, total= 0.0s  

[CV] C=100, gamma=0.001, kernel=sigmoid .....  

[CV] ..... C=100, gamma=0.001, kernel=sigmoid, total= 0.0s  

[CV] C=100, gamma=0.001, kernel=sigmoid .....  

[CV] ..... C=100, gamma=0.001, kernel=sigmoid, total= 0.0s  

[CV] C=100, gamma=0.001, kernel=sigmoid .....  

[CV] ..... C=100, gamma=0.001, kernel=sigmoid, total= 0.0s  

[CV] C=100, gamma=0.001, kernel=sigmoid .....  

[CV] ..... C=100, gamma=0.001, kernel=sigmoid, total= 0.0s  

[CV] C=100, gamma=0.001, kernel=sigmoid .....  

[CV] ..... C=100, gamma=0.001, kernel=sigmoid, total= 0.0s  

Confusion Matrix:  

[[28  0  0]  

 [ 0 27  5]  

 [ 0  2 28]]
```

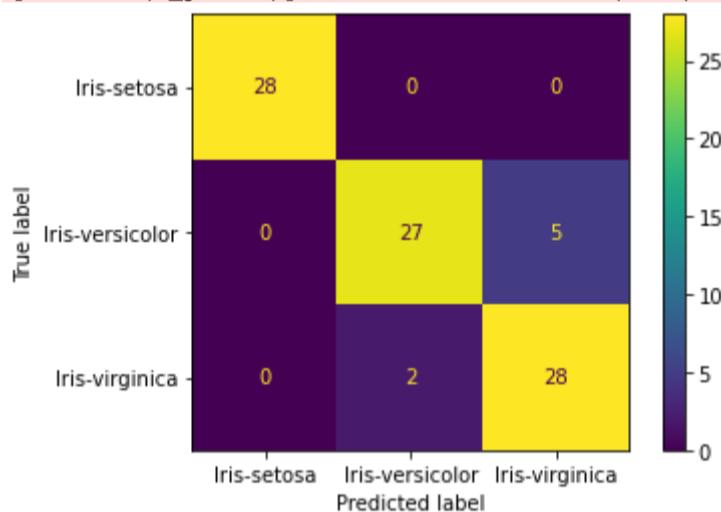
#### Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	28
Iris-versicolor	0.93	0.84	0.89	32
Iris-virginica	0.85	0.93	0.89	30
accuracy			0.92	90
macro avg	0.93	0.93	0.92	90
weighted avg	0.92	0.92	0.92	90

#### Accuracy:

0.9222222222222223

[Parallel(n\_jobs=1)]: Done 240 out of 240 | elapsed: 0.9s finished



In [ ]:

```
# IRIS PLANT DATASET
# SVM(With Tuning)[30-70 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.3, test_size=0.7, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Classification
from sklearn.svm import SVC

classifier = SVC()
```

```
#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

param_grid = {'C': [0.1, 1, 10, 100], 'gamma': [1, 0.1, 0.01, 0.001], 'kernel': ['rbf', 'poly', 'sigmoid']}

pprint(param_grid)

#####

from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = SVC()
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(SVC(), param_grid, refit=True, verbose=2)
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()
```

### Parameters currently in use:

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done    1 out of    1 | elapsed:    0.0s remaining:    0.0s
[CV] ..... C=0.1, gamma=0.001, kernel=rbf, total=  0.0s
[CV] C=0.1, gamma=0.001, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.001, kernel=rbf, total=  0.0s
[CV] C=0.1, gamma=0.001, kernel=rbf .....
[CV] ..... C=0.1, gamma=0.001, kernel=rbf, total=  0.0s
[CV] C=0.1, gamma=0.001, kernel=poly .....
[CV] ..... C=0.1, gamma=0.001, kernel=poly, total=  0.0s
[CV] C=0.1, gamma=0.001, kernel=poly .....
[CV] ..... C=0.1, gamma=0.001, kernel=poly, total=  0.0s
[CV] C=0.1, gamma=0.001, kernel=poly .....
[CV] ..... C=0.1, gamma=0.001, kernel=poly, total=  0.0s
[CV] C=0.1, gamma=0.001, kernel=poly .....
[CV] ..... C=0.1, gamma=0.001, kernel=poly, total=  0.0s
[CV] C=0.1, gamma=0.001, kernel=poly .....
[CV] ..... C=0.1, gamma=0.001, kernel=poly, total=  0.0s
[CV] C=0.1, gamma=0.001, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=0.001, kernel=sigmoid, total=  0.0s
[CV] C=0.1, gamma=0.001, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=0.001, kernel=sigmoid, total=  0.0s
[CV] C=0.1, gamma=0.001, kernel=sigmoid .....
[CV] ..... C=0.1, gamma=0.001, kernel=sigmoid, total=  0.0s
[CV] C=0.1, gamma=0.001, kernel=sigmoid .....
```







[CV] .....  
Confusion Matrix:

```

[[33  0  0]
 [ 0 33  3]
 [ 0  4 32]]

