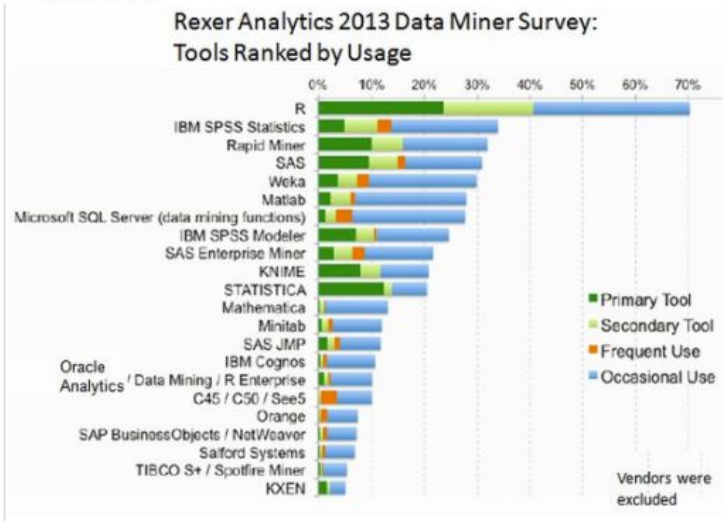
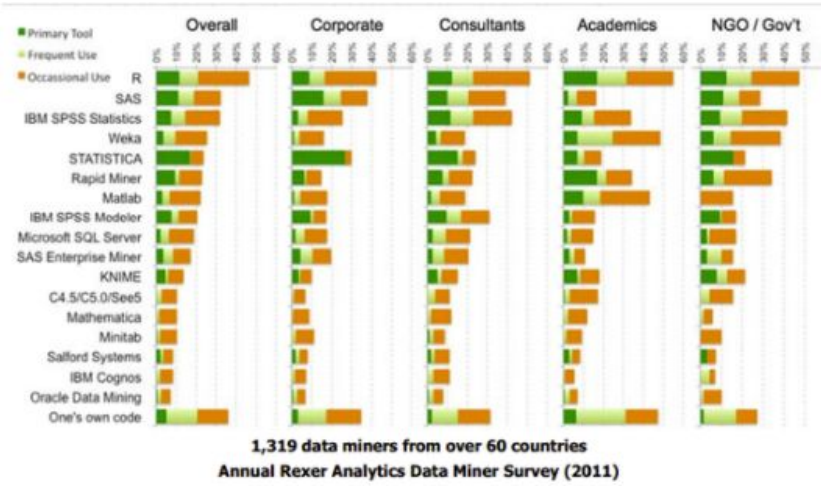
1.- Explain the advantages and disadvantages of writing a program on your own vs using a pre-created suite such as WEKA.

a. In this article: https://www.quora.com/Which-is-better-R-or-Weka they compared the industry tendencies for data mining of 2011 and 2013.. The results are as follow:

On the same forum, they all came to the conclusion (as our own team did) that maybe if WEKA is quite a versatile tool for ML, it's quite inflexible, WEKA is a ready to deploy tool already compiled in JAR quite useful when wanting to do quick queries with no need for customization or optimization but with a friendly user interface (for what a ML tool can call friendly). So basically if you program your own is because you want either: performance, specific algorithms, batch tools or specific OS requirements. But if you don't have time or have a task accomplishable by WEKA then it's quite useful.

Now, look at 2013:





2. Explain what criteria you followed to choose the data sets for your tree and the WEKA tests.

After looking at ML forums, this seemed a recurrent dataset page http://archive.ics.uci.edu/ml/datasets.html so we basically took a classification one which apparently has been cited many times on different articles, so we tested it against our implementation. But then we realized they didn't work. Why? After trying with different datasets we came to the conclusion that our implementation only works with only nominal, and complete datasets. A Lot of WEKA datasets non complete ? datasets or also continuous data instead of discrete/categorical datasets. This narrowed down the datasets quite many, so we went with the typical "contact lense" example.

