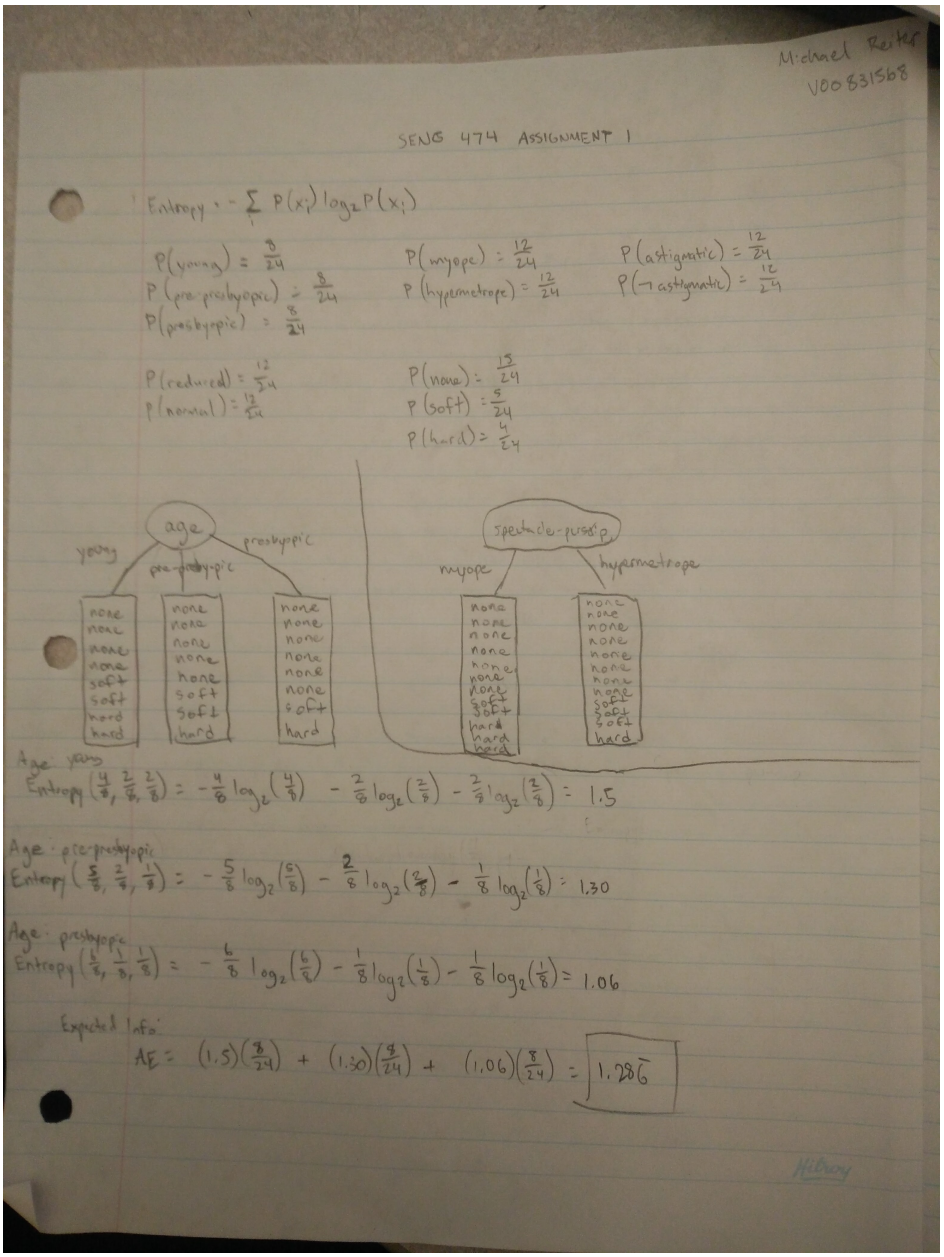
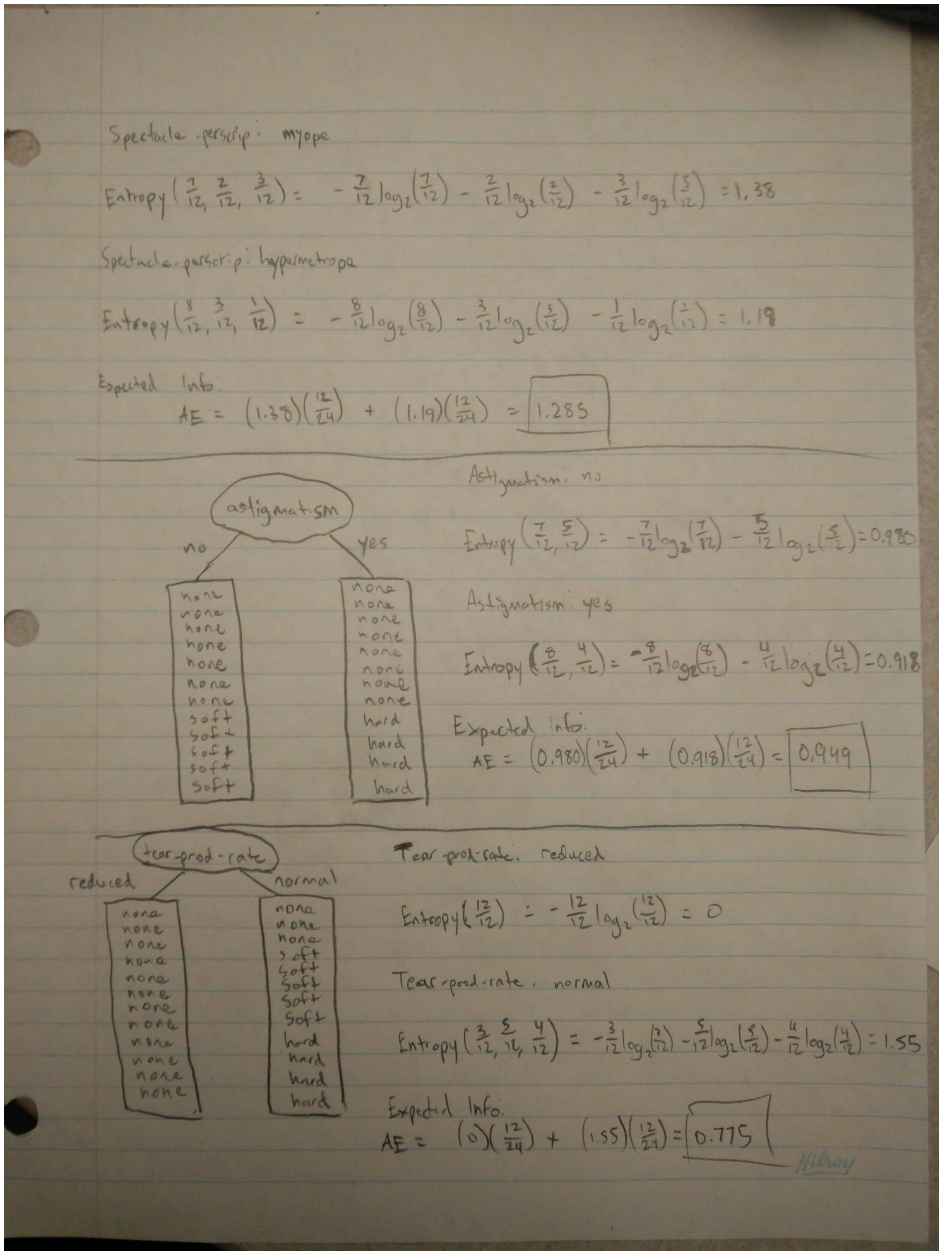
Michael Reiter
V00831568