```

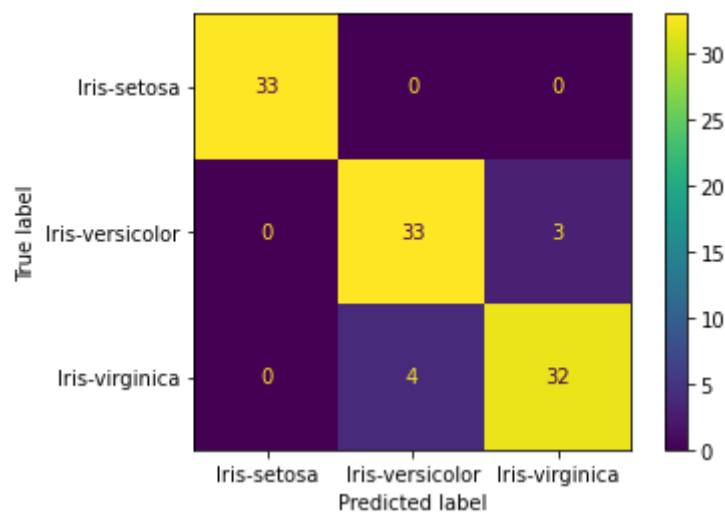
Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	33
Iris-versicolor	0.89	0.92	0.90	36
Iris-virginica	0.91	0.89	0.90	36
accuracy			0.93	105
macro avg	0.94	0.94	0.94	105
weighted avg	0.93	0.93	0.93	105

Accuracy:

Accuracy:  
0.9333333333333333

[Parallel(n\_jobs=1)]: Done 240 out of 240 | elapsed: 0.9s finished



```
In [ ]: #####
```

```
In [ ]: # IRIS PLANT DATASET
# Random Forest Classifier(Without Tuning)[70-30 split]
```

```
import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, test_size=0.3, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Finding the important parameters that contribute to most of the variance in the data.

import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.decomposition import PCA

pca_test = PCA(n_components=4)
pca_test.fit(X_train)
sns.set(style='whitegrid')
plt.plot(np.cumsum(pca_test.explained_variance_ratio_))
plt.xlabel('number of components')
plt.ylabel('cumulative explained variance')
plt.axvline(linewidth=4, color='r', linestyle = '--', x=10, ymin=0, ymax=1)
display(plt.show())

# So we can see that we have 10 important parameters

pca = PCA(n_components=2)
pca.fit(X_train)
X_train = pca.transform(X_train)
X_test = pca.transform(X_test)

# Classification
from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier(n_estimators=20, random_state=0)
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
```

```

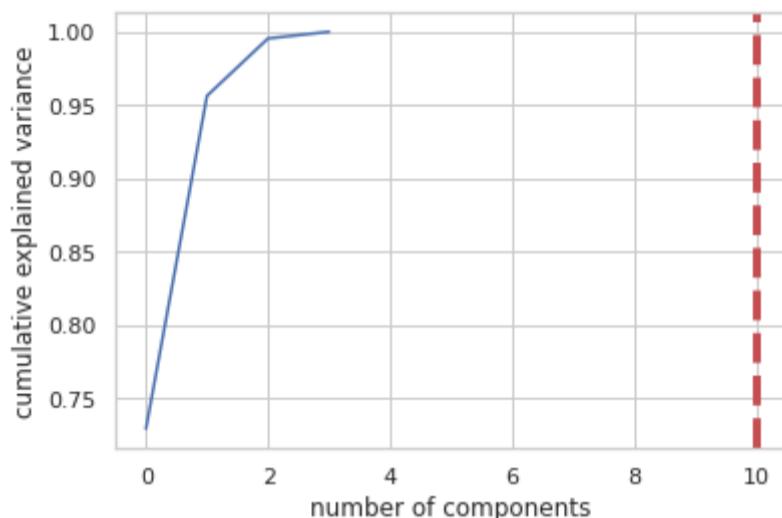
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```



None

Confusion Matrix:

```
[[14  0  0]
 [ 0 14  3]
 [ 0  0 14]]
```

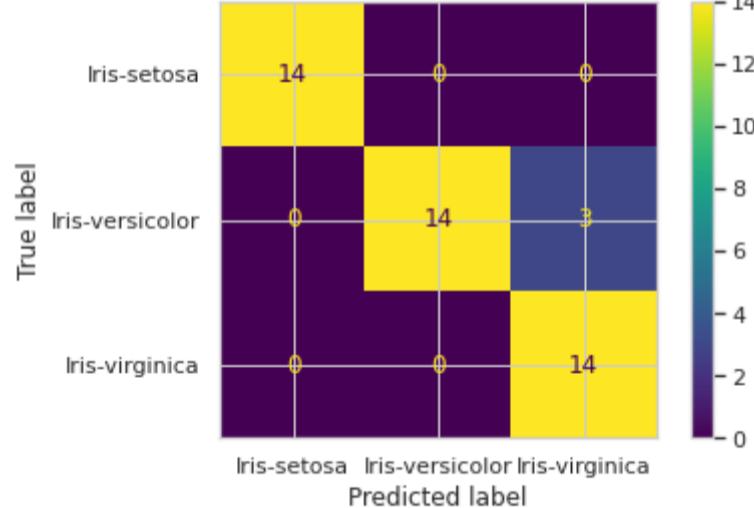
-----

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	14
Iris-versicolor	1.00	0.82	0.90	17
Iris-virginica	0.82	1.00	0.90	14
accuracy			0.93	45
macro avg	0.94	0.94	0.94	45
weighted avg	0.95	0.93	0.93	45

-----

Accuracy:

0.9333333333333333



In [ ]:

```

# IRIS PLANT DATASET
# Random Forest Classifier(With Tuning)[70-30 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, test_size=0.3, random_state=10)

```

```

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Finding the important parameters that contribute to most of the variance in the data.

import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.decomposition import PCA

pca_test = PCA(n_components=4)
pca_test.fit(X_train)
sns.set(style='whitegrid')
plt.plot(np.cumsum(pca_test.explained_variance_ratio_))
plt.xlabel('number of components')
plt.ylabel('cumulative explained variance')
plt.axvline(linewidth=4, color='r', linestyle = '--', x=10, ymin=0, ymax=1)
display(plt.show())

# So we can see that we have 10 important parameters

pca = PCA(n_components=2)
pca.fit(X_train)
X_train = pca.transform(X_train)
X_test = pca.transform(X_test)

# Classification
from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier()

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

from sklearn.model_selection import RandomizedSearchCV
# Number of trees in random forest
n_estimators = [int(x) for x in np.linspace(start = 200, stop = 2000, num = 10)]
# Number of features to consider at every split
max_features = ['auto', 'sqrt']
# Maximum number of levels in tree
max_depth = [int(x) for x in np.linspace(10, 110, num = 11)]
max_depth.append(None)
# Minimum number of samples required to split a node
min_samples_split = [2, 5, 10]
# Minimum number of samples required at each Leaf node
min_samples_leaf = [1, 2, 4]
# Method of selecting samples for training each tree
bootstrap = [True, False]
# Create the random grid
random_grid = {'n_estimators': n_estimators,
               'max_features': max_features,
               'max_depth': max_depth,
               'min_samples_split': min_samples_split,
               'min_samples_leaf': min_samples_leaf,
               'bootstrap': bootstrap}
pprint(random_grid)

#####

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = RandomForestClassifier()
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores
rf_random = RandomizedSearchCV(estimator = classifier, param_distributions = random_grid, n_iter = 100, cv = 3, verbose=2, random_
# Fit the random search model
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")

```

```

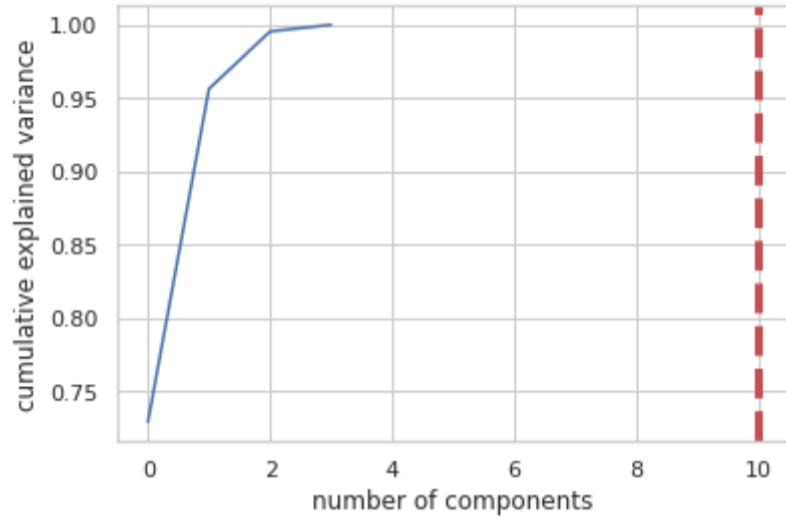
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```



None

Parameters currently in use:

```

{'bootstrap': True,
 'ccp_alpha': 0.0,
 'class_weight': None,
 'criterion': 'gini',
 'max_depth': None,
 'max_features': 'auto',
 'max_leaf_nodes': None,
 'max_samples': None,
 'min_impurity_decrease': 0.0,
 'min_impurity_split': None,
 'min_samples_leaf': 1,
 'min_samples_split': 2,
 'min_weight_fraction_leaf': 0.0,
 'n_estimators': 100,
 'n_jobs': None,
 'oob_score': False,
 'random_state': None,
 'verbose': 0,
 'warm_start': False}
{'bootstrap': [True, False],
 'max_depth': [10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, None],
 'max_features': ['auto', 'sqrt'],
 'min_samples_leaf': [1, 2, 4],
 'min_samples_split': [2, 5, 10],
 'n_estimators': [200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000]}
Fitting 3 folds for each of 100 candidates, totalling 300 fits

```

```

[Parallel(n_jobs=-1)]: Using backend LokyBackend with 2 concurrent workers.
[Parallel(n_jobs=-1)]: Done  37 tasks      | elapsed:   44.7s
[Parallel(n_jobs=-1)]: Done 158 tasks      | elapsed:  3.0min
[Parallel(n_jobs=-1)]: Done 300 out of 300 | elapsed:  5.8min finished

