1. Describe your process and methods. What parameters (testing/training size, classification target, subtree raising etc.) did you try using with the decision tree? What, if anything, did you do to preprocess the dataset? Why did you select those parameters/preprocessing?
   1. **Get the dataset from** [**https://archive.ics.uci.edu/ml/machine-learning-databases/abalone/**](https://archive.ics.uci.edu/ml/machine-learning-databases/abalone/)
   2. **Edit text in abalone.data to get the attributes and relation to decision tree**
   3. **Add this text to the file abalone.data at the top of file**

**@relation abalone**

**@attribute Sex {M, F, I}**

**@attribute Length real**

**@attribute Diameter real**

**@attribute Height real**

**@attribute Whole real**

**@attribute Shucked real**

**@attribute Viscera real**

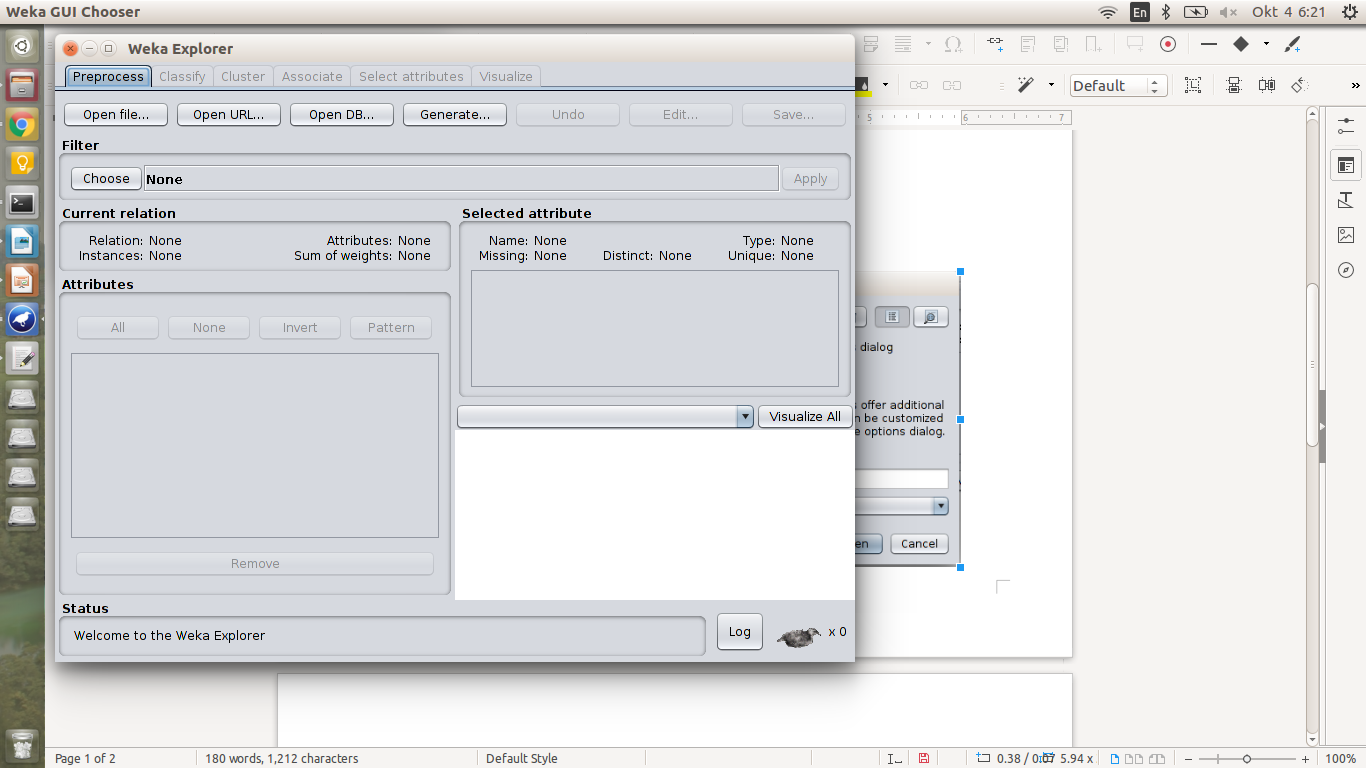
**@attribute Shell real**

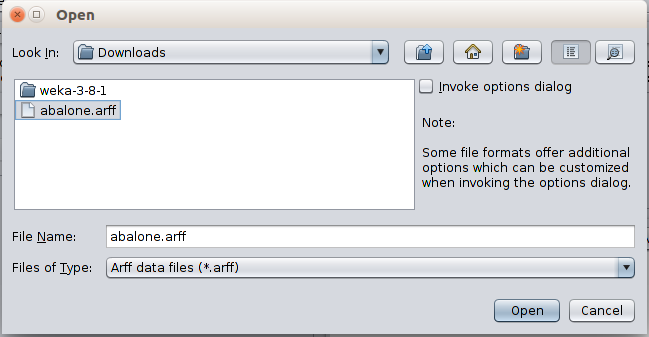
**@attribute Rings real**

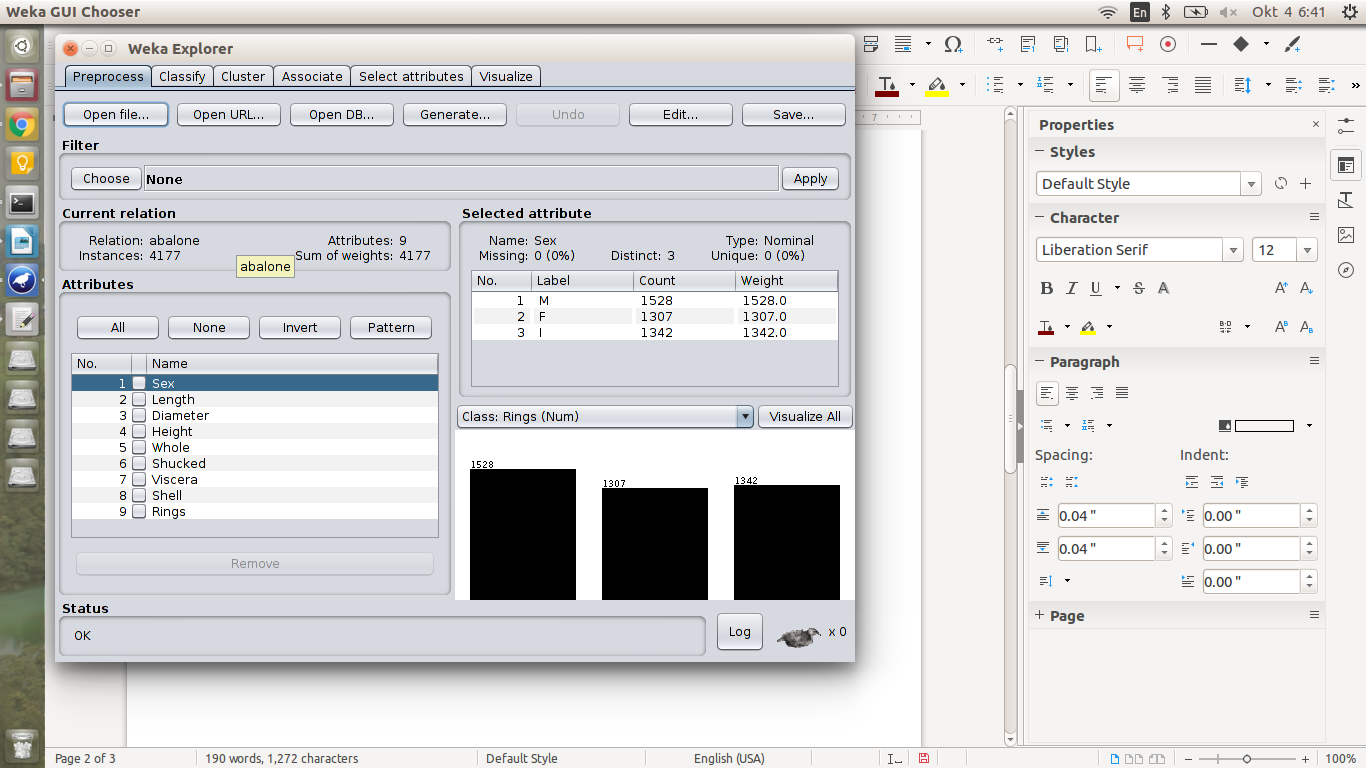
**@DATA**

**to define attributes name for decision tree. Save as .arff file.**

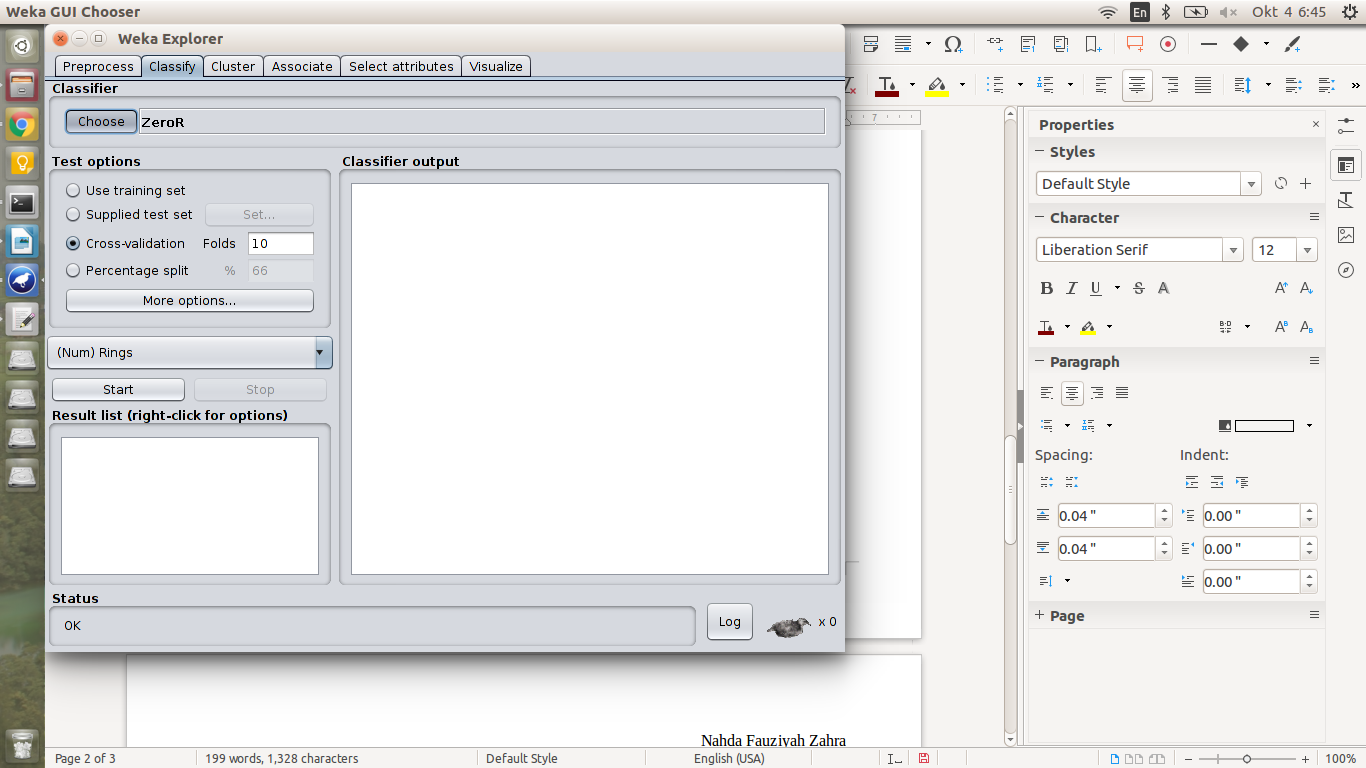
* 1. **Here’s the step to make decision tree using weka**