SENG 474 Assignment 1

1. a)



Michael Reiter
V00831568



Spectacle-perscrip: myope

$$\text{Entropy}\left(\frac{7}{12}, \frac{2}{12}, \frac{3}{12}\right) = -\frac{7}{12}\log_2\left(\frac{7}{12}\right) - \frac{2}{12}\log_2\left(\frac{2}{12}\right) - \frac{3}{12}\log_2\left(\frac{3}{12}\right) = 1.38$$

Spectacle-perscrip: hypermetrope

$$\text{Entropy}\left(\frac{8}{12}, \frac{3}{12}, \frac{1}{12}\right) = -\frac{8}{12}\log_2\left(\frac{8}{12}\right) - \frac{3}{12}\log_2\left(\frac{3}{12}\right) - \frac{1}{12}\log_2\left(\frac{1}{12}\right) = 1.19$$

Expected info:

$$AE = (1.38)\left(\frac{12}{24}\right) + (1.19)\left(\frac{12}{24}\right) = \boxed{1.285}$$

Astigmatism: no

$$\text{Entropy}\left(\frac{7}{12}, \frac{5}{12}\right) = -\frac{7}{12}\log_2\left(\frac{7}{12}\right) - \frac{5}{12}\log_2\left(\frac{5}{12}\right) = 0.980$$

