supervised_learning

February 19, 2022

1 Supervised Learning

1.1 Functional Approach

Sven Gehrke Feb 2022

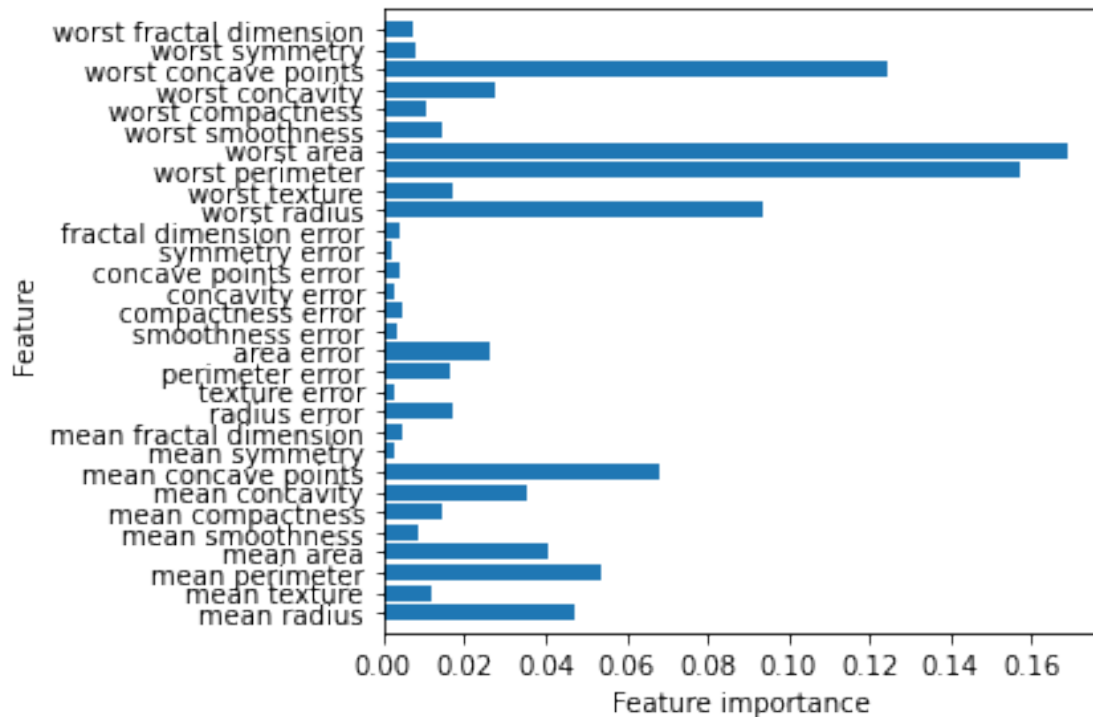## 1.2 Main Part

1.2.1 Load the data

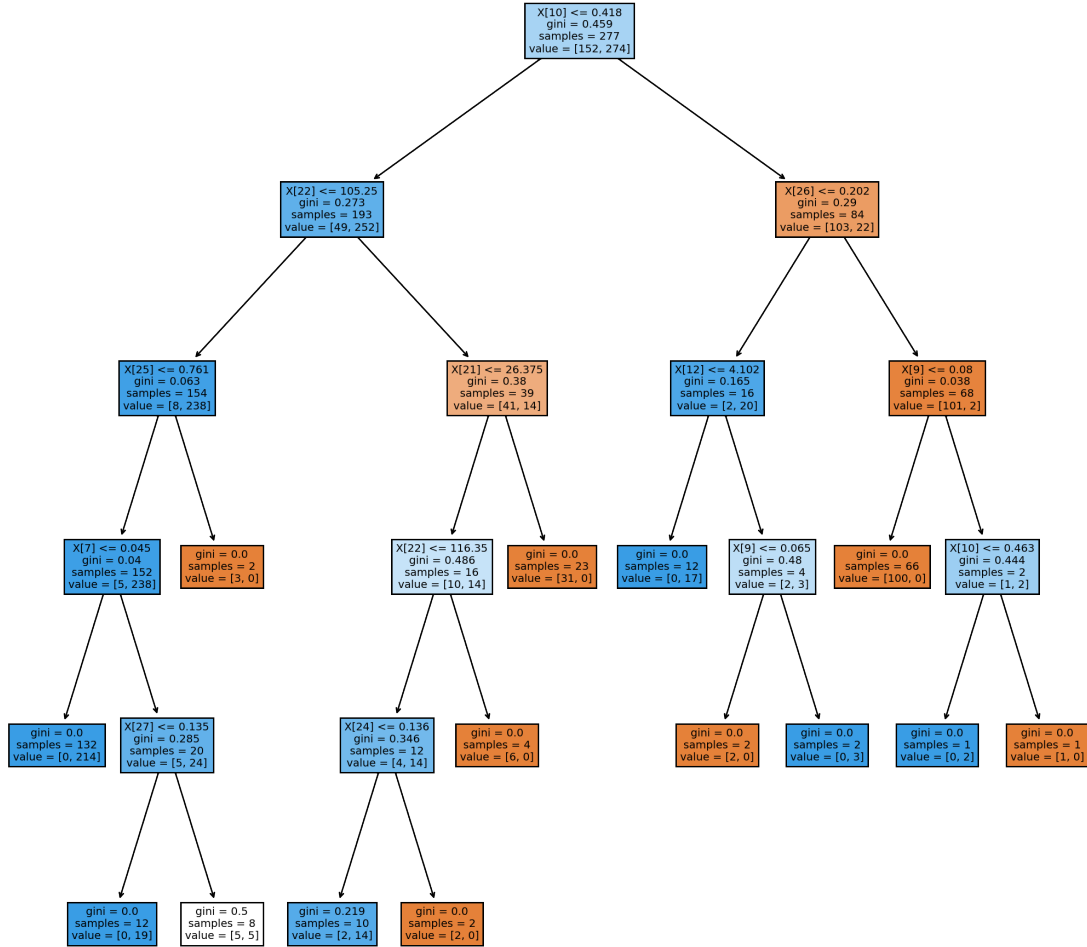### 1.2.2 Data Understanding

1.2.3 Data Preparation

1.2.4 Modelling

1.2.5 Evaluation

plotting feature importance

DecisionTreeClassifier(max_depth=5, max_features='auto', random_state=586074676)

X[10] <= 0.418
gini = 0.459
samples = 277
value = [152, 274]

X[22] <= 105.25
gini = 0.273
samples = 193
value = [49, 252]

X[26] <= 0.202
gini = 0.29
samples = 84
value = [103, 22]

X[25] <= 0.761
gini = 0.063
samples = 154
value = [8, 238]

X[21] <= 26.375
gini = 0.38
samples = 39
value = [41, 14]

X[12] <= 4.102
gini = 0.165
samples = 16
value = [2, 20]

X[9] <= 0.08
gini = 0.038
samples = 68
value = [101, 2]

X[7] <= 0.045
gini = 0.04
samples = 152
value = [5, 238]

gini = 0.0
samples = 2
value = [3, 0]

X[22] <= 116.35
gini = 0.486
samples = 16
value = [10, 14]

gini = 0.0
samples = 23
value = [31, 0]

gini = 0.0
samples = 12
value = [0, 17]

X[9] <= 0.065
gini = 0.48
samples = 4
value = [2, 3]

gini = 0.0
samples = 66
value = [100, 0]

X[10] <= 0.463
gini = 0.444
samples = 2
value = [1, 2]

gini = 0.0
samples = 132
value = [0, 214]

X[27] <= 0.135
gini = 0.285
samples = 20
value = [5, 24]

X[24] <= 0.136
gini = 0.346
samples = 12
value = [4, 14]

gini = 0.0
samples = 4
value = [6, 0]

gini = 0.0
samples = 2
value = [2, 0]

gini = 0.0
samples = 2
value = [0, 3]

gini = 0.0
samples = 1
value = [0, 2]

gini = 0.0
samples = 1
value = [1, 0]

gini = 0.0
samples = 12
value = [0, 19]

gini = 0.5
samples = 8
value = [5, 5]

gini = 0.219
samples = 10
value = [2, 14]

gini = 0.0
samples = 2
value = [2, 0]

```
|--- radius error <= 0.42
|   |--- worst perimeter <= 105.25
|   |   |--- worst compactness <= 0.76
|   |   |   |--- mean concave points <= 0.05
|   |   |   |   |--- class: 1.0
|   |   |   |--- mean concave points >  0.05
|   |   |   |   |--- worst concave points <= 0.14
|   |   |   |   |   |--- class: 1.0
|   |   |   |   |--- worst concave points >  0.14
|   |   |   |   |   |--- class: 0.0
|   |   |--- worst compactness >  0.76
```
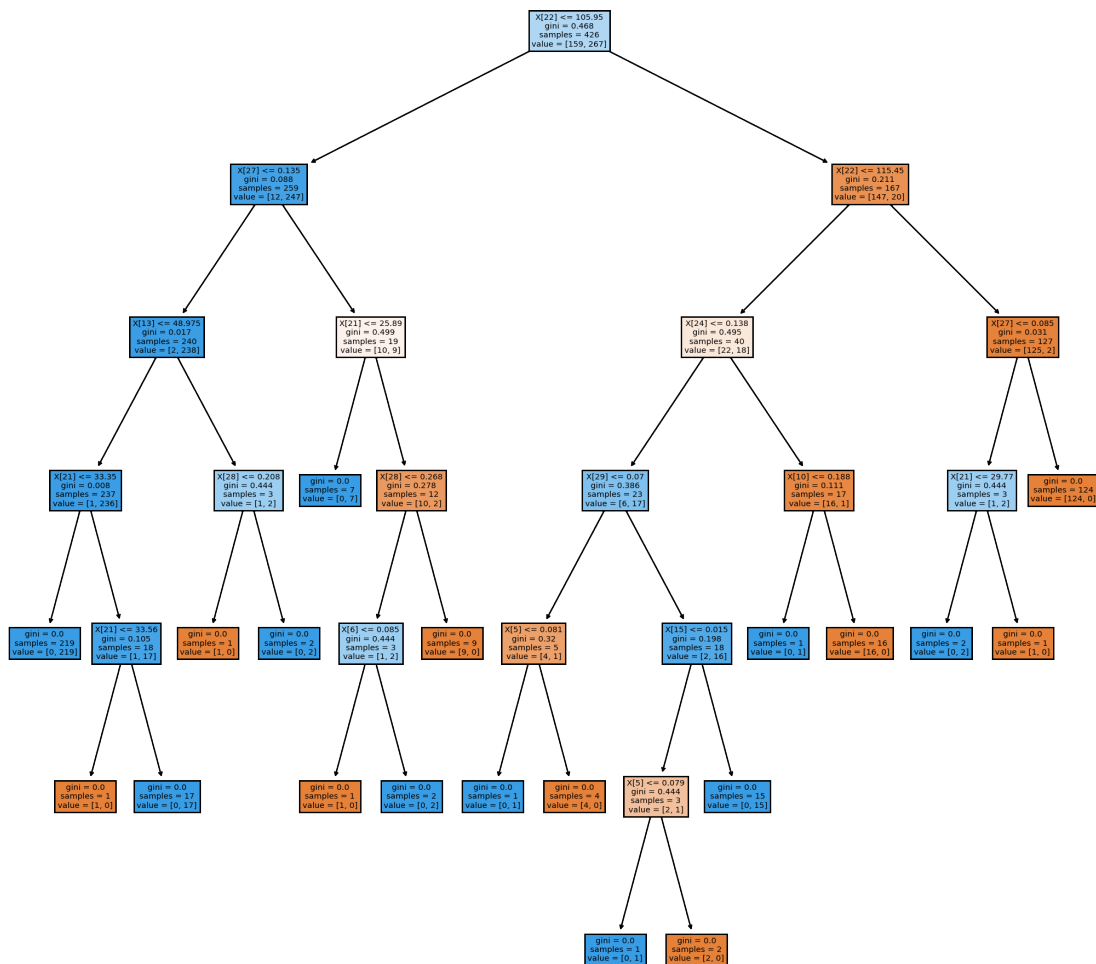
```
|   |   |   |--- class: 0.0
|   |--- worst perimeter >  105.25
|   |   |--- worst texture <= 26.38
|   |   |   |--- worst perimeter <= 116.35
|   |   |   |   |--- worst smoothness <= 0.14
|   |   |   |   |   |--- class: 1.0
|   |   |   |   |--- worst smoothness >  0.14
|   |   |   |   |   |--- class: 0.0
|   |   |   |--- worst perimeter >  116.35
|   |   |   |   |--- class: 0.0
|   |   |--- worst texture >  26.38
|   |   |   |--- class: 0.0
|--- radius error >  0.42
|   |--- worst concavity <= 0.20
|   |   |--- perimeter error <= 4.10
|   |   |   |--- class: 1.0
|   |   |--- perimeter error >  4.10
|   |   |   |--- mean fractal dimension <= 0.06
|   |   |   |   |--- class: 0.0
|   |   |   |--- mean fractal dimension >  0.06
|   |   |   |   |--- class: 1.0
|   |--- worst concavity >  0.20
|   |   |--- mean fractal dimension <= 0.08
|   |   |   |--- class: 0.0
|   |   |--- mean fractal dimension >  0.08
|   |   |   |--- radius error <= 0.46
|   |   |   |   |--- class: 1.0
|   |   |   |--- radius error >  0.46
|   |   |   |   |--- class: 0.0
```