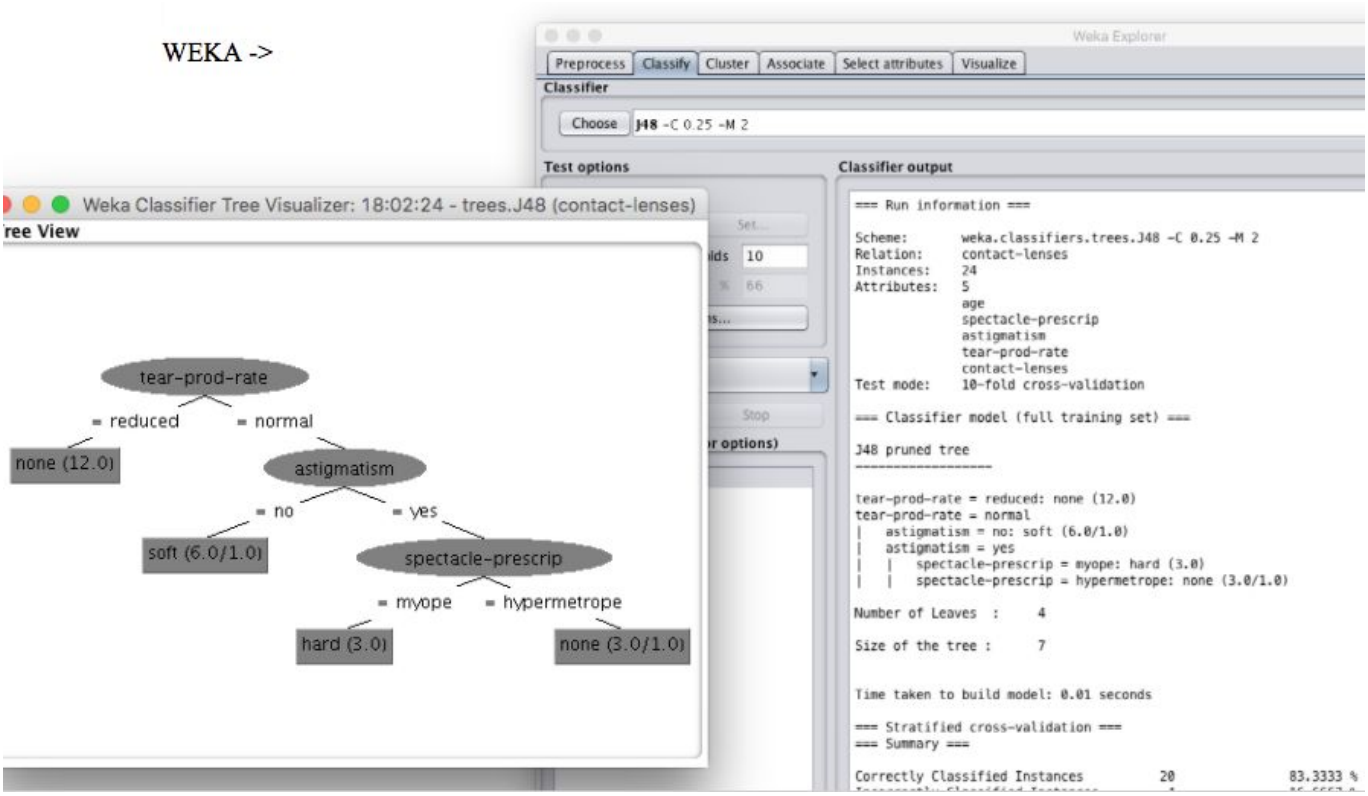
3. Include the graphics of the trees or part of the trees you generated in WEKA and your own program. Are they different, and if so, why?

IALab5 — -bash — 127×40

```
Albertos-MacBook-Pro:IALab5 albertolopez$ ./run < /Users/albertolopez/Desktop/IALab5/testcases/contact-lenses.arff.in
tear-prod-rate: reduced
  ANSWER: none
tear-prod-rate: normal
  astigmatism: no
    age: young
      ANSWER: soft
    age: pre-presbyopic
      ANSWER: soft
    age: presbyopic
      spectacle-prescrip: myope
        ANSWER: none
      spectacle-prescrip: hypermetrope
        ANSWER: soft
  astigmatism: yes
    spectacle-prescrip: myope
      ANSWER: hard
    spectacle-prescrip: hypermetrope
      age: young
        ANSWER: hard
      age: pre-presbyopic
        ANSWER: none
      age: presbyopic
        ANSWER: none
Albertos-MacBook-Pro:IALab5 albertolopez$
```

<- Our implementation

WEKA ->



Results: Weka doesn't take into consideration the age.. Which is weird. WEKA is also able to chose from a big list of algorithms and methods from branches like Bayes, lazy, meta identification, rules and trees with tons of ways of doing the same thing.

4) Based in what you have learned so far where would you use decision trees?

b. For almost anything in which we have pre labeled examples as well as a somewhat extensive set of learning dataset and we aren't confident enough of what to do.. Wether were to go dining , if I'm susceptible for a credit card, more advanced stuff such as where to invest on the stock market depending on previous events and how they've affected the market. If we go into a more computer science approach, several online shopping companies use decision trees to specify whether they should offer specific offers to a type of customer based on his habits (time online, frequency of buying online, etc) to actually make effective a sale; other application is fault diagnosis on engineering , fraud detection and business management. More can be found here http://what-when-how.com/artificial-intelligence/decision-tree-applications-for-data-modelling-artificial-intelligence/