1. **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?**

Proses dan metode yang digunakan adalah mencari akurasi dari sebuah dataset yang berjumlah sebanyak 4000++ data. Dataset yang digunakan adalah data "Abalone" yang didownload dari *https://archive.ics.uci.edu/ml/machine-learning-databases/abalone/abalone.data* dengan format *.arff* supaya bisa ditraining di aplikasi weka. Dikarenakan dataset "Abalone" ini bukan merupakan data kategori nominal (data numerikal), maka data harus difilter terlebih dahulu menjadi data numerikal supaya bisa diklasifikasi menggunakan *Decision Tree*.

**Method** yang digunakan :

1. Evaluation on Test Split (split = 0,66)
2. Stratified K-Fold Cross Validation (K = 10)

**Parameters** yang ada :

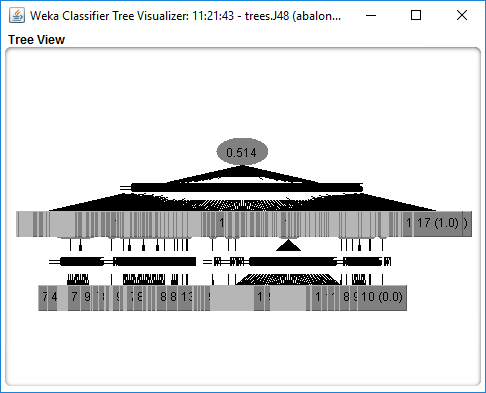
1. Jumlah dataset = 4176 data
2. Training size :

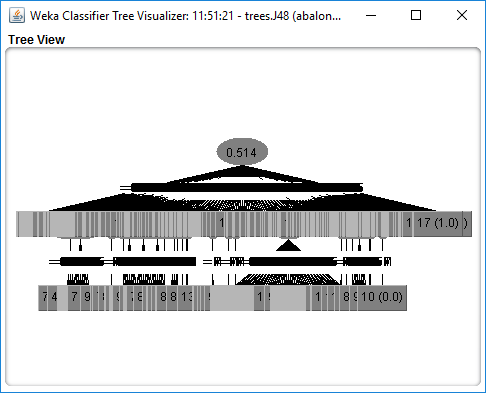
* Test Split = 2756 data
* Cross Validation = 10 Folds dari seluruh dataset

1. Test size

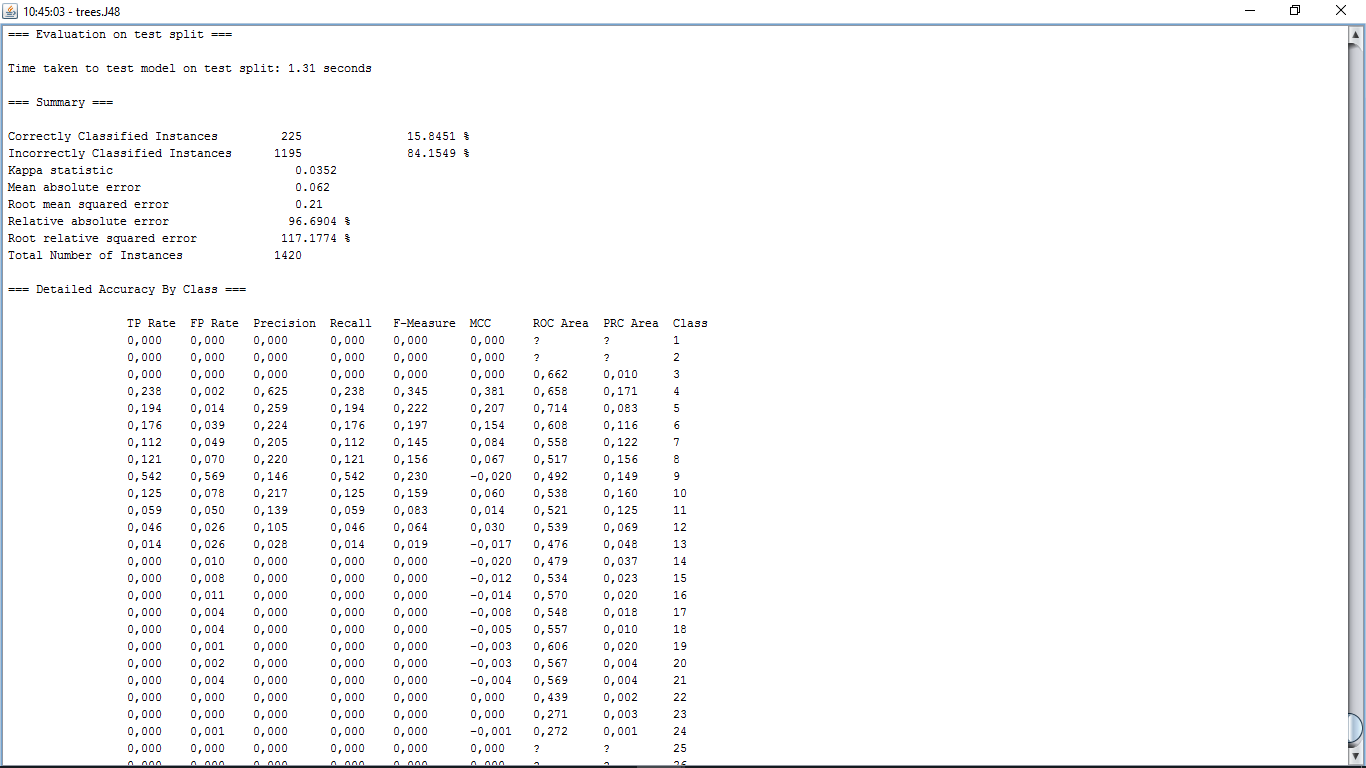
* Test Split = 1420 data
* Cross Validation = 10 Folds dari seluruh dataset