Astigmatism: yes

$$\text{Entropy}\left(\frac{8}{12}, \frac{4}{12}\right) = -\frac{8}{12}\log_2\left(\frac{8}{12}\right) - \frac{4}{12}\log_2\left(\frac{4}{12}\right) = 0.918$$

Expected info:

$$AE = (0.980)\left(\frac{12}{24}\right) + (0.918)\left(\frac{12}{24}\right) = \boxed{0.949}$$

Tear-prod-rate: reduced

$$\text{Entropy}\left(\frac{12}{12}\right) = -\frac{12}{12}\log_2\left(\frac{12}{12}\right) = 0$$

Tear-prod-rate: normal

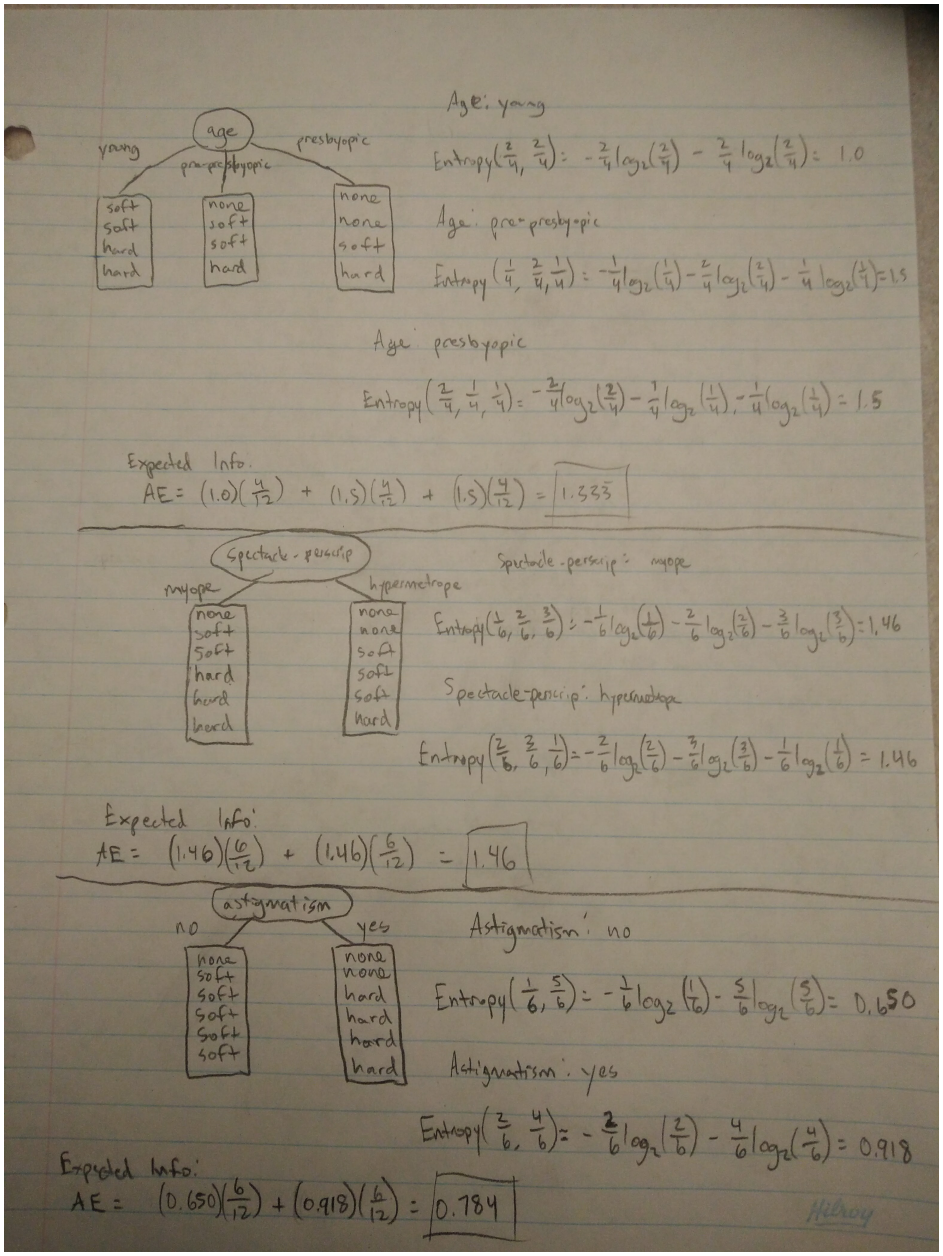
$$\text{Entropy}\left(\frac{3}{12}, \frac{5}{12}, \frac{4}{12}\right) = -\frac{3}{12}\log_2\left(\frac{3}{12}\right) - \frac{5}{12}\log_2\left(\frac{5}{12}\right) - \frac{4}{12}\log_2\left(\frac{4}{12}\right) = 1.55$$