```

Confusion Matrix:

```

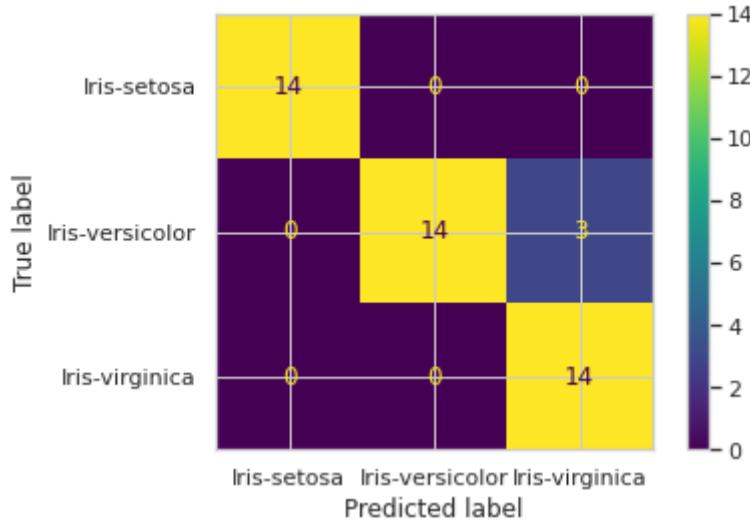
[[14  0  0]
 [ 0 14  3]
 [ 0  0 14]]
-----
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	14
Iris-versicolor	1.00	0.82	0.90	17
Iris-virginica	0.82	1.00	0.90	14
accuracy			0.93	45
macro avg	0.94	0.94	0.94	45
weighted avg	0.95	0.93	0.93	45

Accuracy:

0.9333333333333333



```
In [ ]:
# IRIS PLANT DATASET
# Multi Layer Perceptron(Without Tuning)[70-30 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, test_size=0.30, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Finding the important parameters that contribute to most of the variance in the data.

import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.decomposition import PCA

pca_test = PCA(n_components=4)
pca_test.fit(X_train)
sns.set(style='whitegrid')
plt.plot(np.cumsum(pca_test.explained_variance_ratio_))
plt.xlabel('number of components')
plt.ylabel('cumulative explained variance')
plt.axvline(width=4, color='r', linestyle = '--', x=10, ymin=0, ymax=1)
display(plt.show())

# So we can see that we have 10 important parameters

pca = PCA(n_components=2)
pca.fit(X_train)
X_train = pca.transform(X_train)
X_test = pca.transform(X_test)

# Classification using MLP
from sklearn.neural_network import MLPClassifier

classifier = MLPClassifier()
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))
```

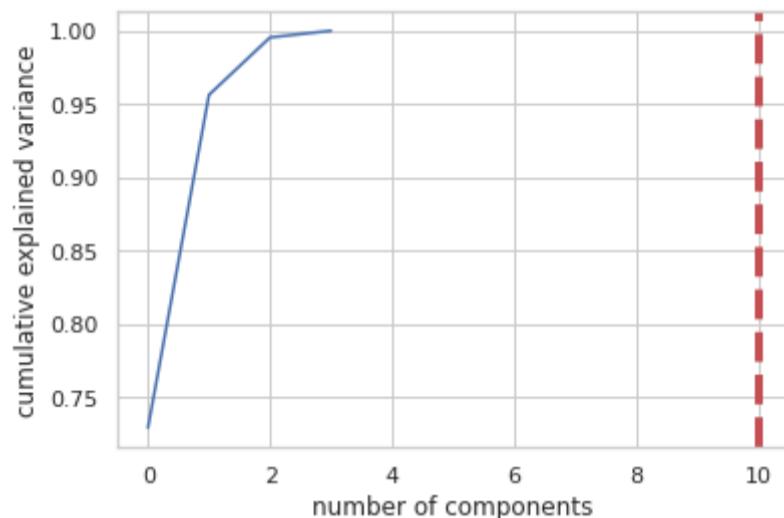
```

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```



None

```
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200) reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
```

Confusion Matrix:

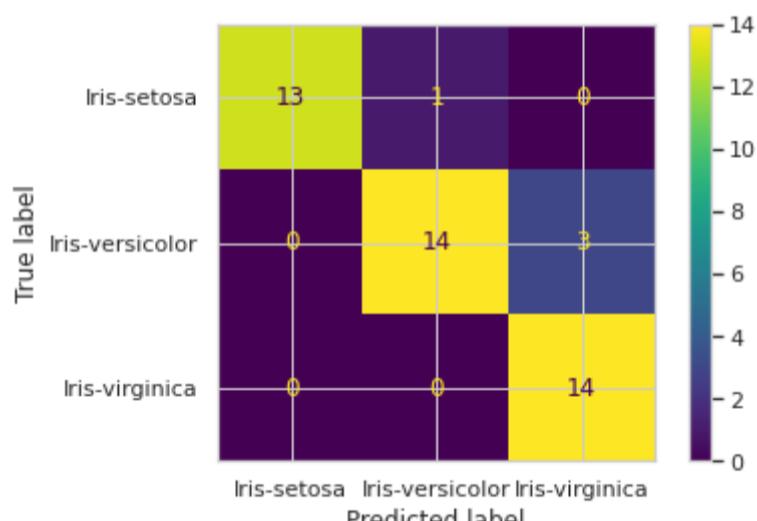
```
[[13  1  0]
 [ 0 14  3]
 [ 0  0 14]]
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	0.93	0.96	14
Iris-versicolor	0.93	0.82	0.87	17
Iris-virginica	0.82	1.00	0.90	14
accuracy			0.91	45
macro avg	0.92	0.92	0.91	45
weighted avg	0.92	0.91	0.91	45

Accuracy:

