

Decision Trees

Decision tree is a representation for classification, it is used for inductive inference. This is a method for approximating discrete-valued target functions. The decision tree is modeled as a tree-shaped diagram.

The team made an implementation of ID3 in Python, the output of this program is the decision tree of the dataset given as input. Making this implementation, we could identify some advantages and disadvantages using our program and a pre-created suite such as WEKA.

Advantages:

- The output can be modified to the needs of the problem given.
- It is easier to know how the algorithm works.
- The code can be optimized.
- The implementation helps us to improve our programming skills.
- You don't need to learn how the program works.

Disadvantages:

- Our program does not have different algorithms like Weka.
- It could have a low level of accuracy.
- For building the program, we need to spend a certain of time to code it.

For this report, we used two different datasets to compare our implementation and WEKA.

We decided to use the datasets given by the lab to test that our program works well. In the following section there is the result of our program and the output of Weka. In these examples, the outputs are the same for both programs.

1. Test 2

The input given was:

```

@relation contact-lenses

@attribute age          {young, pre-presbyopic, presbyopic}
@attribute spectacle-prescrip {myope, hypermetrope}
@attribute astigmatism    {no, yes}
@attribute tear-prod-rate  {reduced, normal}
@attribute contact-lenses {soft, hard, none}

@data
young,myope,no,reduced,none
young,myope,no,normal,soft
young,myope,yes,reduced,none
young,myope,yes,normal,hard
young,hypermetrope,no,reduced,none
young,hypermetrope,no,normal,soft
young,hypermetrope,yes,reduced,none
young,hypermetrope,yes,normal,hard
pre-presbyopic,myope,no,reduced,none
pre-presbyopic,myope,no,normal,soft
pre-presbyopic,myope,yes,reduced,none
pre-presbyopic,myope,yes,normal,hard
pre-presbyopic,hypermetrope,no,reduced,none
pre-presbyopic,hypermetrope,no,normal,soft
pre-presbyopic,hypermetrope,yes,reduced,none
pre-presbyopic,hypermetrope,yes,normal,none
presbyopic,myope,no,reduced,none
presbyopic,myope,no,normal,none
presbyopic,myope,yes,reduced,none
presbyopic,myope,yes,normal,hard
presbyopic,hypermetrope,no,reduced,none
presbyopic,hypermetrope,no,normal,soft
presbyopic,hypermetrope,yes,reduced,none
presbyopic,hypermetrope,yes,normal,none

```

a. Weka

```

J48 pruned tree
-----

tear-prod-rate = reduced: none (12.0)
tear-prod-rate = normal
|   astigmatism = no: soft (6.0/1.0)
|   astigmatism = yes
|   |   spectacle-prescrip = myope: hard (3.0)
|   |   spectacle-prescrip = hypermetrope: none (3.0/1.0)

Number of Leaves   :    4

Size of the tree   :    7

```

b. Our program

```

tear-prod-rate: reduced
  ANSWER: none
tear-prod-rate: normal
  astigmatism: no
    age: young
      ANSWER: soft
    age: pre-presbyopic
      ANSWER: soft
    age: presbyopic
      spectacle-prescrip: myope
        ANSWER: none
      spectacle-prescrip: hypermetrope
        ANSWER: soft
  astigmatism: yes
    spectacle-prescrip: myope
      ANSWER: hard
    spectacle-prescrip: hypermetrope
      age: young
        ANSWER: hard
      age: pre-presbyopic
        ANSWER: none
      age: presbyopic
        ANSWER: none

```

2. Test 3

c. Weka

J48 pruned tree

```

-----

outlook = sunny
|  humidity = high: no (3.0)
|  humidity = normal: yes (2.0)
outlook = overcast: yes (4.0)
outlook = rainy
|  windy = TRUE: no (2.0)
|  windy = FALSE: yes (3.0)

```

Number of Leaves : 5

d. Our program

```

outlook: sunny
  humidity: high
    ANSWER: no
  humidity: normal
    ANSWER: yes
outlook: overcast
  ANSWER: yes
outlook: rainy
  windy: TRUE
    ANSWER: no
  windy: FALSE
    ANSWER: yes

```

As we know, a decision tree is a decision tree model that shows us the decisions and their possible consequences. They are an important type of algorithm that are used in machine learning, it is a technique for supervised classification learning.

One of the uses of them is to create a predictive model. Predictive modeling maps observations about an item's target value; it is use in data mining, statistics and machine learning for several purposes.

One example of an application is having a dataset about engine's temperatures, oil, information about if the engine is damaged or not; a decision tree can helps us to determine the probability that the engine needs maintenance or the probability to fail in a flight, etc. These information can help to a company to prevent incidents or to perform maintenance after an engine's failed.

Decision trees are very helpful in different areas such as economy, engineering, etc. because it would give you important information like probabilities, costs, and alternatives.

REFERENCES:

<http://www.cs.princeton.edu/courses/archive/spr07/cos424/papers/mitchell-dectrees.pdf>