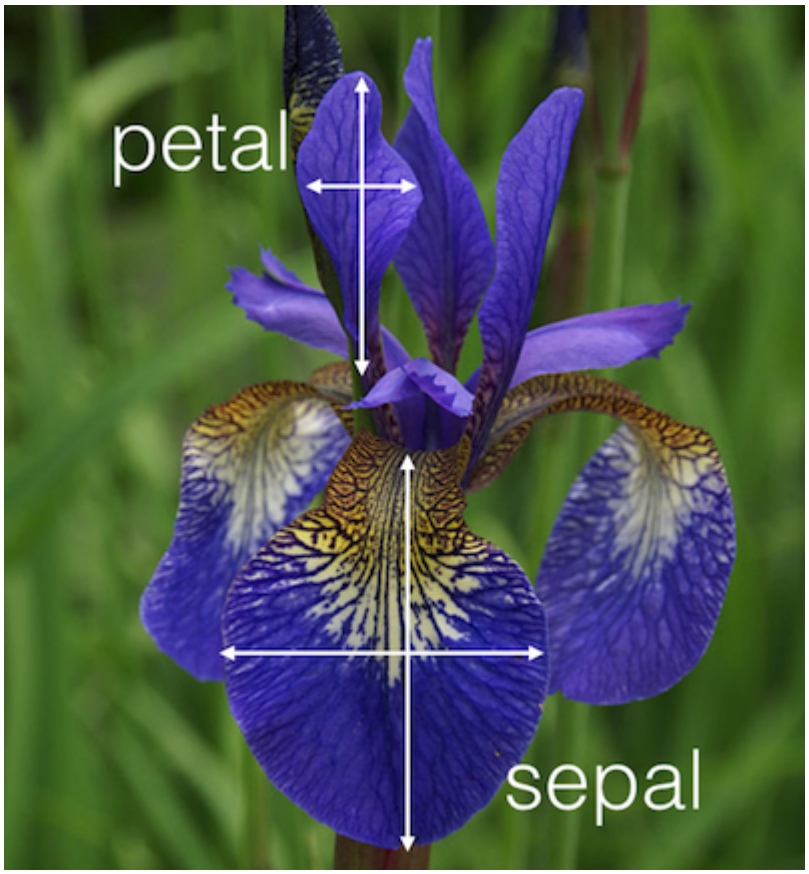
Exercise



Use famous iris flower dataset from sklearn.datasets to predict flower species using random forest classifier.

1. Measure prediction score using default n_estimators (100)
2. Now tune your model by changing number of trees in your classifer and tell me what best score you can get using how many trees

In [1]:
```
from sklearn.datasets import load_iris
iris = load_iris()
dir(iris)
```

Out[1]:
```
['DESCR',
 'data',
 'feature_names',
 'filename',
 'frame',
 'target',
 'target_names']
```

In [2]:
```
import pandas as pd
df = pd.DataFrame(iris.data, columns=iris.feature_names)
df.head()
```

Out[2]:

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

In [3]:
```
df['target'] = iris.target
df.head()
```

Out[3]:

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

train test split

In [4]:
```
X = df.drop(['target'],axis='columns')
y = iris.target
```

In [5]:
```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.2)
```

In [6]:
```
from sklearn.ensemble import RandomForestClassifier
model = RandomForestClassifier()
model.fit(X_train, y_train)
```

Out[6]:
```
RandomForestClassifier()
```

In [7]:
```
model.score(X_test,y_test)
```

Out[7]:
```
0.9
```

In [21]:
```
model = RandomForestClassifier(n_estimators=40)
model.fit(X_train, y_train)
model.score(X_test,y_test)
```

Out[21]:
```
0.9
```

Date	Author
2021-09-19	Ehsan Zia