

CS 760: Machine Learning - Fall 2020

Homework 4: Decision Trees

Due : 10/29/2020

Zijie Zhang

October 28, 2020

All the code involved is in this [GitHub repository](https://github.com/z-zijie/2020Fall/tree/master/COMP760/Homework4).

<https://github.com/z-zijie/2020Fall/tree/master/COMP760/Homework4>

All the code is in this [python file](https://github.com/z-zijie/2020Fall/blob/master/COMP760/Homework4/code.py).

<https://github.com/z-zijie/2020Fall/blob/master/COMP760/Homework4/code.py>

Problem 1

- **Pclass:**
 $1 \rightarrow [1, 0, 0];$
 $2 \rightarrow [0, 1, 0];$
 $3 \rightarrow [0, 0, 1].$
- **Sex:**
 $0 \rightarrow [1, 0];$
 $1 \rightarrow [0, 1].$
- **Age:** Binning in $[0, 20.25, 40.5, 60.75, 81]$ then convert to **One-Hot**.
- **Siblings/Spouses Aboard:** Convert to One-Hot like **Sex**.
- **Parents/Children Aboard:** Convert to One-Hot like **Sex**.
- **Fare:** Binning in $[0, 10, 20, 30, 40, 50, 70, 115, 160, 205, 250, 295, 340, 385, 430, 475, 520]$ then convert to **One-Hot**.

Explanation:

1. Pclass, Sex, Siblings/Spouses Aboard and Parents/Children Aboard are **CATEGORICAL FEATURES**, so we can directly encode them One-Hot.
2. Age and Fare are **NUMERICAL FEATURES**, we should binning them first. It is worth noticing that the distribution of *Fare* is **skewed distribution**, we can't binning it uniformly.

Problem 2

Please check the relevant [code](#) in the file.

Problem 3

Please check the relevant [code](#) in the file.

Stopping criteria are

1. If the entropy of the response y in the current subset of data is close to zero.
2. If the current subset of data contains too few samples. (less than 5% of the total data)
3. If the depth of the Tree greater than the number of features.

Problem 4

Please check the relevant [code](#) and [output](#) in the file.

```
|—Sex—|
  |—Fare—|
    |—Pclass—|
      | False | *
      | False | *
      |—Age—|
        |—Parents/Children Aboard—|
          | False | *
          | False | *
          | False | *
          | False | *
          | False | *
          | False | *
          | False | *
        |—Siblings/Spouses Aboard—|
          | False | *
          | False | *
          | False | *
          | False | *
          | False | *
          | False | *
          | False | *
          | False | *
        | False | *
        | False | *
      |—Parents/Children Aboard—|
        |—Age—|
          | False | *
          |—Siblings/Spouses Aboard—|
            | False | *
            | False | *
            | False | *
            | False | *
            | False | *
            | False | *
            | False | *
            | False | *
            | False | *
          | False | *
          | False | *
        | True | *
        | False | *
        | False | *
        | False | *
        | False | *
        | False | *
      |—Siblings/Spouses Aboard—|
        | False | *
```

```

| False|*
| False|*
| False|*
| False|*
| False|*
| False|*
| False|*
| False|*
| True|*
| False|*
| False|*
| False|*
| False|*
| False|*
| False|*
| True|*
| -Pclass-|
| -Fare-|
| False|*
| False|*
| True|*
| True|*
| True|*
| True|*
| True|*
| True|*
| True|*
| True|*
| True|*
| False|*
| False|*
| False|*
| False|*
| True|*
| -Age-|
| True|*
| -Fare-|
| False|*
| False|*
| True|*
| True|*
| True|*
| True|*
| True|*
| False|*
| False|*
| False|*
| False|*
| False|*
| False|*
| False|*
| False|*
| False|*
| False|*
| True|*
| False|*
| -Fare-|
| -Age-|
| True|*
| True|*
| False|*
| True|*
| True|*
| False|*
| False|*
| False|*
| False|*

```

[illegible]

Problem 5

Please check the relevant [code](#) in the file.

Accuracy = 0.812852311612176

Problem 6

Please check the relevant **code** in the file.

My own feature vector $\mathbf{x} = [1, 1, 22, 1, 0, 71.2833]$

Predicted result = **YES**.

Problem 7

Please check the relevant **code** in the file.

[illegible]

| | | | | |
|----------|---------------------------|---------------------------|-----------|-----------|
| | | | | False * |
| | | | | False * |
| | | | | False * |
| | | | | False * |
| | -Pclass- | | | |
| | | | False * | |
| | | | False * | |
| | | -Siblings/Spouses Aboard- | | |
| | | | | False * |
| | | | | True * |
| | | | | False * |
| | | | | False * |
| | | | | False * |
| | | | | False * |
| | | | | False * |
| | | | | False * |
| | | | | False * |
| | | | | False * |
| | False * | | | |
| | False * | | | |
| | -Parents/Children Aboard- | | | |
| | | -Age- | | |
| | | | False * | |
| | | | False * | |
| | | | False * | |
| | | | False * | |
| | | | True * | |
| | | | False * | |
| | | | False * | |
| | | | False * | |
| | | | False * | |
| | | | False * | |
| | -Pclass- | | | |
| | | True * | | |
| | | False * | | |
| | | False * | | |
| | False * | | | |
| | False * | | | |
| | False * | | | |
| | True * | | | |
| | False * | | | |
| | False * | | | |
| | False * | | | |
| | False * | | | |
| | False * | | | |
| | False * | | | |
| | True * | | | |
| -Pclass- | | | | |
| | -Fare- | | | |
| | | False * | | |
| | | False * | | |
| | | True * | | |
| | | True * | | |
| | | False * | | |
| | | True * | | |
| | | True * | | |
| | | True * | | |
| | | True * | | |
| | | True * | | |
| | | True * | | |
| | | True * | | |
| | | True * | | |
| | | False * | | |
| | | False * | | |
| | | False * | | |
| | | False * | | |
| | | True * | | |
| | -Age- | | | |
| | | True * | | |
| | | True * | | |
| | | True * | | |
| | | False * | | |


```

| True | *
| True | *
| False | *
| True | *
| True | *
| False | *
| True | *
| False | *
| False | *
| False | *
| False | *
| False | *
| True | *
| True | *
| False | *
| False | *
| False | *
| -Age- |
| True | *
| True | *
| True | *
| False | *
| -Fare- |
| -Age- |
| True | *
| False | *
| False | *
| True | *
| True | *
| False | *
| False | *
| False | *
| False | *
| False | *
| False | *
| False | *
| False | *
| False | *
| False | *
| False | *
| False | *
| False | *
| -Sex- | 3
| -Fare- |
| -Siblings/Spouses Aboard- |
| -Pclass- |
| False | *
| False | *
| -Age- |
| False | *
| False | *
| False | *
| False | *
| False | *
| False | *
| False | *
| -Parents/Children Aboard- |
| -Pclass- |
| False | *
| -Age- |
| False | *

```


[illegible]

[illegible]