M. Azka Yasin

5115100038

Kecerdasan Komputasional F

**Abalone data**

1. Deskripsi data

Number of Instances: 4177

Number of Attributes: 8

Attribute information:

Given is the attribute name, attribute type, the measurement unit and a

brief description. The number of rings is the value to predict: either

as a continuous value or as a classification problem.

Name Data Type Meas. Description

---- --------- ----- ----- ---------------

Sex nominal M, F, and I (infant)

Length continuous mm Longest shell measurement

Diameter continuous mm perpendicular to length

Height continuous mm with meat in shell

Whole weight continuous grams whole abalone

Shucked weight continuous grams weight of meat

Viscera weight continuous grams gut weight (after bleeding)

Shell weight continuous grams after being dried

Rings integer +1.5 gives the age in years

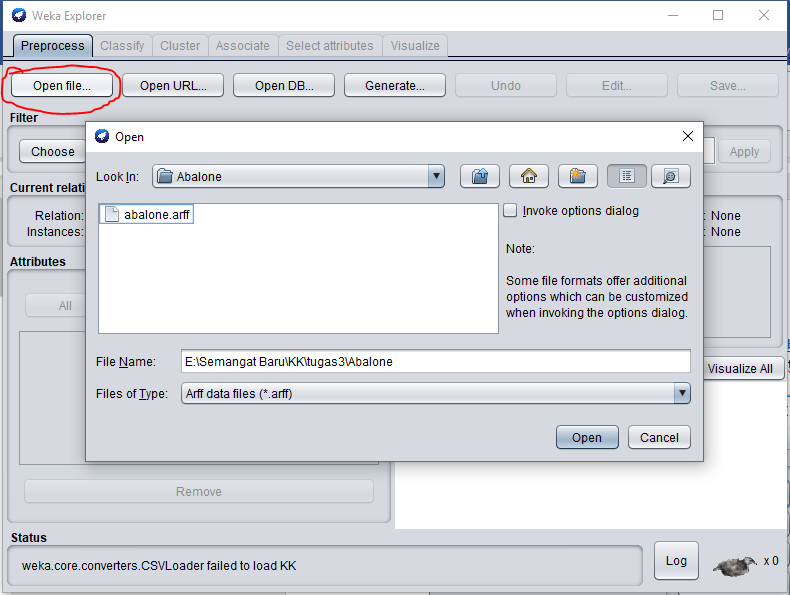
Missing Attribute Values: None

1. Langkah-langkah dan metode

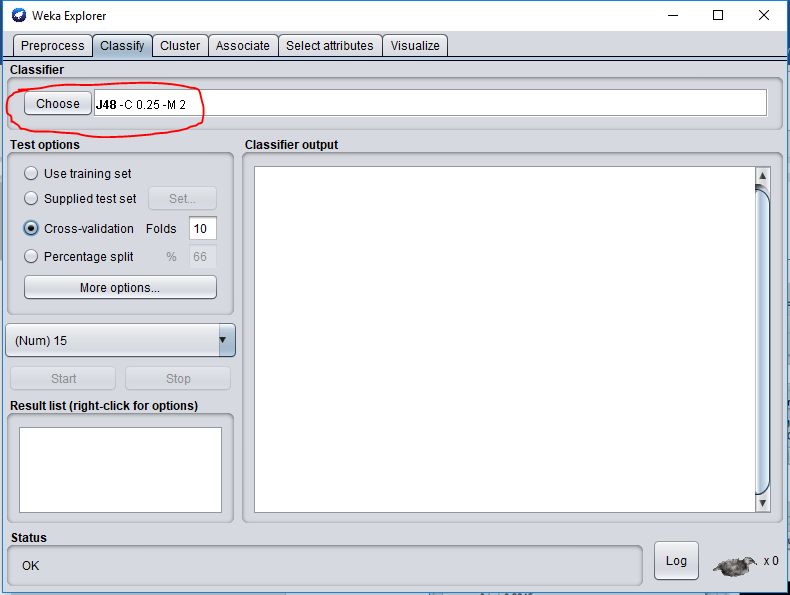
* Dapatkan abalone dataset dari UCI ML Repo di link berikut <http://archive.ics.uci.edu/ml/machine-learning-databases/abalone/>
* Ubah file abalone.data menjadi abalone.arff karena weka tidakdapat membaca file.data
* Buka aplikasi weka dan pilih menu explorer



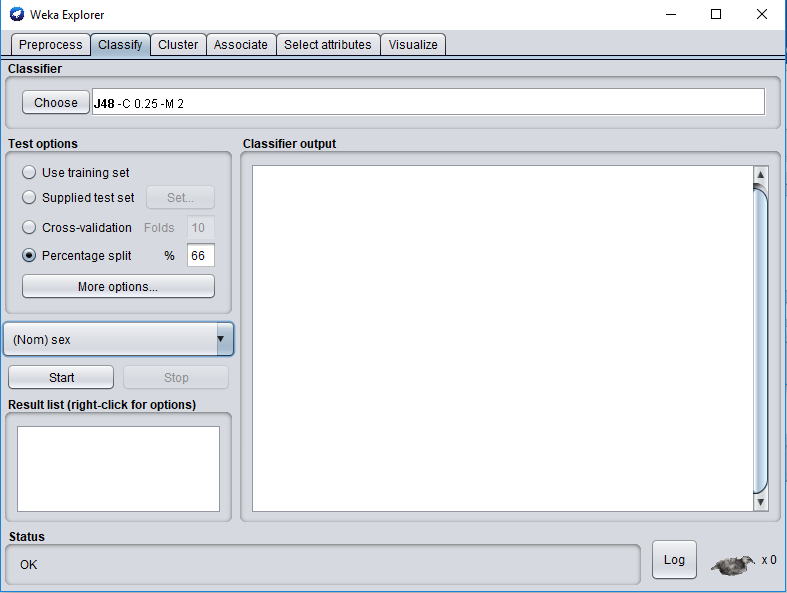
* Kemudian load file tadi dengan memilih menu open file dan kemudian cari destinasi file tersebut



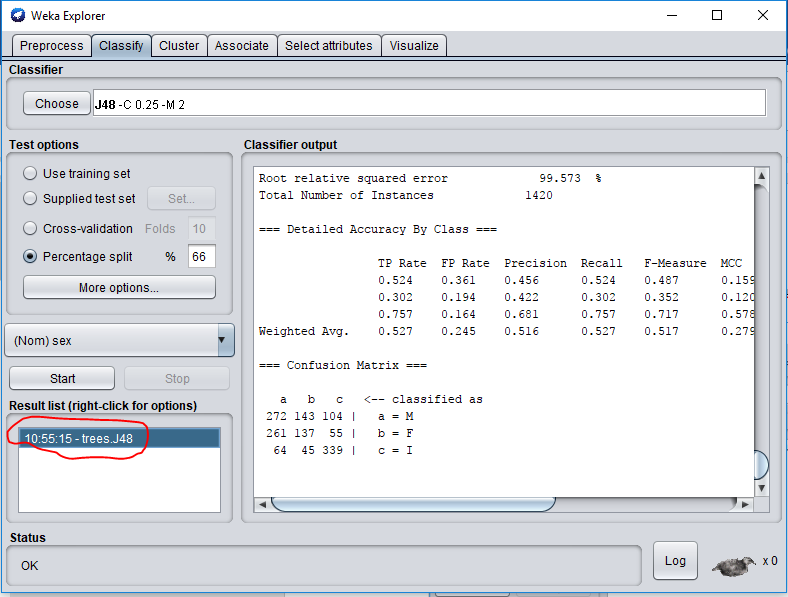
* Masuk ke tab classify,kemudian pilih classifiernya berupa j48 dengan cara choose – classifiers – trees – j48



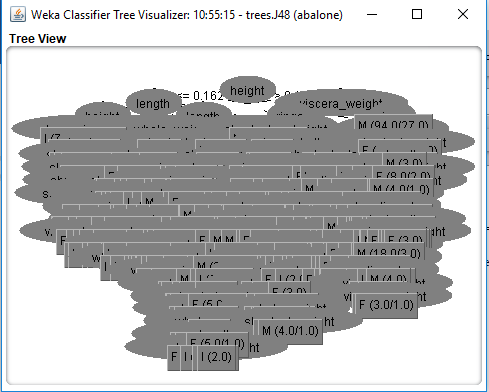
* Kemudian pada test option kita memilih presentasi pembagian data, kita pilih percentage split dengan besar 66%



* Kemudian klik start, dan untuk meilihat hasilnya klik kanan pada hasil pilih visualize tree



* Maka akan keluar treenya sebagai berikut



1. Hasil data :

Number of Leaves 349

Size of the tree 697

Correctly Classified Instances 748 52.6761 %

Incorrectly Classified Instances 672 47.3239 %

Kappa statistic 0.2859

Mean absolute error 0.3467

Root mean squared error 0.4689

Relative absolute error 78.1941 %

Root relative squared error 99.573 %

Total Number of Instances 1420

1. Untuk hasil presentasi sebesar 52,6761%, saya pikir metode j48 yang tidak sesuai untuk mengklasifikasikan abalone data sehingga hasilnya kurang maksimal

**Breast Cancer**

1. Deskripsi data

Number of Instances: 286

Number of Attributes: 9 + the class attribute

|  |
| --- |
| Attribute Information: |
|  | 1. Class: no-recurrence-events, recurrence-events | |
|  | 2. age: 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90-99. | |
|  | 3. menopause: lt40, ge40, premeno. | |
|  | 4. tumor-size: 0-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, | |
|  | 45-49, 50-54, 55-59. | |
|  | 5. inv-nodes: 0-2, 3-5, 6-8, 9-11, 12-14, 15-17, 18-20, 21-23, 24-26, | |
|  | 27-29, 30-32, 33-35, 36-39. | |
|  | 6. node-caps: yes, no. | |
|  | 7. deg-malig: 1, 2, 3. | |
|  | 8. breast: left, right. | |
|  | 9. breast-quad: left-up, left-low, right-up, right-low, central. | |
|  | 10. irradiat: yes, no. | |
| Class Distribution: | |
|  | | 1. no-recurrence-events: 201 instances |
|  | | 2. recurrence-events: 85 instances |

|  |
| --- |
| Num Instances: 286 |
|  | Num Attributes: 10 |
|  | Num Continuous: 0 (Int 0 / Real 0) |
|  | Num Discrete: 10 |
|  | Missing values: 9 / 0.3% |

1. Langkah-langkah dan metode

Untuk langkah- dan metode sama dengan abalone data namun file breast cancer kali ini tidak perlu di rubah exstensinya

1. Hasil data :

Number of Leaves 4

Size of the tree 6

Correctly Classified Instances 66 68.0412 %

Incorrectly Classified Instances 31 31.9588 %

Kappa statistic 0.2001

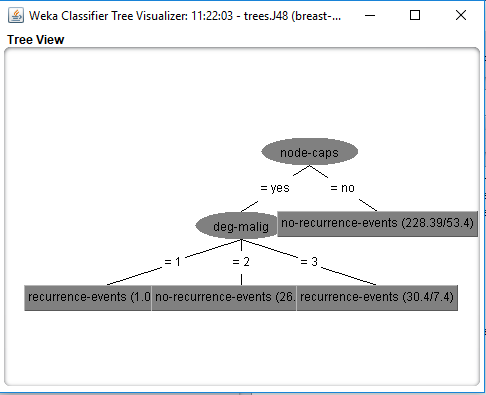
Mean absolute error 0.3966

Root mean squared error 0.4879

Relative absolute error 92.4804 %

Root relative squared error 102.0849 %

Total Number of Instances 97



1. Hasil akurasi yang didapatkan cukup memuaskan yaitu sebesar 68.0412% dikarenakan j48 ini cukup efektif untuk dataset breast cancer