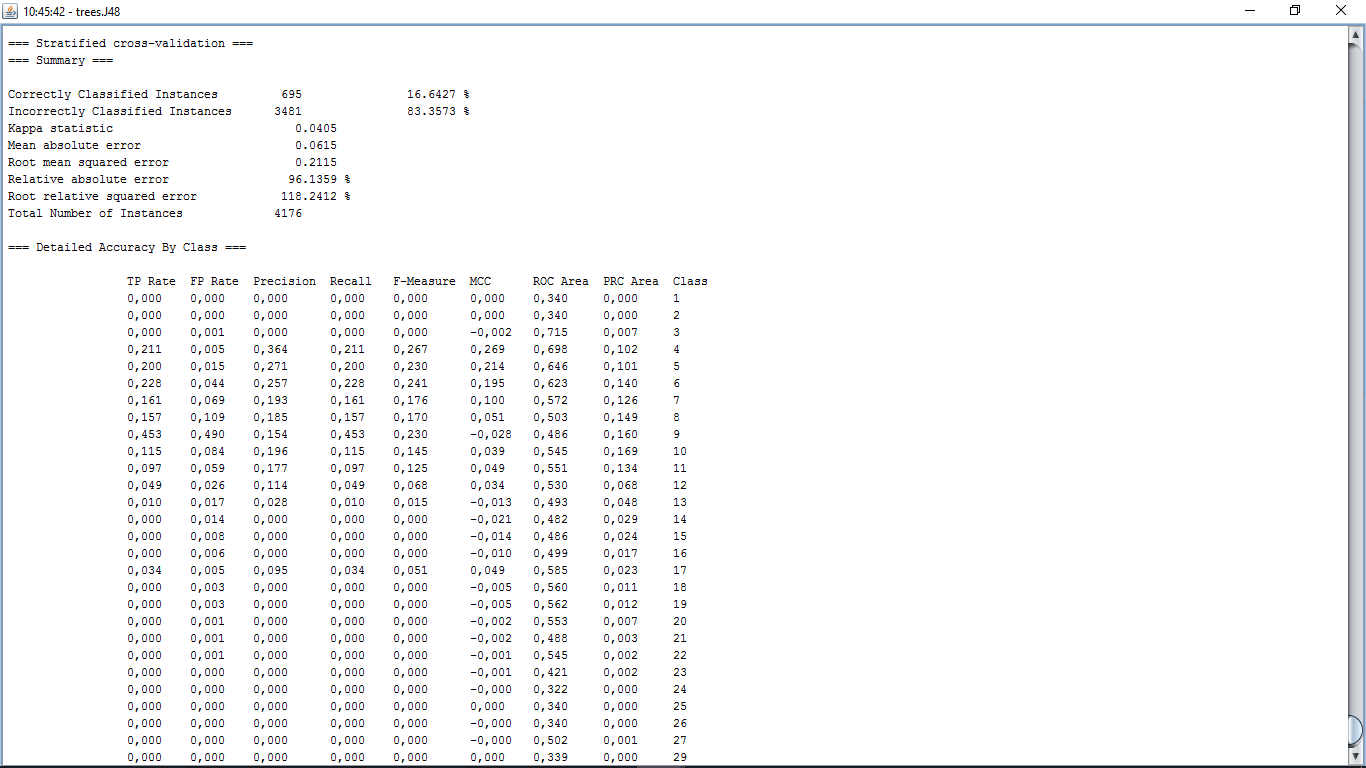
1. Classification Target = Rings.
2. **What were your results? Show what decision trees you found.**





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





Berdasarkan gambar hasil *screenshot* di atas, hasil dari keakuratannya tidak berbeda jauh. Dengan menggunakan Split Test (66%) akurasinya sebesar 15,85% dan dengan menggunakan K-Fold Validation akurasinya sebesar 16,64%. Kenapa saat merubah parameter akurasinya bisa berubah? Karena dalam Decision Tree apabila sudah menklasifikasikan dengan perbandingan tiap atributnya terkadang masih belum ditemukan hasil pastinya (masih bisa dikatakan *gambling*) jadi kita perlu mencari alternatif dengan merubah parameter lainnya.