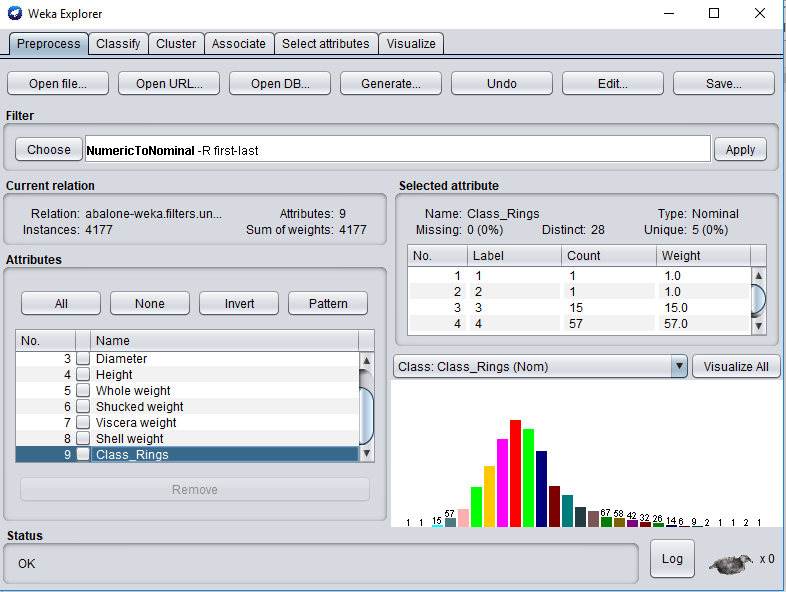
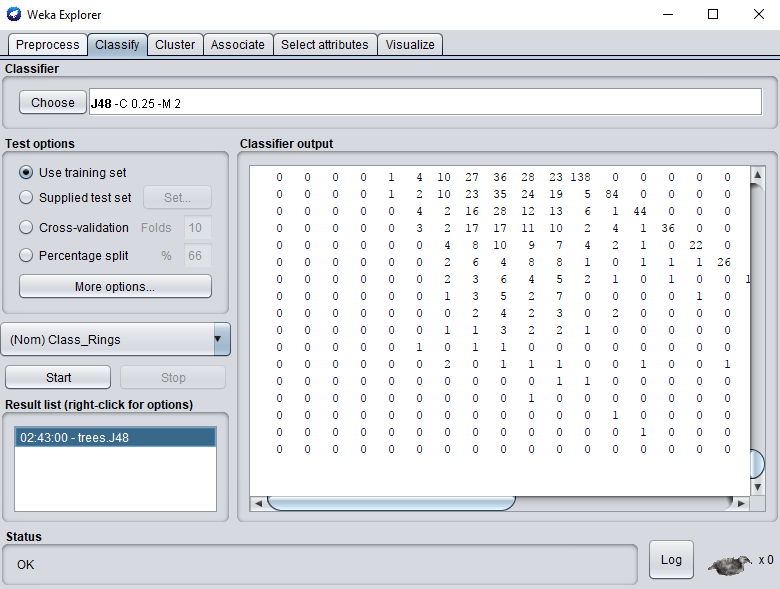
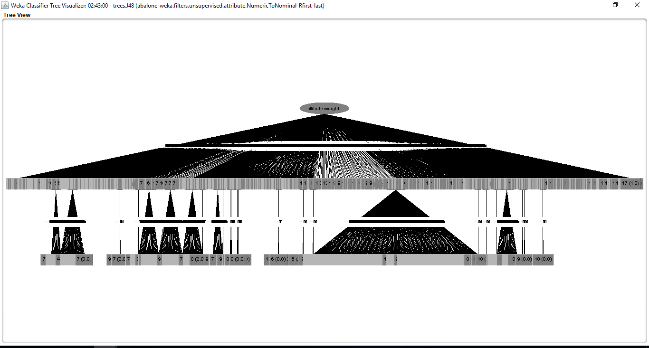
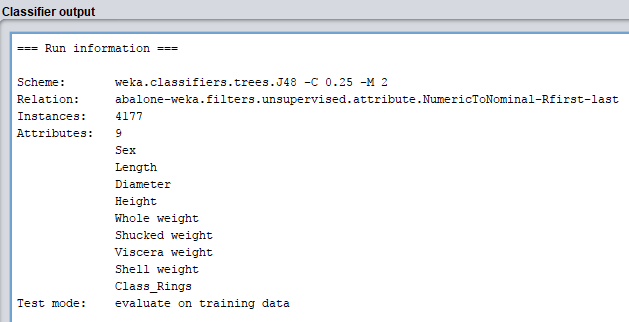
1. Get the "Abalone" dataset from UCI ML Repo

* Karena masalah koneksi internet dan server yang tidak bisa diakses, saya menggambil abalone dataset dari <https://github.com/renatopp/arff-datasets/blob/master/regression/abalone.arff>

2. Use a decision tree (C4.5/J48) learner in WEKA to predict the number of rings Abalone will have

* Untuk memasukan data tersebut saya menggunakan Weka Explorer. Selanjutnya saya mengubah nilai dari Class\_Rings yang terdapat pada dataset menjadi Nominal dengan memanfaatkan Filter.
* Untuk membuat decision tree, kita dapat memilih tab **Classify => Classifier => Choose => trees => J48.** Untuk test options saya menggunakan Use training set, setelah itu klik tombol **Start** dan lihat hasil decision tree.

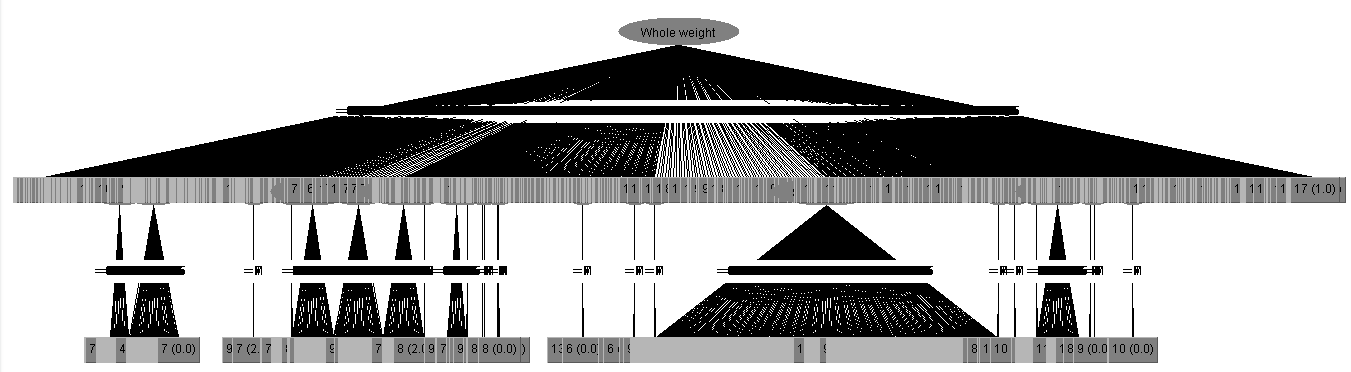
3. Write the description of your process and methods. What parameters (testing/training size, classification target, subtree raising etc.) did you use? Did you preprocess the dataset? Why did you select those parameters/preprocessing?

* Parameters yang diperoleh setelah melakukan Classifier adalah sebagai berikut.
* Dari informasi tersebut didapatkan bahwa abalone dataset memiliki jumlah data sebanyak **4177** dengan atribut setiap data adalah **9**.
* Rincian atribut tersebut adalah sebagai berikut.

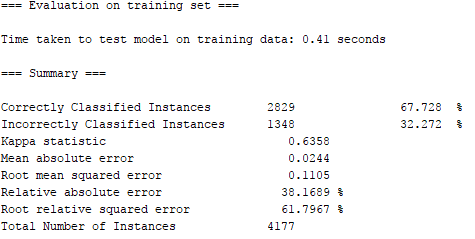
|  |  |  |
| --- | --- | --- |
| Nama | Data type | Description |
| Sex | Nominal | M, F, and I (infant) |
| Length | Continuous | Longest shell measurement |
| Diameter | Continuous | perpendicular to length |
| Height | Continuous | with meat in shell |
| Whole weight | Continuous | whole abalone |
| Shucked weight | Continuous | weight of meat |
| Viscera weight | Continuous | gut weight (after bleeding) |
| Shell weight | Continuous | after being dried |
| Rings | Integer | +1.5 gives the age in years |

* Saya menggunakan Use training set pada classifier ini, karena dengan menggunakan Use training set kita akan menggunakan semua data sebagai training set terlebih dahulu untuk menentukan tree yang akan dibangun dengan harapan memiliki tingkat akurasi yang tinggi.

4. What were your results? Show what decision trees you found.

* Decision tree yang didapat dengan menggunakan Use training set adalah sebagai berikut.

5. What do the results tell us? Why are the results (in)accurate? Why did changing parameter(s) improve/degrade accuracy?

* Hasilnya dapat ditunjukan sebagai berikut
* Dari gambar diatas didapatkan akurasi sebesar 67.728%, hal ini dipengaruhi oleh kesalahan pada Root relative squared error.