Expected info:

$$AE = (0)\left(\frac{12}{24}\right) + (1.55)\left(\frac{12}{24}\right) = \boxed{0.775}$$

Milroy

Michael Reiter
V00831568



Age: young

$\text{Entropy}(\frac{2}{4}, \frac{2}{4}) = -\frac{2}{4}\log_2(\frac{2}{4}) - \frac{2}{4}\log_2(\frac{2}{4}) = 1.0$

Age: pre-presbyopic

$\text{Entropy}(\frac{1}{4}, \frac{2}{4}, \frac{1}{4}) = -\frac{1}{4}\log_2(\frac{1}{4}) - \frac{2}{4}\log_2(\frac{2}{4}) - \frac{1}{4}\log_2(\frac{1}{4}) = 1.5$

Age: presbyopic

$\text{Entropy}(\frac{2}{4}, \frac{1}{4}, \frac{1}{4}) = -\frac{2}{4}\log_2(\frac{2}{4}) - \frac{1}{4}\log_2(\frac{1}{4}) - \frac{1}{4}\log_2(\frac{1}{4}) = 1.5$

Expected Info.

$AE = (1.0)(\frac{4}{12}) + (1.5)(\frac{4}{12}) + (1.5)(\frac{4}{12}) = \boxed{1.333}$

Spectacle-prescrip: myope

$\text{Entropy}(\frac{1}{6}, \frac{2}{6}, \frac{3}{6}) = -\frac{1}{6}\log_2(\frac{1}{6}) - \frac{2}{6}\log_2(\frac{2}{6}) - \frac{3}{6}\log_2(\frac{3}{6}) = 1.46$

Spectacle-prescrip: hypermetrope

$\text{Entropy}(\frac{2}{6}, \frac{3}{6}, \frac{1}{6}) = -\frac{2}{6}\log_2(\frac{2}{6}) - \frac{3}{6}\log_2(\frac{3}{6}) - \frac{1}{6}\log_2(\frac{1}{6}) = 1.46$

Expected Info:

$AE = (1.46)(\frac{6}{12}) + (1.46)(\frac{6}{12}) = \boxed{1.46}$

Astigmatism: no

$\text{Entropy}(\frac{1}{6}, \frac{5}{6}) = -\frac{1}{6}\log_2(\frac{1}{6}) - \frac{5}{6}\log_2(\frac{5}{6}) = 0.650$