```
0.9111111111111111
```



In [ ]:

```

# IRIS PLANT DATASET
# Multi Layer Perceptron(With Tuning)[70-30 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, test_size=0.3, random_state=10)

```

```

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Finding the important parameters that contribute to most of the variance in the data.

import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.decomposition import PCA

pca_test = PCA(n_components=4)
pca_test.fit(X_train)
sns.set(style='whitegrid')
plt.plot(np.cumsum(pca_test.explained_variance_ratio_))
plt.xlabel('number of components')
plt.ylabel('cumulative explained variance')
plt.axvline(linewidth=4, color='r', linestyle = '--', x=10, ymin=0, ymax=1)
display(plt.show())

# So we can see that we have 10 important parameters

pca = PCA(n_components=2)
pca.fit(X_train)
X_train = pca.transform(X_train)
X_test = pca.transform(X_test)

# Classification
from sklearn.neural_network import MLPClassifier
classifier = MLPClassifier(max_iter=100)

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

parameter_space = {
    'hidden_layer_sizes': [(50,50,50), (50,100,50), (100,)],
    'activation': ['tanh', 'relu'],
    'solver': ['sgd', 'adam'],
    'alpha': [0.0001, 0.05],
    'learning_rate': ['constant','adaptive'],
}
pprint(parameter_space)

#####

from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = MLPClassifier(max_iter=100)
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(classifier, parameter_space, n_jobs=-1, cv=3)
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

print("Performance Evaluation")
print(classification_report(y_test, y_pred))

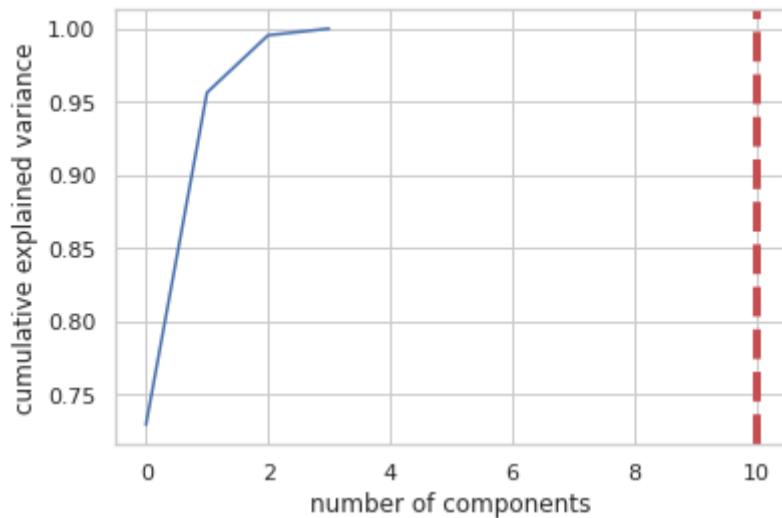
print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix

```

```
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()
```



None

Parameters currently in use:

```
{'activation': 'relu',
 'alpha': 0.0001,
 'batch_size': 'auto',
 'beta_1': 0.9,
 'beta_2': 0.999,
 'early_stopping': False,
 'epsilon': 1e-08,
 'hidden_layer_sizes': (100,),
 'learning_rate': 'constant',
 'learning_rate_init': 0.001,
 'max_fun': 15000,
 'max_iter': 100,
 'momentum': 0.9,
 'n_iter_no_change': 10,
 'nesterovs_momentum': True,
 'power_t': 0.5,
 'random_state': None,
 'shuffle': True,
 'solver': 'adam',
 'tol': 0.0001,
 'validation_fraction': 0.1,
 'verbose': False,
 'warm_start': False}
{'activation': ['tanh', 'relu'],
 'alpha': [0.0001, 0.05],
 'hidden_layer_sizes': [(50, 50, 50), (50, 100, 50), (100,)],
 'learning_rate': ['constant', 'adaptive'],
 'solver': ['sgd', 'adam']}
```

```
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:571: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (100) reached and the optimization hasn't converged yet.
  % self.max_iter, ConvergenceWarning)
```

Confusion Matrix:

```
[[13  1  0]
 [ 0 14  3]
 [ 0  0 14]]
```

-----

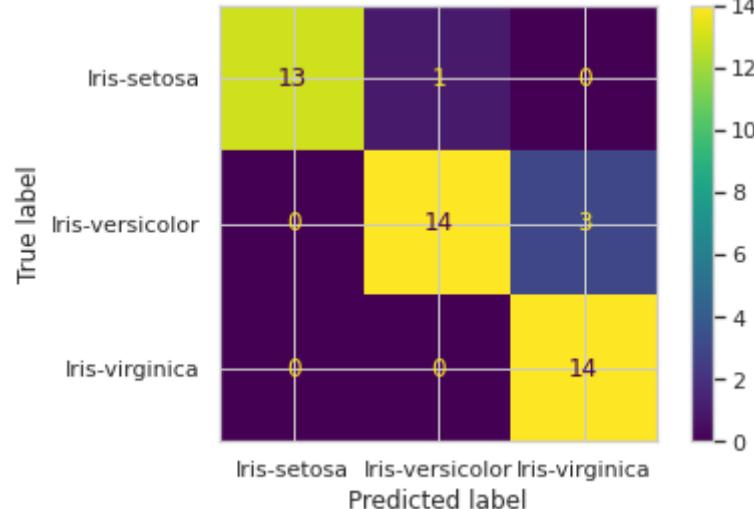
Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	0.93	0.96	14
Iris-versicolor	0.93	0.82	0.87	17
Iris-virginica	0.82	1.00	0.90	14
accuracy			0.91	45
macro avg	0.92	0.92	0.91	45
weighted avg	0.92	0.91	0.91	45

-----

Accuracy:

```
0.9111111111111111
```



In [ ]:

```
# IRIS PLANT DATASET
# SVM(Without Tuning)[70-30 split]
```

```

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, test_size=0.3, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Finding the important parameters that contribute to most of the variance in the data.

import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.decomposition import PCA

pca_test = PCA(n_components=4)
pca_test.fit(X_train)
sns.set(style='whitegrid')
plt.plot(np.cumsum(pca_test.explained_variance_ratio_))
plt.xlabel('number of components')
plt.ylabel('cumulative explained variance')
plt.axvline(linewidth=4, color='r', linestyle = '--', x=10, ymin=0, ymax=1)
display(plt.show())

# So we can see that we have 10 important parameters

pca = PCA(n_components=2)
pca.fit(X_train)
X_train = pca.transform(X_train)
X_test = pca.transform(X_test)

# Classification
from sklearn.svm import SVC

classifier = SVC()
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

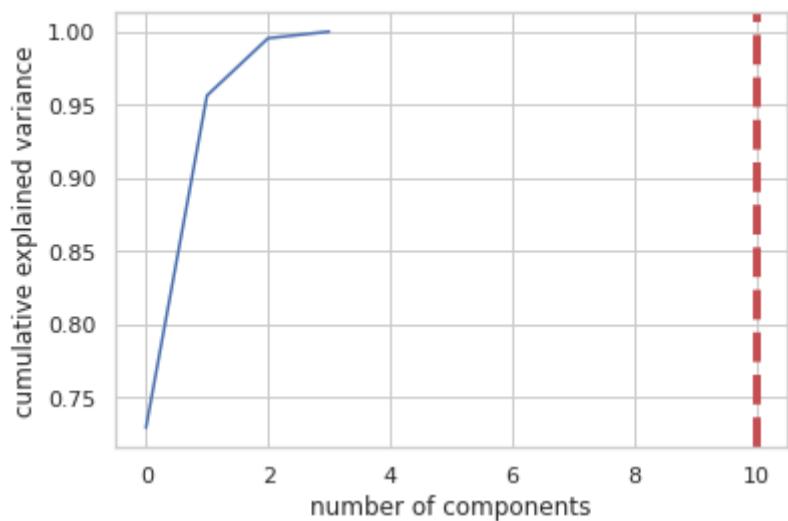
print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(classifier, X_test, y_test)
plt.show()

```



None

Confusion Matrix:

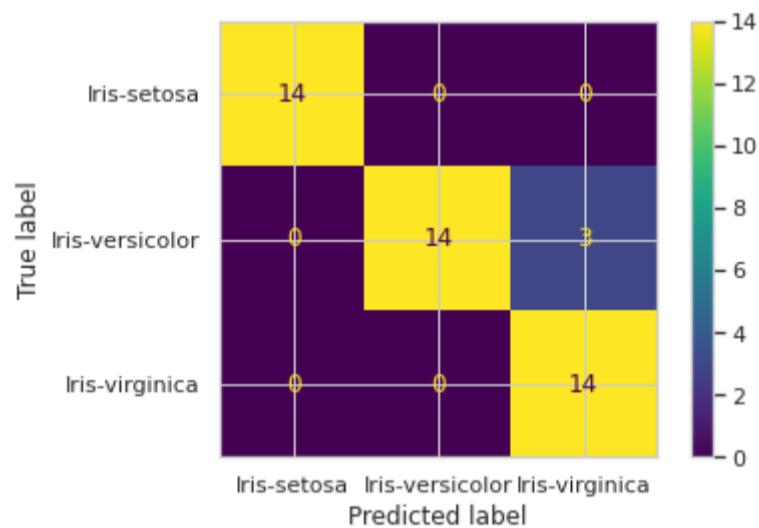
```
[[14  0  0]
 [ 0 14  3]
 [ 0  0 14]]
```

Performance Evaluation

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	14
Iris-versicolor	1.00	0.82	0.90	17
Iris-virginica	0.82	1.00	0.90	14
accuracy			0.93	45
macro avg	0.94	0.94	0.94	45
weighted avg	0.95	0.93	0.93	45

Accuracy:

0.9333333333333333



In [ ]:

```
# IRIS PLANT DATASET
# SVM(With Tuning)[70-30 split]

import pandas as pd
import numpy as np

# Dataset Preparation
df = pd.read_csv("iris.data", header=None)

col_name = ['Sepal Length', 'Sepal Width', 'Petal Length', 'Petal Width', 'Class']

df.columns = col_name

X = df.drop(['Class'], axis=1)
y = df['Class']

from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, train_size=0.7, test_size=0.3, random_state=10)

# Feature Scaling
from sklearn.preprocessing import StandardScaler

sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)

# Finding the important parameters that contribute to most of the variance in the data.

import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.decomposition import PCA
```

```

pca_test = PCA(n_components=4)
pca_test.fit(X_train)
sns.set(style='whitegrid')
plt.plot(np.cumsum(pca_test.explained_variance_ratio_))
plt.xlabel('number of components')
plt.ylabel('cumulative explained variance')
plt.axvline(linewidth=4, color='r', linestyle = '--', x=10, ymin=0, ymax=1)
display(plt.show())

# So we can see that we have 10 important parameters

pca = PCA(n_components=2)
pca.fit(X_train)
X_train = pca.transform(X_train)
X_test = pca.transform(X_test)

# Classification
from sklearn.svm import SVC

classifier = SVC()

#####
# Showing all the parameters

from pprint import pprint
# Look at parameters used by our current forest
print('Parameters currently in use:\n')
pprint(classifier.get_params())

#####
# Creating a set of important sample features

param_grid = {'C': [0.1, 1, 10, 100], 'gamma': [1, 0.1, 0.01, 0.001], 'kernel': ['rbf', 'poly', 'sigmoid']}
pprint(param_grid)

#####

from sklearn.model_selection import GridSearchCV

# Use the random grid to search for best hyperparameters
# First create the base model to tune
classifier = SVC()
# Random search of parameters, using 3 fold cross validation,
# search across 100 different combinations, and use all available cores

rf_random = GridSearchCV(SVC(), param_grid, refit=True, verbose=2)
rf_random.fit(X_train, y_train)

y_pred = rf_random.predict(X_test)

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))

print("-----")
print("-----")

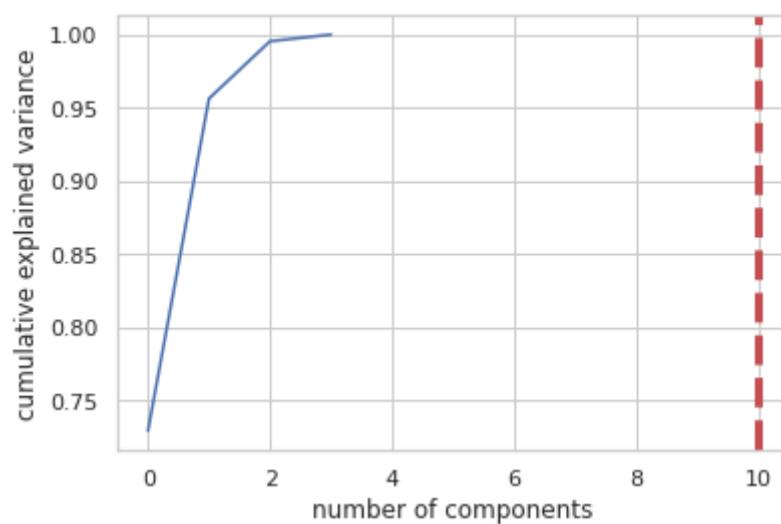
print("Performance Evaluation")
print(classification_report(y_test, y_pred))

print("-----")
print("-----")

print("Accuracy:")
print(accuracy_score(y_test, y_pred))

import matplotlib.pyplot as plt
from sklearn.metrics import plot_confusion_matrix
plot_confusion_matrix(rf_random, X_test, y_test)
plt.show()

```



None

### Parameters currently in use:

```
[Parallel(n_jobs=1)]: Using backend SequentialBackend with 1 concurrent workers.
[Parallel(n_jobs=1)]: Done    1 out of    1 | elapsed:    0.0s remaining:    0.0s
[cv] ..... C=1, gamma=1, kernel=rbf, total= 0.0s
[cv] C=1, gamma=1, kernel=rbf .....
[cv] ..... C=1, gamma=1, kernel=rbf, total= 0.0s
[cv] C=1, gamma=1, kernel=rbf .....
[cv] ..... C=1, gamma=1, kernel=rbf, total= 0.0s
[cv] C=1, gamma=1, kernel=poly .....
[cv] ..... C=1, gamma=1, kernel=poly, total= 0.0s
[cv] C=1, gamma=1, kernel=poly .....
[cv] ..... C=1, gamma=1, kernel=poly, total= 0.0s
[cv] C=1, gamma=1, kernel=poly .....
[cv] ..... C=1, gamma=1, kernel=poly, total= 0.0s
[cv] C=1, gamma=1, kernel=poly .....
[cv] ..... C=1, gamma=1, kernel=poly, total= 0.0s
[cv] C=1, gamma=1, kernel=poly .....
[cv] ..... C=1, gamma=1, kernel=poly, total= 0.0s
[cv] C=1, gamma=1, kernel=sigmoid .....
[cv] ..... C=1, gamma=1, kernel=sigmoid, total= 0.0s
[cv] C=1, gamma=1, kernel=sigmoid .....
[cv] ..... C=1, gamma=1, kernel=sigmoid, total= 0.0s
[cv] C=1, gamma=1, kernel=sigmoid .....
[cv] ..... C=1, gamma=1, kernel=sigmoid, total= 0.0s
[cv] C=1, gamma=1, kernel=sigmoid .....
[cv] ..... C=1, gamma=1, kernel=sigmoid, total= 0.0s
[cv] C=1, gamma=1, kernel=sigmoid .....
[cv] ..... C=1, gamma=1, kernel=sigmoid, total= 0.0s
[cv] C=1, gamma=1, kernel=sigmoid .....
[cv] ..... C=1, gamma=1, kernel=sigmoid, total= 0.0s
[cv] C=1, gamma=0.1, kernel=rbf .....
[cv] ..... C=1, gamma=0.1, kernel=rbf, total= 0.0s
[cv] C=1, gamma=0.1, kernel=rbf .....
[cv] ..... C=1, gamma=0.1, kernel=rbf, total= 0.0s
[cv] C=1, gamma=0.1, kernel=rbf .....
[cv] ..... C=1, gamma=0.1, kernel=rbf, total= 0.0s
[cv] C=1, gamma=0.1, kernel=rbf .....
[cv] ..... C=1, gamma=0.1, kernel=rbf, total= 0.0s
```







```
[CV] C=100, gamma=0.001, kernel=rbf .....  
[CV] ..... C=100, gamma=0.001, kernel=rbf, total= 0.0s  
[CV] C=100, gamma=0.001, kernel=poly .....  
[CV] ..... C=100, gamma=0.001, kernel=poly, total= 0.0s  
[CV] C=100, gamma=0.001, kernel=poly .....  
[CV] ..... C=100, gamma=0.001, kernel=poly, total= 0.0s  
[CV] C=100, gamma=0.001, kernel=poly .....  
[CV] ..... C=100, gamma=0.001, kernel=poly, total= 0.0s  
[CV] C=100, gamma=0.001, kernel=poly .....  
[CV] ..... C=100, gamma=0.001, kernel=poly, total= 0.0s  
[CV] C=100, gamma=0.001, kernel=sigmoid .....  
[CV] ..... C=100, gamma=0.001, kernel=sigmoid, total= 0.0s  
[CV] C=100, gamma=0.001, kernel=sigmoid .....  
[CV] ..... C=100, gamma=0.001, kernel=sigmoid, total= 0.0s  
[CV] C=100, gamma=0.001, kernel=sigmoid .....  
[CV] ..... C=100, gamma=0.001, kernel=sigmoid, total= 0.0s  
[CV] C=100, gamma=0.001, kernel=sigmoid .....  
[CV] ..... C=100, gamma=0.001, kernel=sigmoid, total= 0.0s  
[CV] C=100, gamma=0.001, kernel=sigmoid .....  
[CV] ..... C=100, gamma=0.001, kernel=sigmoid, total= 0.0s
```

Confusion Matrix:

```
[[14  0  0]  
 [ 0 14  3]  
 [ 0  0 14]]
```

-----

Performance Evaluation

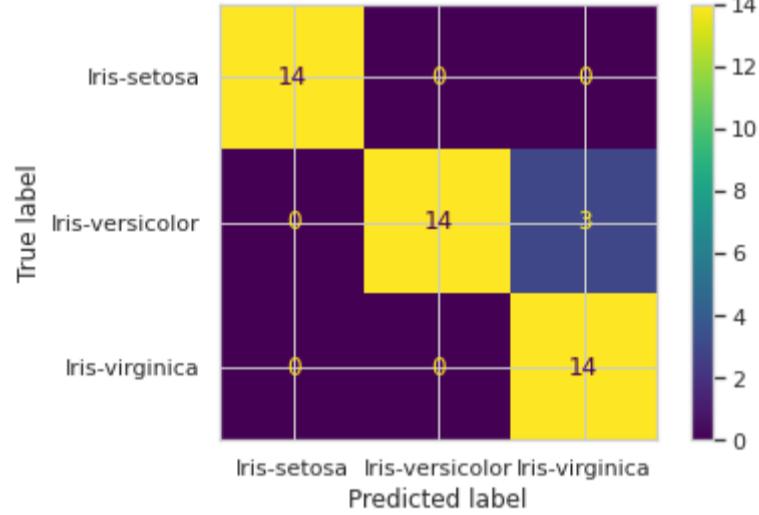
	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	14
Iris-versicolor	1.00	0.82	0.90	17
Iris-virginica	0.82	1.00	0.90	14
accuracy			0.93	45
macro avg	0.94	0.94	0.94	45
weighted avg	0.95	0.93	0.93	45

-----

Accuracy:

0.9333333333333333

[Parallel(n\_jobs=1)]: Done 240 out of 240 | elapsed: 0.7s finished



In [ ]: