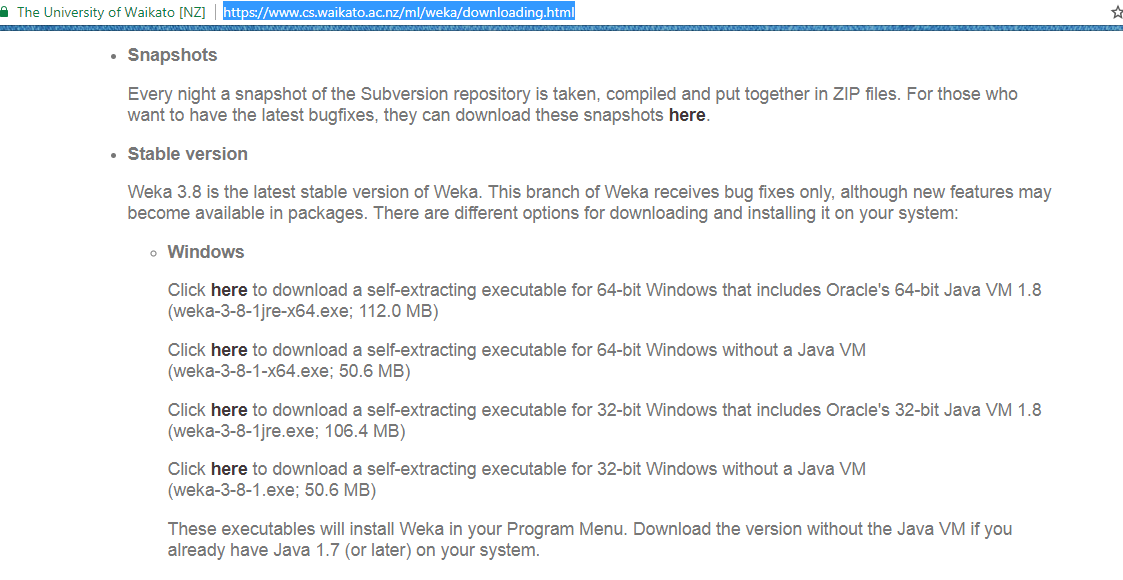
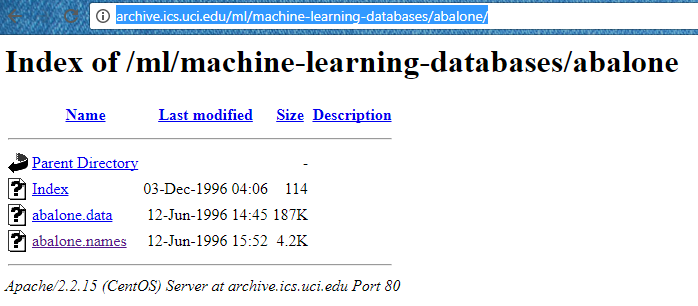
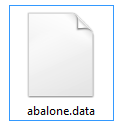
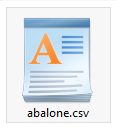
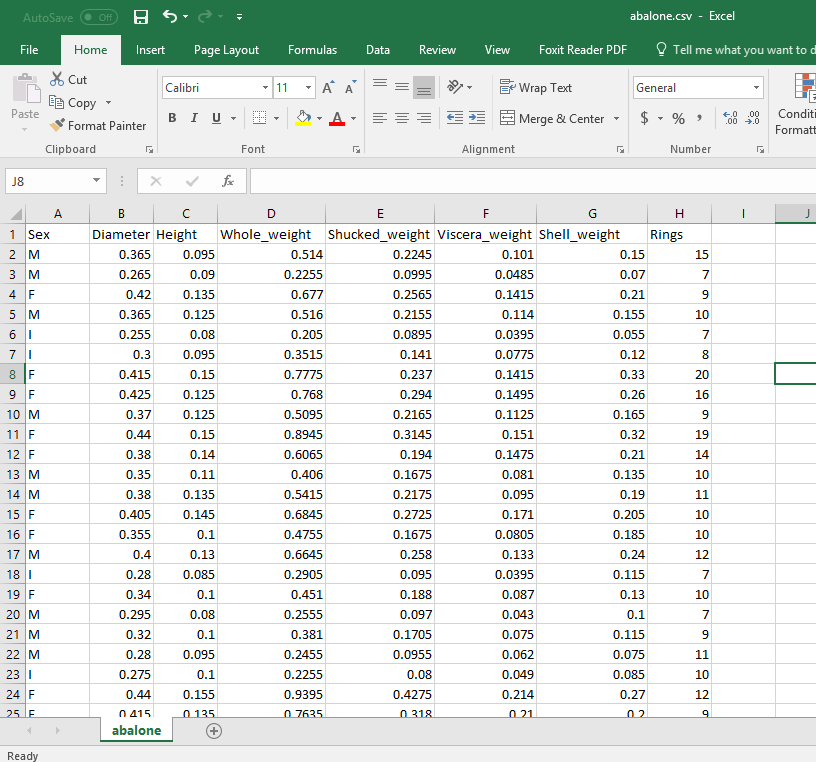
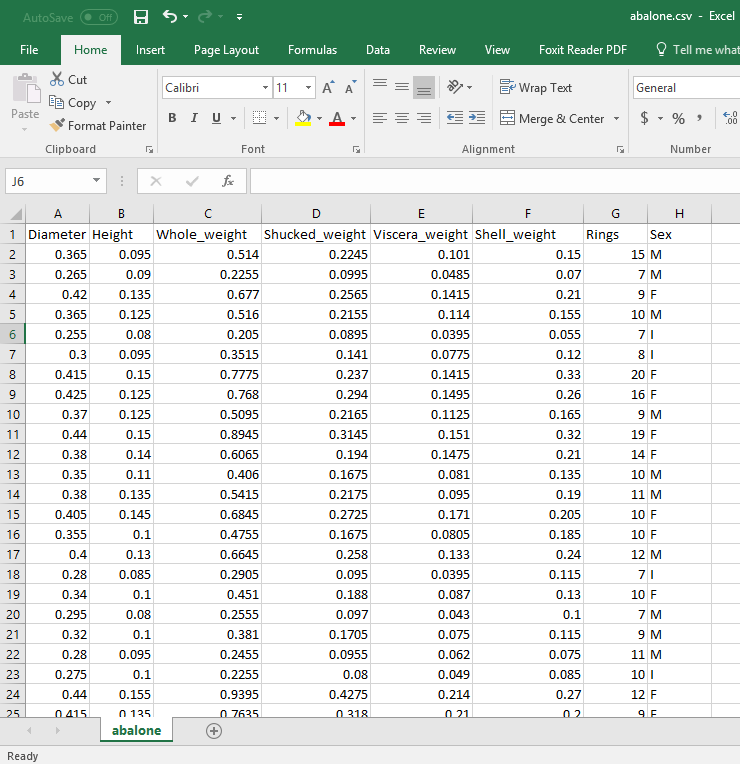
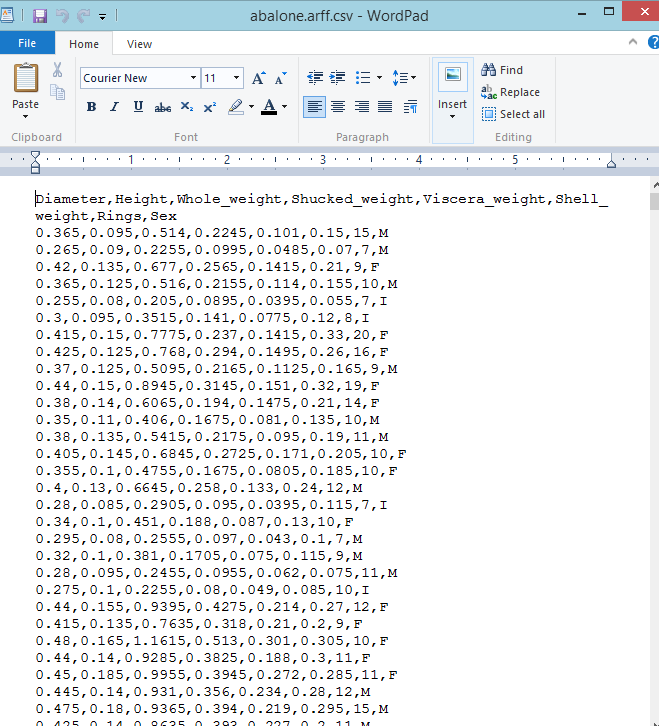
**Question**

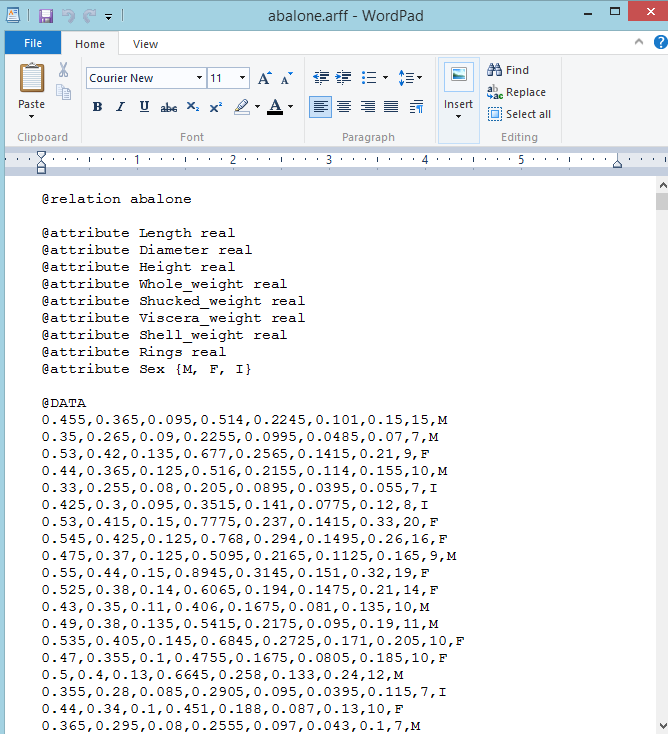
1. Get the "Abalone" dataset from UCI ML Repo
2. Use a decision tree (C4.5/J48) learner in WEKA to predict the number of rings Abalone will have
3. Write the description of your process and methods. What parameters (testing/training size, classification target, subtree raising etc.) did you use? Did you preprocess the dataset? Why did you select those parameters/preprocessing?
4. What were your results? Show what decision trees you found.

**Answer**

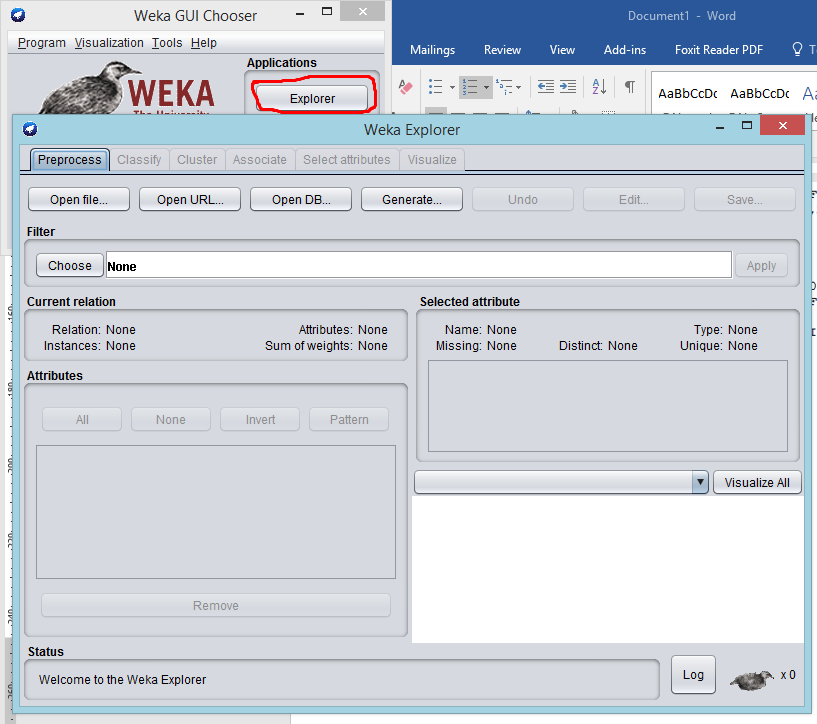
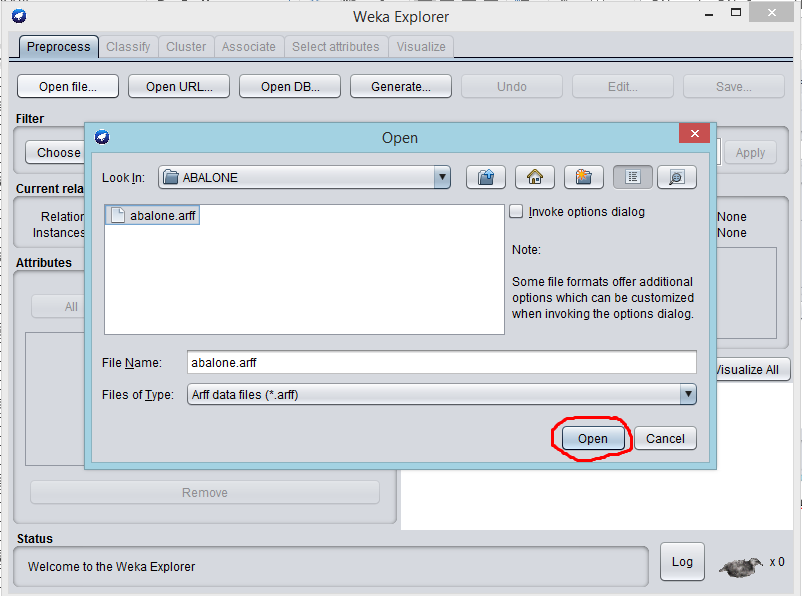
1. **TUTORIAL PADA ABALONE DATA**
2. Download WEKA dari alamat: <https://www.cs.waikato.ac.nz/ml/weka/downloading.html>

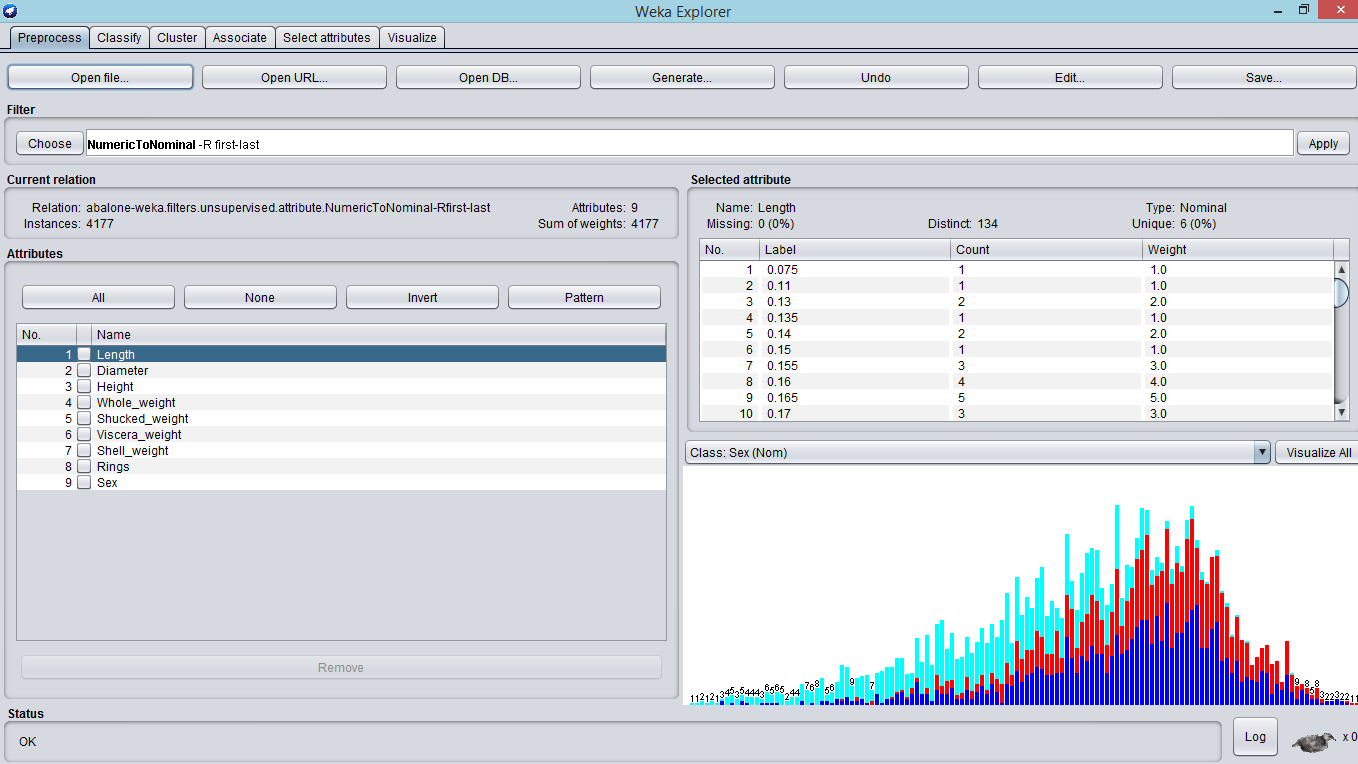
Download sesuai system operasi pada laptop anda.