DecisionTreeClassifier()


```
|--- worst perimeter <= 105.95
|   |--- worst concave points <= 0.14
|   |   |--- area error <= 48.98
|   |   |   |--- worst texture <= 33.35
|   |   |   |   |--- class: 1
|   |   |   |--- worst texture >  33.35
|   |   |   |   |--- worst texture <= 33.56
|   |   |   |   |   |--- class: 0
|   |   |   |   |--- worst texture >  33.56
|   |   |   |   |   |--- class: 1
|   |   |--- area error >  48.98
|   |   |   |--- worst symmetry <= 0.21
|   |   |   |   |--- class: 0
```

```
|   |   |   |--- worst symmetry >  0.21
|   |   |   |   |--- class: 1
|   |--- worst concave points >  0.14
|   |   |--- worst texture <= 25.89
|   |   |   |--- class: 1
|   |   |--- worst texture >  25.89
|   |   |   |--- worst symmetry <= 0.27
|   |   |   |   |--- mean concavity <= 0.08
|   |   |   |   |   |--- class: 0
|   |   |   |   |--- mean concavity >  0.08
|   |   |   |   |   |--- class: 1
|   |   |   |--- worst symmetry >  0.27
|   |   |   |   |--- class: 0
|--- worst perimeter >  105.95
|   |--- worst perimeter <= 115.45
|   |   |--- worst smoothness <= 0.14
|   |   |   |--- worst fractal dimension <= 0.07
|   |   |   |   |--- mean compactness <= 0.08
|   |   |   |   |   |--- class: 1
|   |   |   |   |--- mean compactness >  0.08
|   |   |   |   |   |--- class: 0
|   |   |   |--- worst fractal dimension >  0.07
|   |   |   |   |--- compactness error <= 0.02
|   |   |   |   |   |--- mean compactness <= 0.08
|   |   |   |   |   |   |--- class: 1
|   |   |   |   |   |--- mean compactness >  0.08
|   |   |   |   |   |   |--- class: 0
|   |   |   |   |--- compactness error >  0.02
|   |   |   |   |   |--- class: 1
|   |   |--- worst smoothness >  0.14
|   |   |   |--- radius error <= 0.19
|   |   |   |   |--- class: 1
|   |   |   |--- radius error >  0.19
|   |   |   |   |--- class: 0
|   |--- worst perimeter >  115.45
|   |   |--- worst concave points <= 0.09
|   |   |   |--- worst texture <= 29.77
|   |   |   |   |--- class: 1
|   |   |   |--- worst texture >  29.77
|   |   |   |   |--- class: 0
|   |   |--- worst concave points >  0.09
|   |   |   |--- class: 0


Mean accuracy for evaluation/ model set: 0.958/ 0.9953 (-0.03726)
```