Astigmatism: yes

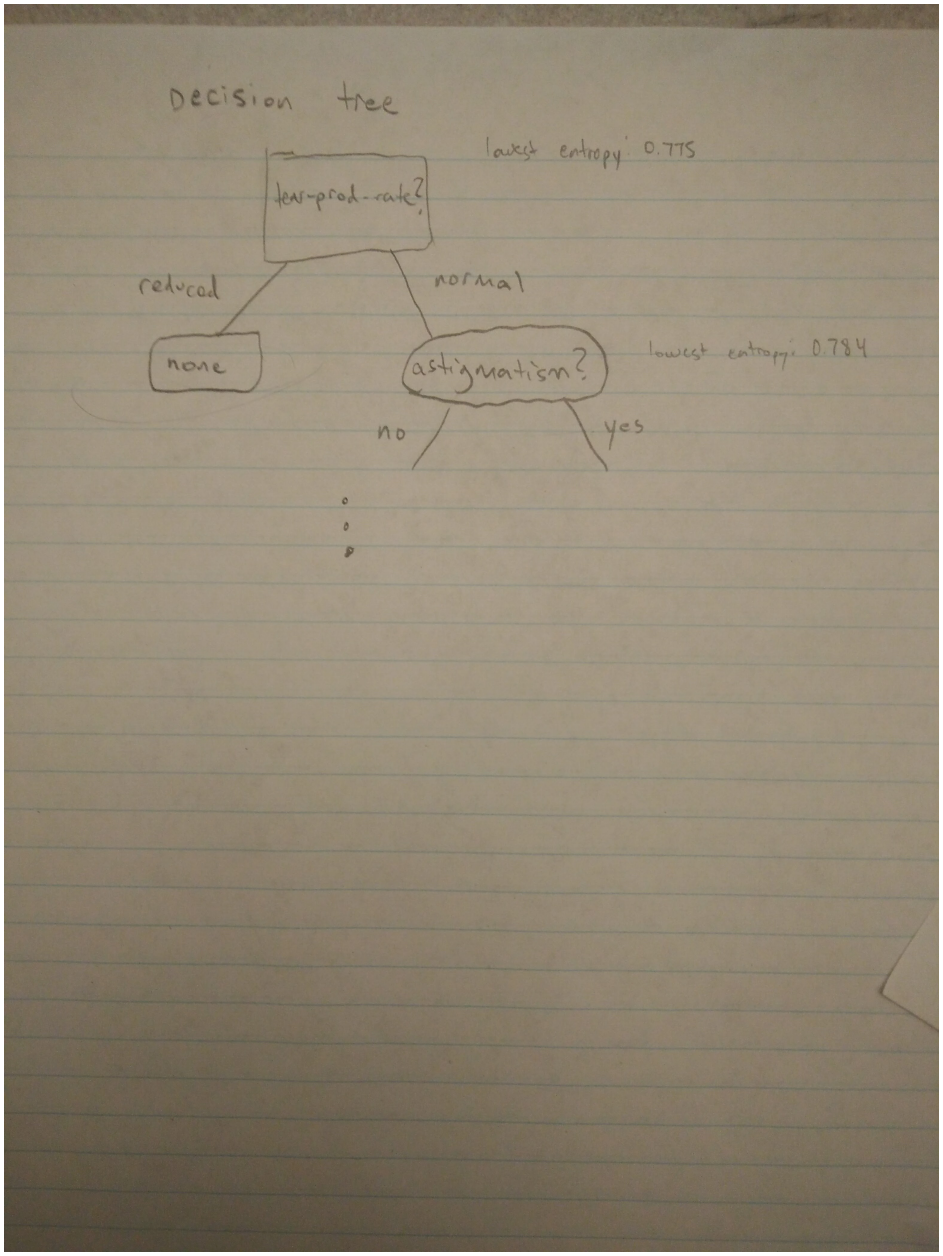
$\text{Entropy}(\frac{2}{6}, \frac{4}{6}) = -\frac{2}{6}\log_2(\frac{2}{6}) - \frac{4}{6}\log_2(\frac{4}{6}) = 0.918$

Expected Info:

$AE = (0.650)(\frac{6}{12}) + (0.918)(\frac{6}{12}) = \boxed{0.784}$

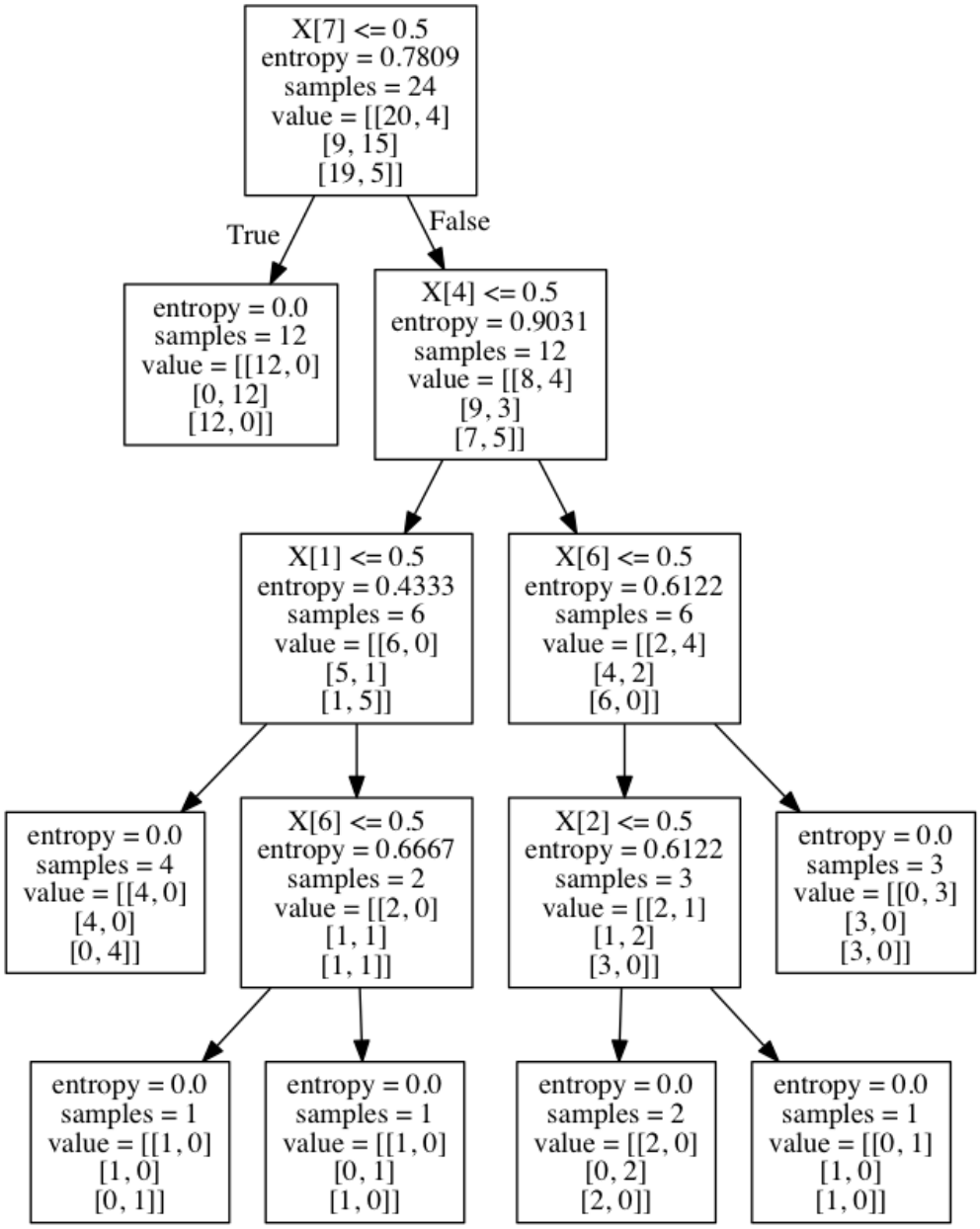
Hilroy

Michael Reiter
V00831568



Michael Reiter
V00831568

1. b)



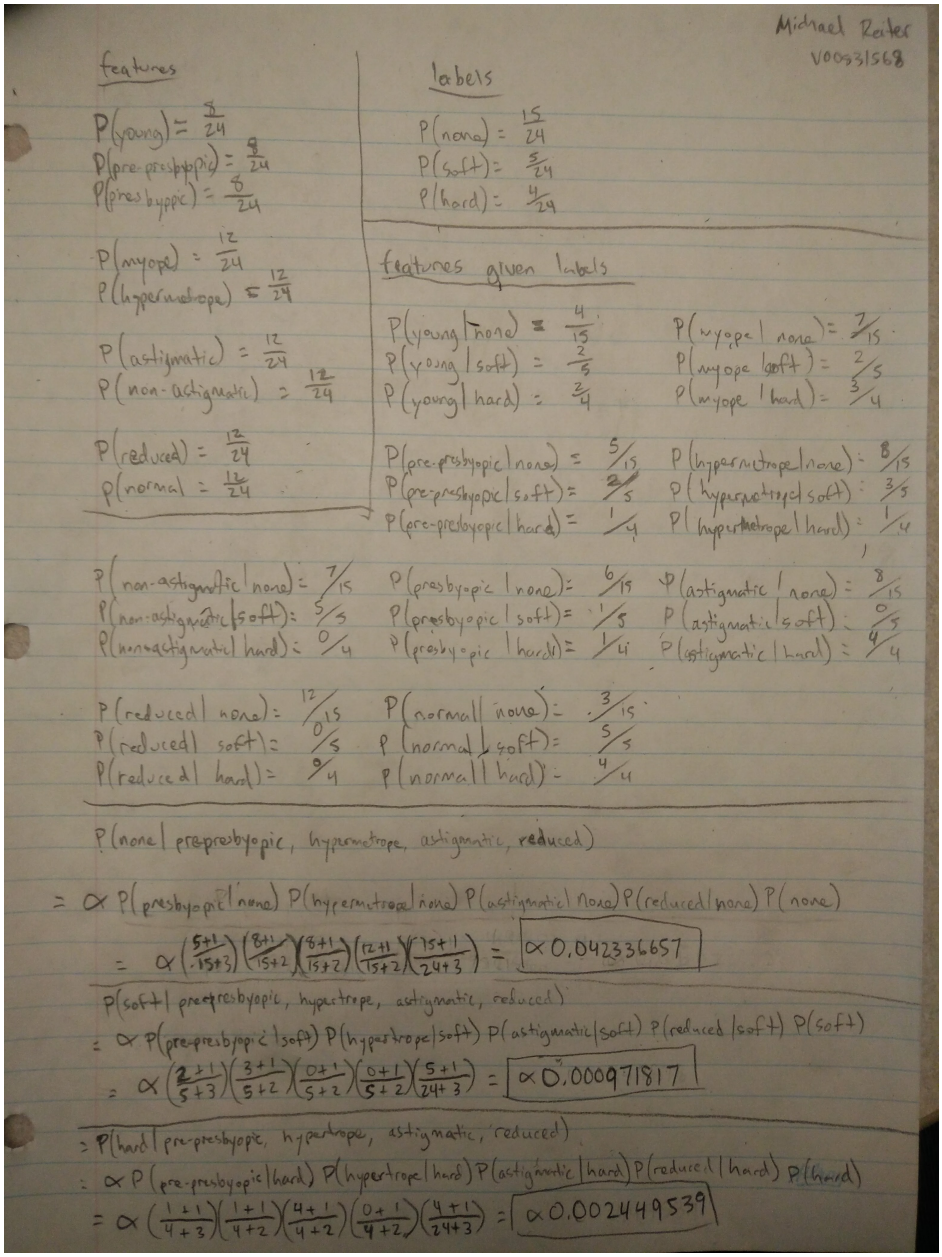
This is a decision tree generated by the following code.

```
from util2 import Arff2Skl
from sklearn import tree
cvt = Arff2Skl('contact-lenses.arff')
label = cvt.meta.names()[-1]
X, y = cvt.transform(label)
dtc = tree.DecisionTreeClassifier(criterion='entropy')
dtc.fit(X, y)
tree.export_graphviz(dtc)
```