1. Download data Abalon di UCI Machine Learning Repository(abalone.data) dari alamat: <http://archive.ics.uci.edu/ml/machine-learning-databases/abalone/>
2. Setelah terdownload, rename abalone.data menjadi abalone.csv
3. Langkah selanjutnya adalah buka file abalone.scv dengan Microsoft excel
4. Berilah nama tiap kolom attribute dan class nya sesuai dengan yang ada pada abalone.name pada database UCI.
5. Pindahlah kolom Sex dari paling kiri ke kolom paling kanan(setelah Rings)
6. Langkah selanjutnya adalah save as data dengan nama abalone.arrf dan selanjutnya buka file dengan Wordpad dan tambahkan @attribute dan real sebagai relation sebagai berikut:

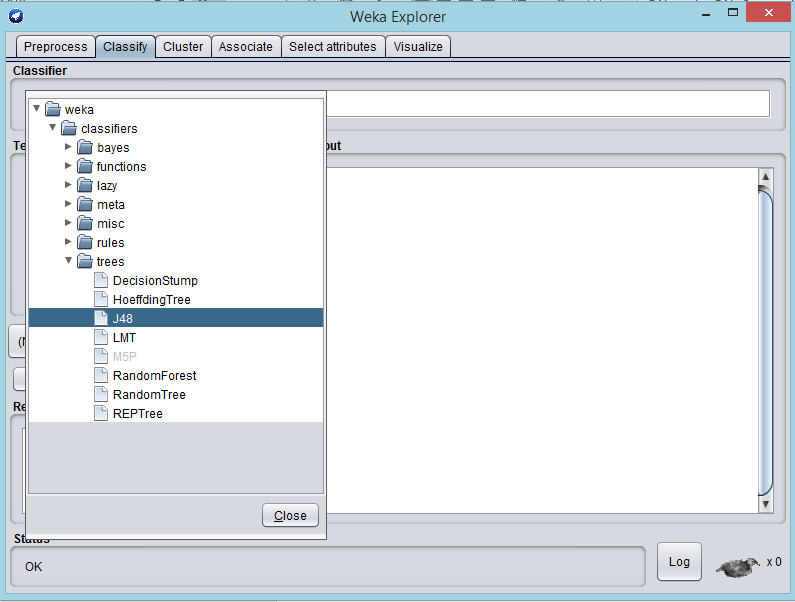
MENJADI

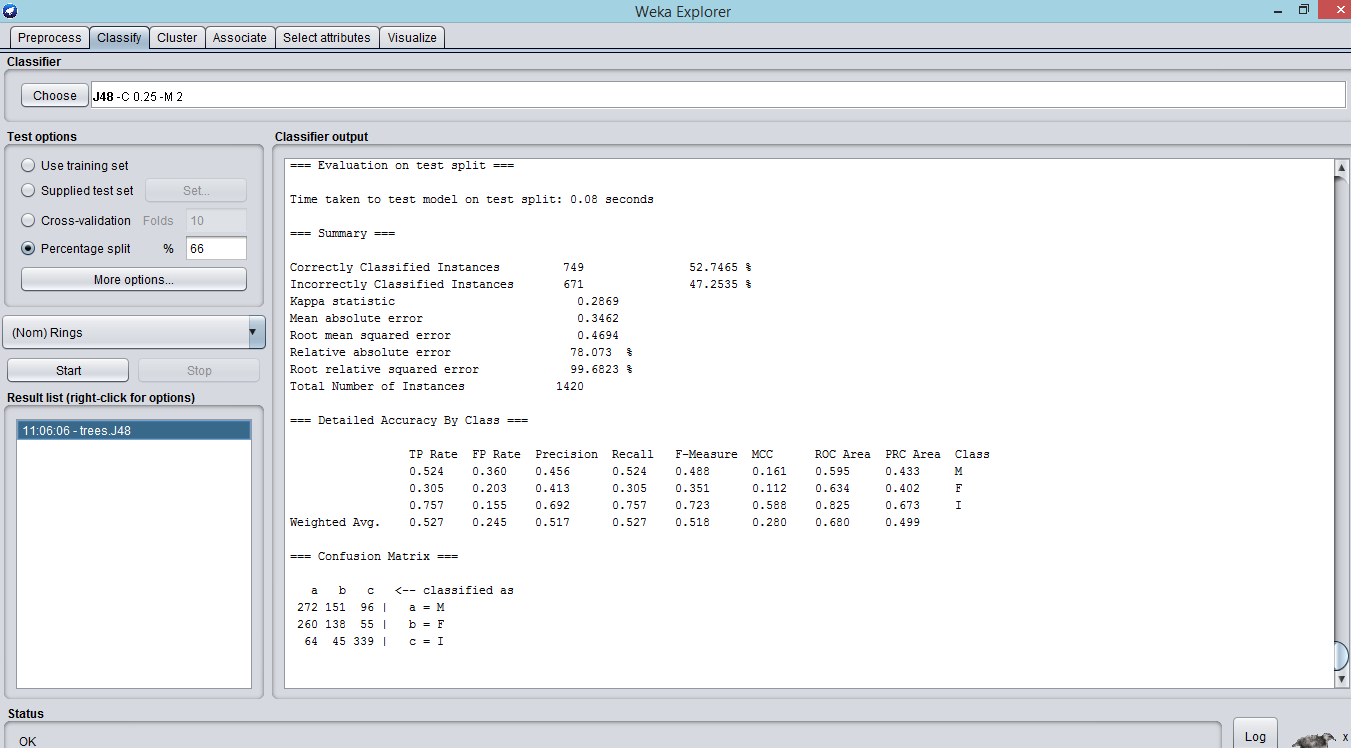
Dan simpan file abalone.arff

1. Langkah selanjutnya buka WEKA
2. Setelah terbuka, load file abalone.arff dan ikuti langkah sebagai berikut:

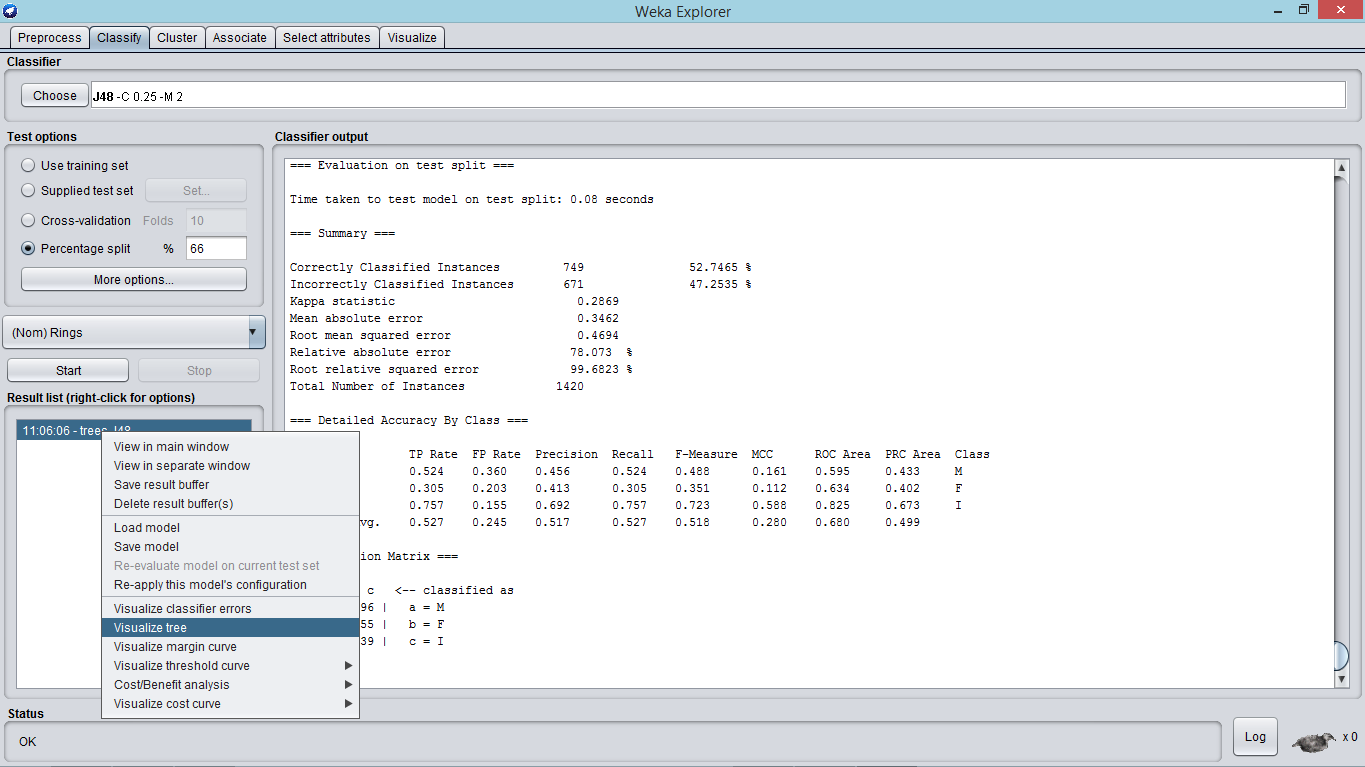


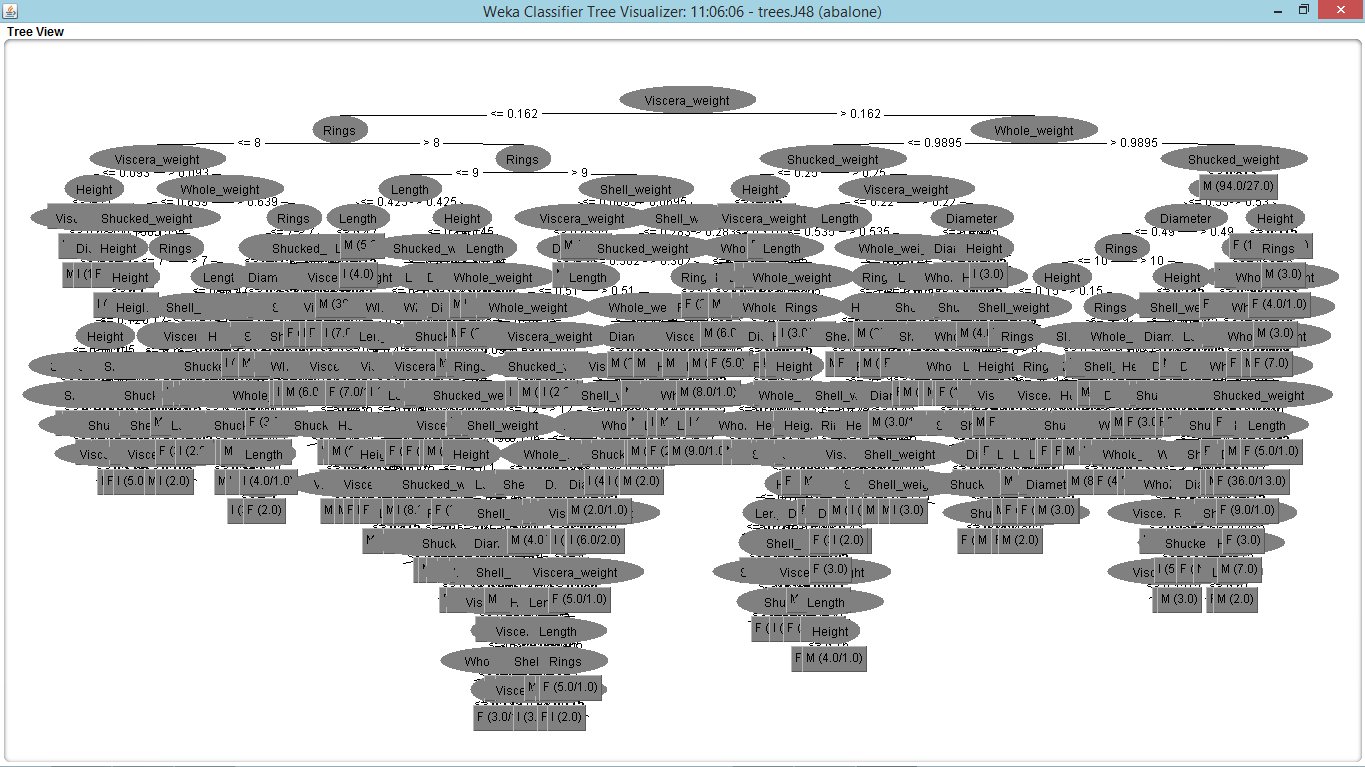
Classify

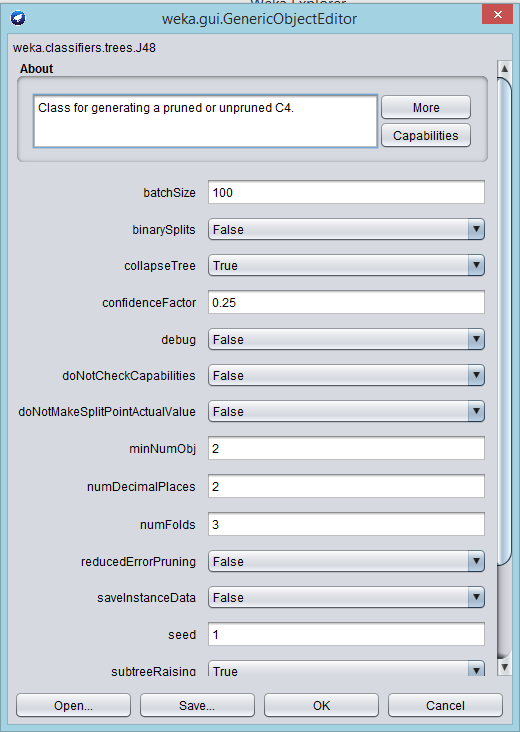




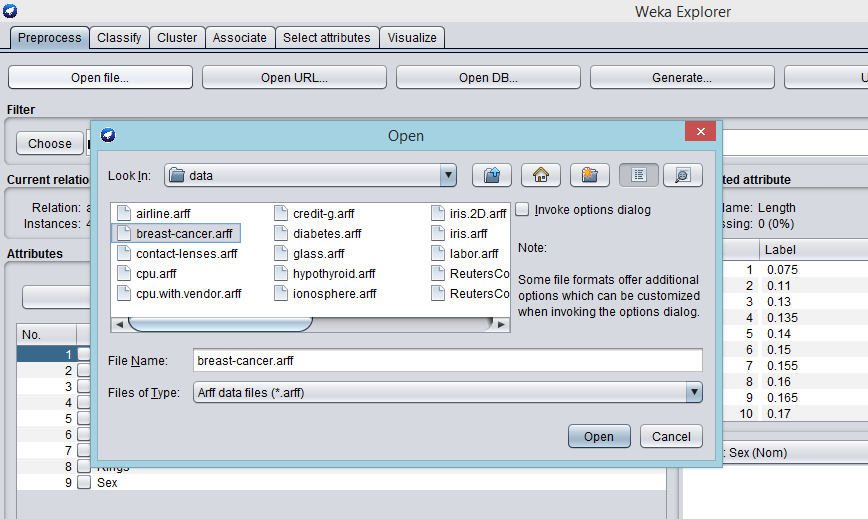
Klik kanan pada result list dan pilih Visualize tree, seperti berikut:

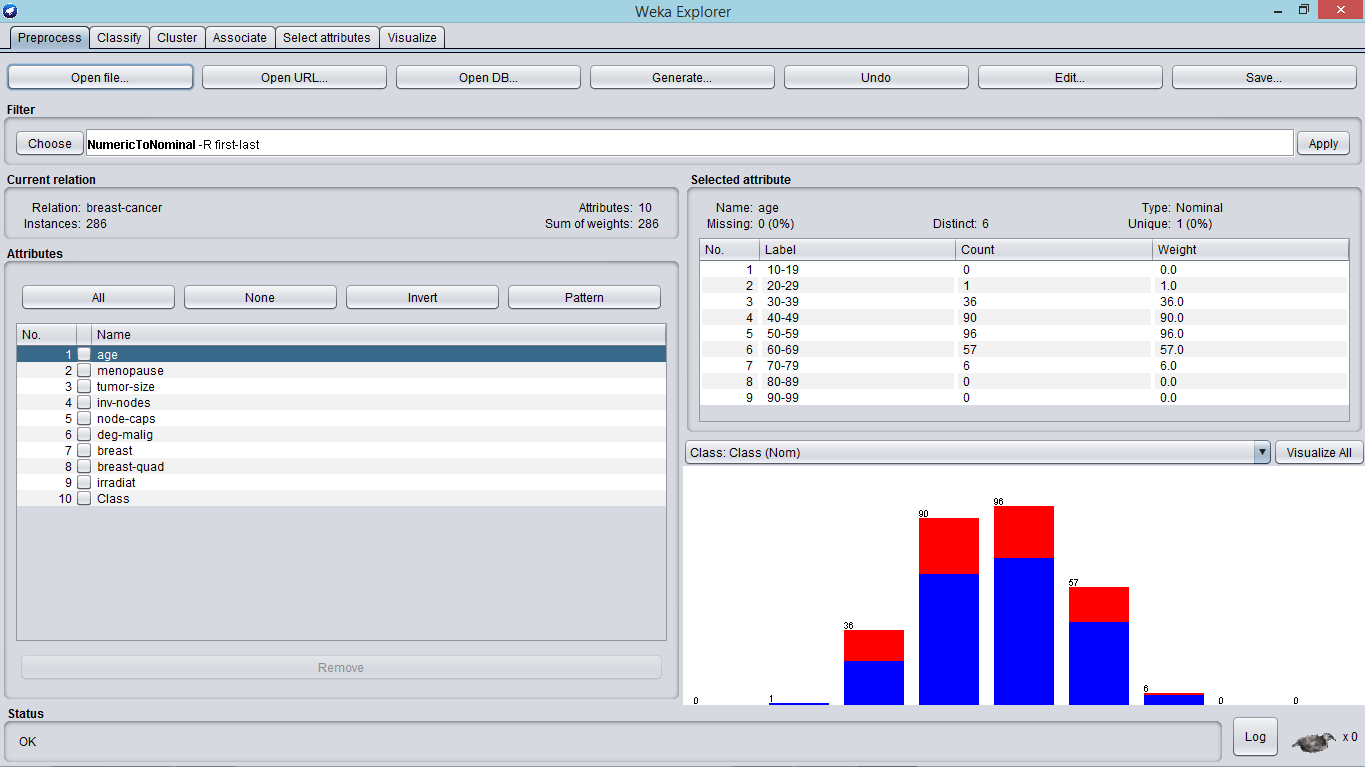


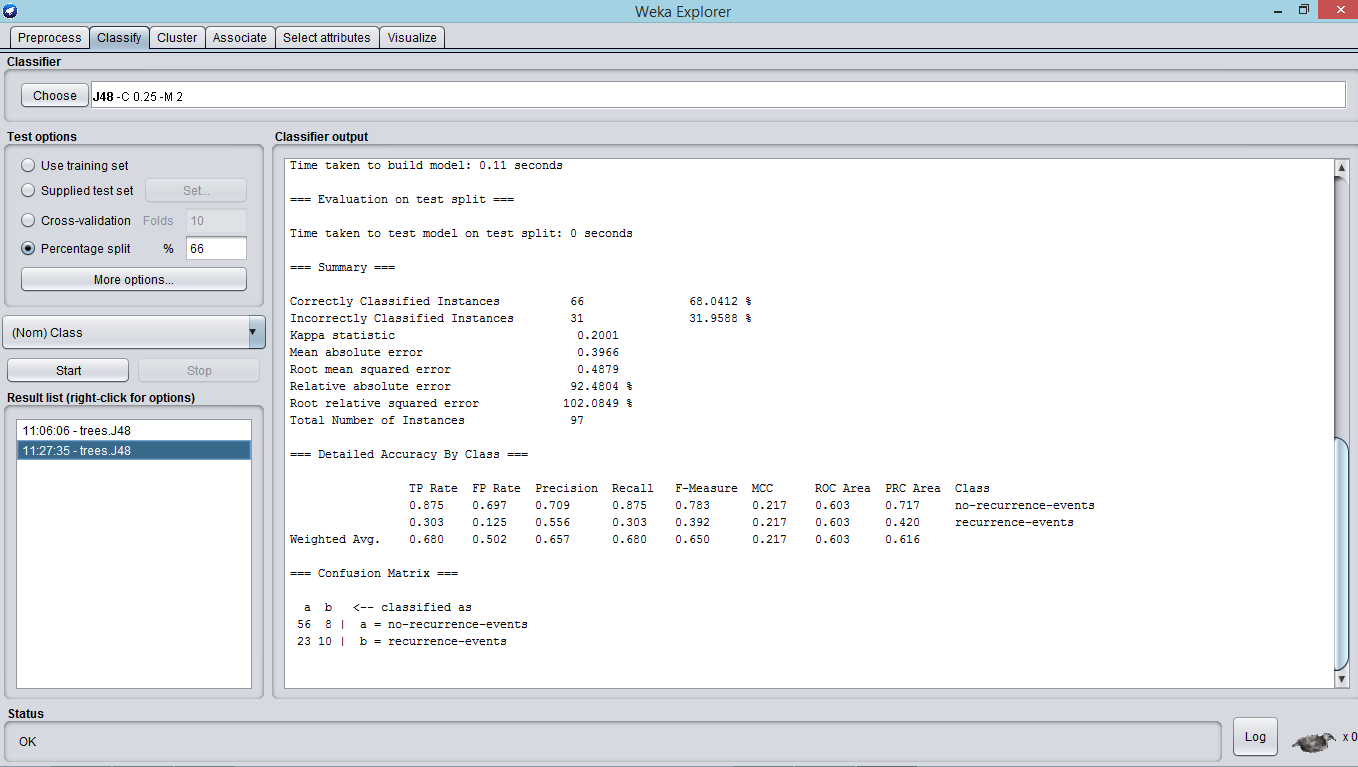


**Parameter**

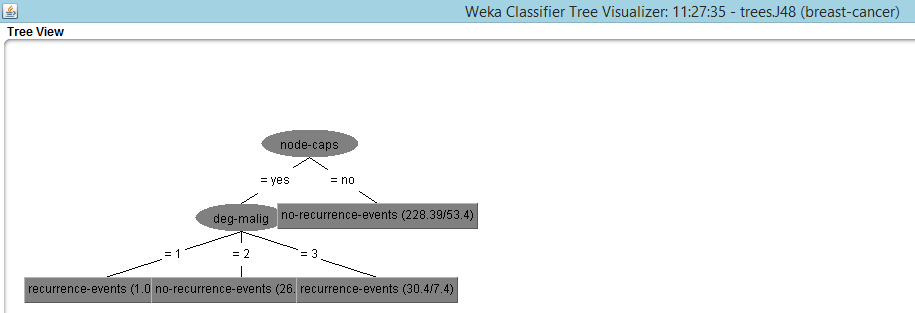
1. **TUTORIAL DATA BREAST CANCER**
2. Untuk data breast cancer, datanya sudah tersedia pada weka jadi langkah-langkahnya tidak perlu mengubah data menjadi .arff karena dalam weka udah menjadi breast-cancer.arff



1. Setelah open file akan muncul gambar seperti di bawah ini:
2. Langkah selanjutnya klik Classify dan ikuti langkah berikut:



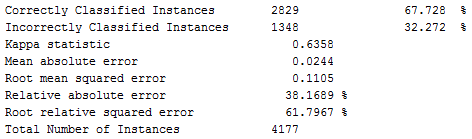
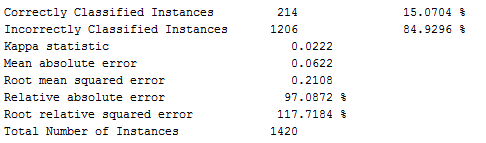
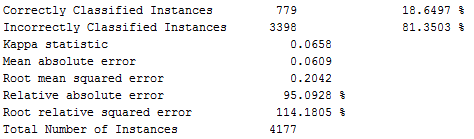
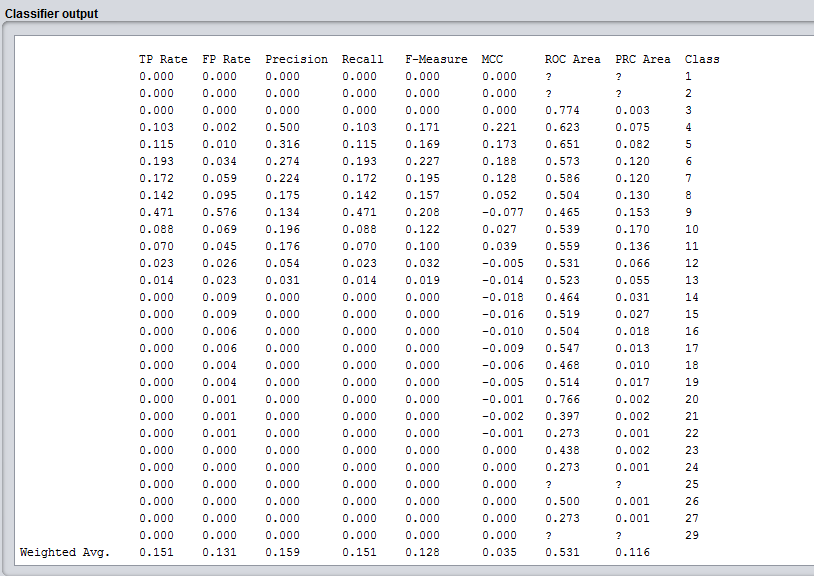
Klik Visualize Tree dari klik kanan pada result dan hasilnya sebagai berikut:



**Question**

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

**Answer**

* Use Training Set
* Cross Validation Folds: 10
* Percantage Splits 66%

Dilihat dari hasil akurasi yang telah diujikan terdapat perbedaan di masing-masing parameter. Hal tersebut terjadi disebabkan karena *misclassified.*