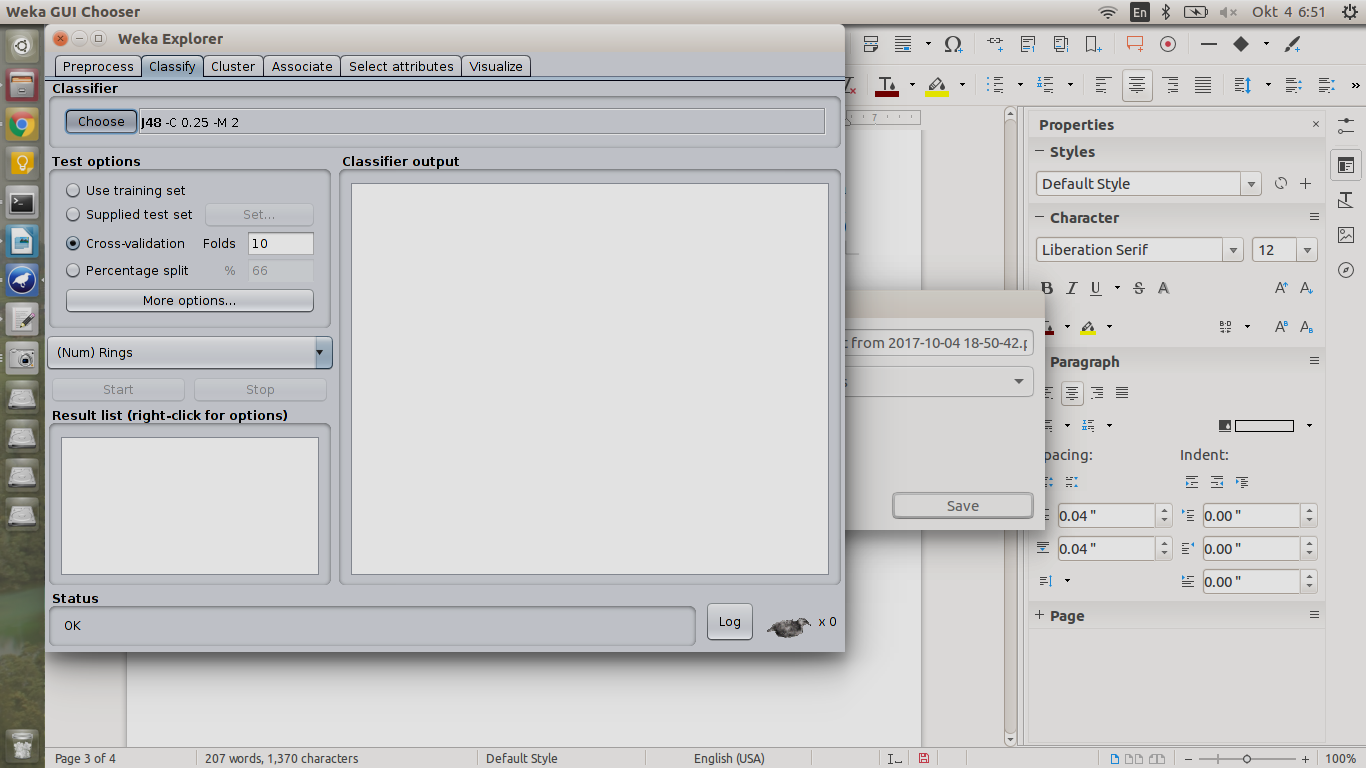
**choose dataset file**

****

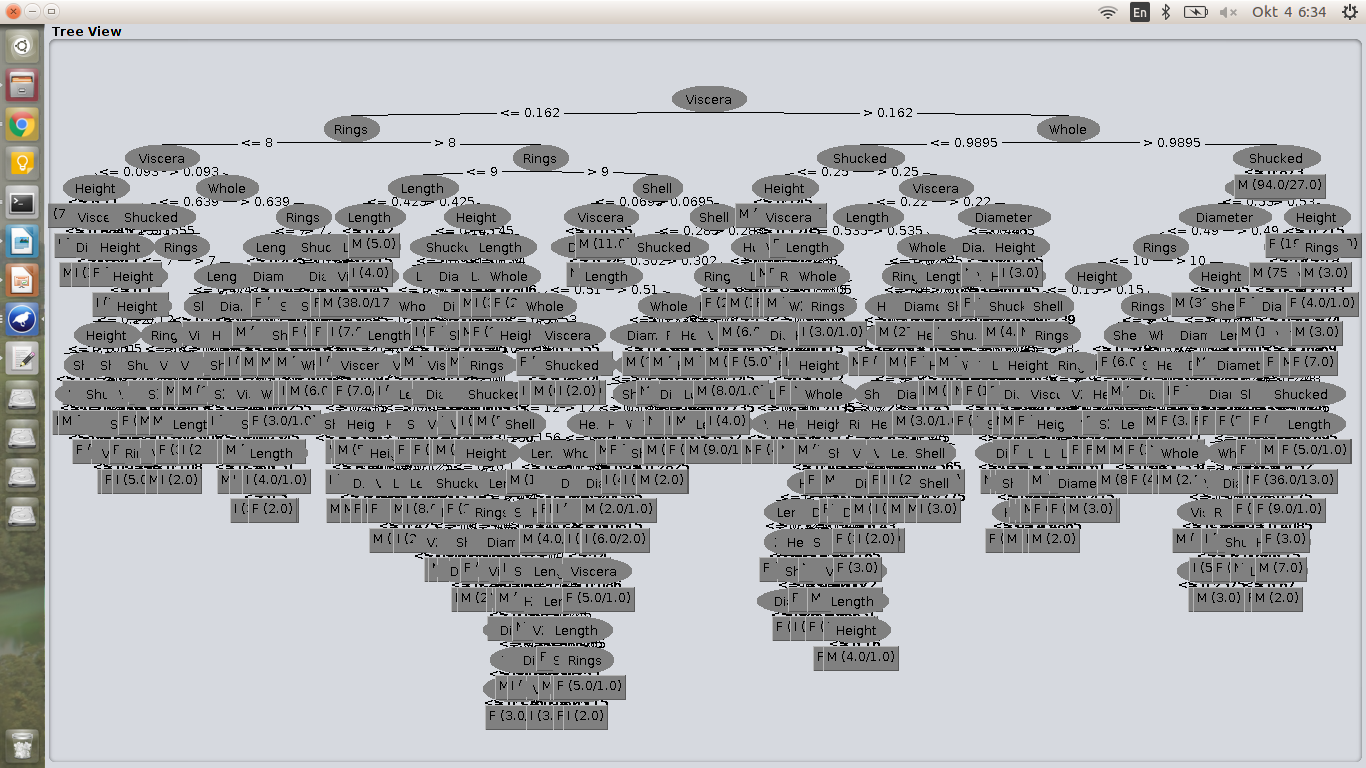
****

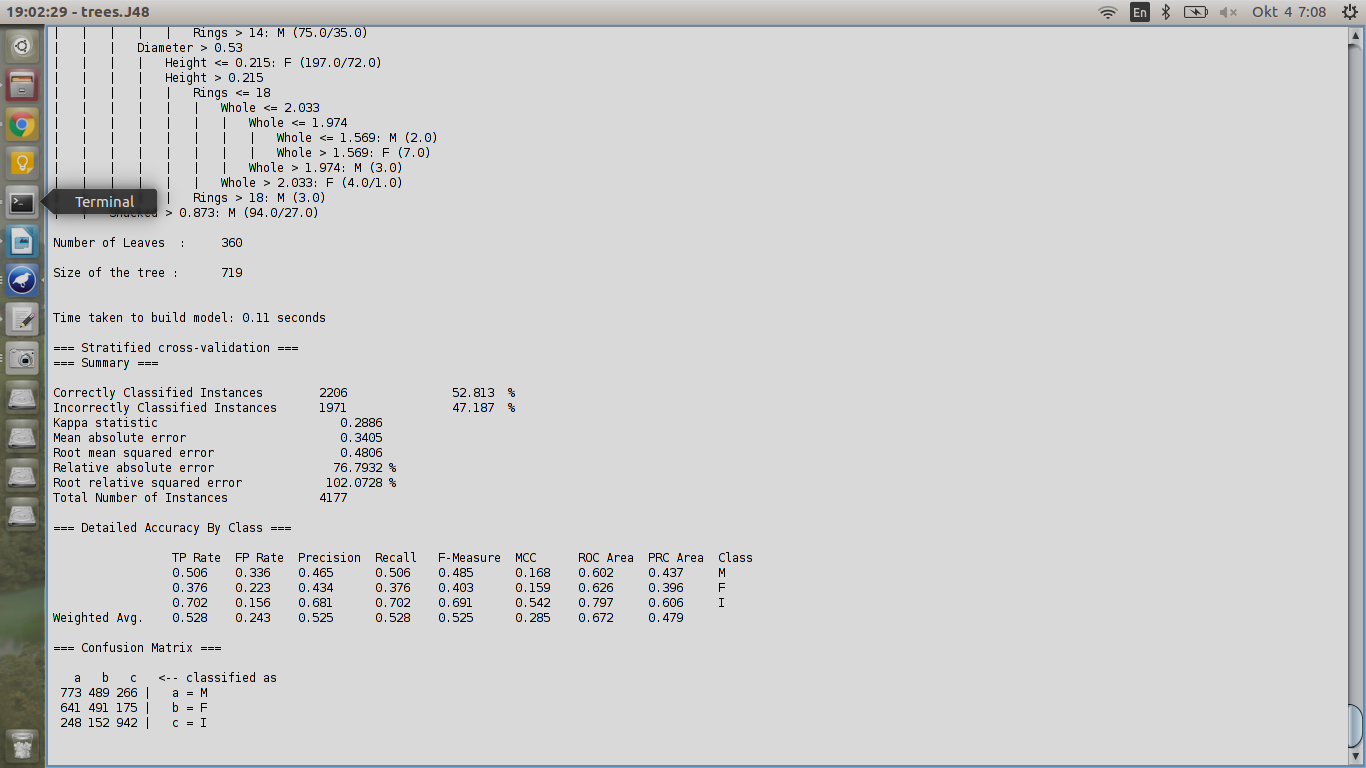
**result after choosing dataset file**

**move to “Classify” tab**

**choose what classifier that we want to use, the example test using cross-validation with 10 folds**

1. What were your results? (Overall accuracy, accuracy of each leaf/branch, false positives, false negatives, etc. on the training and testing sets). Show what decision trees you found.

**decision tree result**

**classifier result**

1. What do the results tell us? Why are the results (in)accurate? What class were some samples misclassified as, and why? Why did changing parameter(s) improve/degrade accuracy?

**Because each parameter has type of itself data class, so the accuracy value may be changed.**