sk_classification&regressionTree

April 15, 2019

1 Classification and Regression Trees

Classification and Regression Trees (CART) are constructed from a dataset by making splits that best separate the data for the classes or predictions being made. The CART algorithm can be used for classification or regression.

```
In [1]: from sklearn import datasets
        from sklearn import metrics
        from sklearn.tree import DecisionTreeClassifier
```

2 Iris flowers Dataset

```
In [2]: dataset = datasets.load_iris()
```

3 Model

```
In [3]: model = DecisionTreeClassifier()
        model.fit(dataset.data, dataset.target)
```

```
Out[3]: DecisionTreeClassifier(class_weight=None, criterion='gini', max_depth=None,
                    max_features=None, max_leaf_nodes=None,
                    min_impurity_decrease=0.0, min_impurity_split=None,
                    min_samples_leaf=1, min_samples_split=2,
                    min_weight_fraction_leaf=0.0, presort=False, random_state=None,
                    splitter='best')
```

4 Prediction/Classification

```
In [4]: expected = dataset.target
        predicted = model.predict(dataset.data)
        print(metrics.classification_report(expected, predicted))
        print(metrics.confusion_matrix(expected, predicted))
```

```
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        50
```

```
           1       1.00      1.00      1.00        50
           2       1.00      1.00      1.00        50

   micro avg       1.00      1.00      1.00       150
   macro avg       1.00      1.00      1.00       150
weighted avg       1.00      1.00      1.00       150

[[50  0  0]
 [ 0 50  0]
 [ 0  0 50]]
```

4.1 References

1. https://machinelearningmastery.com/get-your-hands-dirty-with-scikit-learn-now/