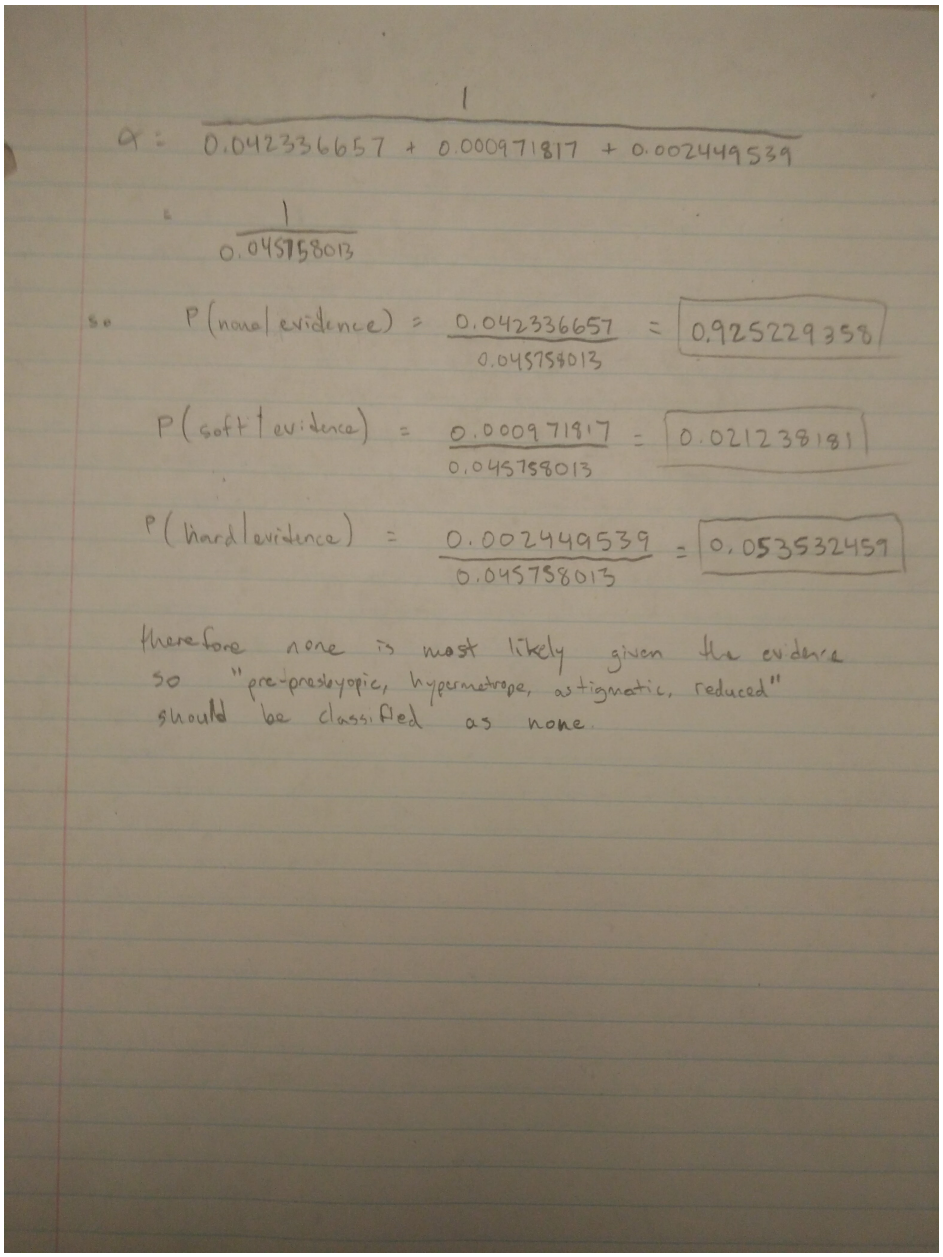
Michael Reiter
V00831568

In my tree, the root level entropy was 0.775. In the scikit-learn generated tree, the root level entropy was 0.7809. These trees differ because scikit-learn uses a different algorithm to generate it than I did. I used ID3, but it used CART. Essentially, these algorithms differ in the way they split. CART tries to balance the tree (and ensure it's binary), while ID3 does not.

2.



Michael Reiter
V00831568



$$\alpha = \frac{1}{0.042336657 + 0.000971817 + 0.002449539}$$

$$= \frac{1}{0.045758013}$$

so $P(\text{none} \mid \text{evidence}) = \frac{0.042336657}{0.045758013} = \boxed{0.925229358}$

$P(\text{soft} \mid \text{evidence}) = \frac{0.000971817}{0.045758013} = \boxed{0.021238181}$

$P(\text{hard} \mid \text{evidence}) = \frac{0.002449539}{0.045758013} = \boxed{0.053532459}$

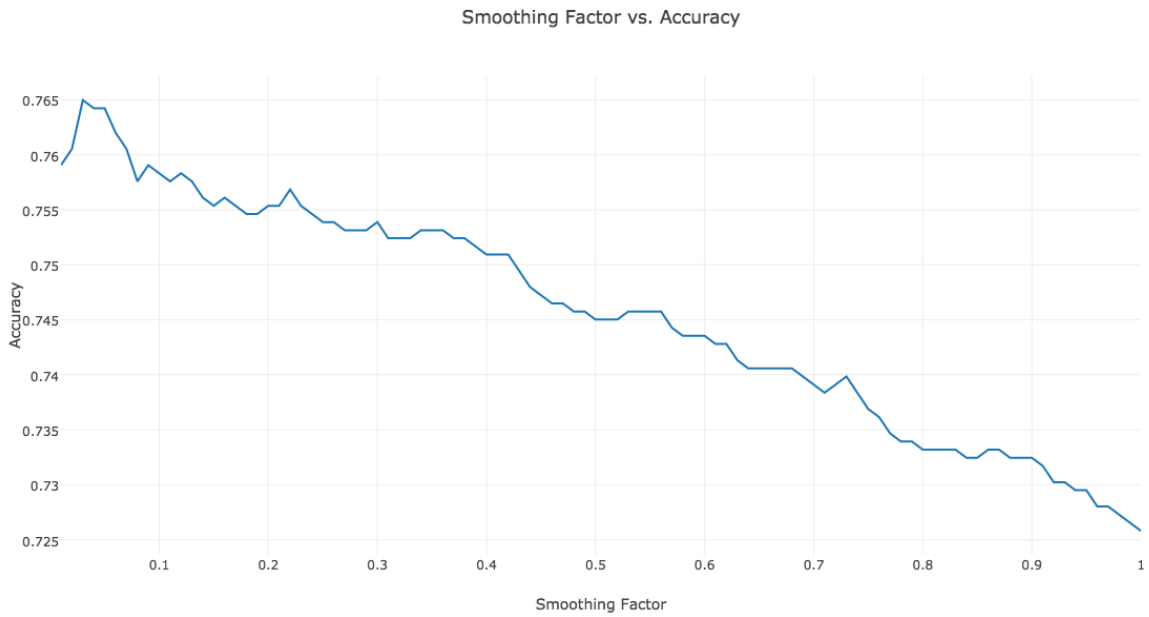
therefore none is most likely given the evidence so "pre-presbyopic, hypermetrope, astigmatic, reduced" should be classified as none.

Michael Reiter
V00831568

3.

a) See uploaded code

b)



This is a plot of smoothing factor (0.01 through 1.0 in increments of 0.01). It is clear that a smoothing factor of 0.03 is optimal. It produces a maximum accuracy of 0.765 given the fetch20_newsgroups dataset included in scikit-learn.