A.

1

	Entropy of Poisonus			Entropy of Edible	
	(-5/8 * Log5/8) + (-3/8 * Log3/8)			(-3/8 * Log3/8) + (-5/8 * Log	
0.625	-0.625		0.375	-0.375	
Log			Log		
-0.678072			-1.4150370		
0.423795			0.5306389		
0.375	-0.375		0.625	-0.625	
Log			Log		
-1.4150370			-0.678072		
0.5306389			0.423795		
0.9544339	Edible		0.9544339		

2 Gill Size

3 ?

B. Weka Exercise

1 The result of the tree is:

What is the result tree? W
?
Are you satisfied with the
or why not?

Scheme:weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: mushroom
Instances: 933
Attributes: 23

cap-shape
cap-surface
cap-color
bruises?
odor
gill-attachment
gill-spacing
gill-size
gill-color
stalk-shape
stalk-root
stalk-surface-above-ring
stalk-surface-below-ring
stalk-color-above-ring
stalk-color-below-ring
veil-type
veil-color
ring-number
ring-type
spore-print-color
population

habitat
class
Test mode:10-fold cross-validation

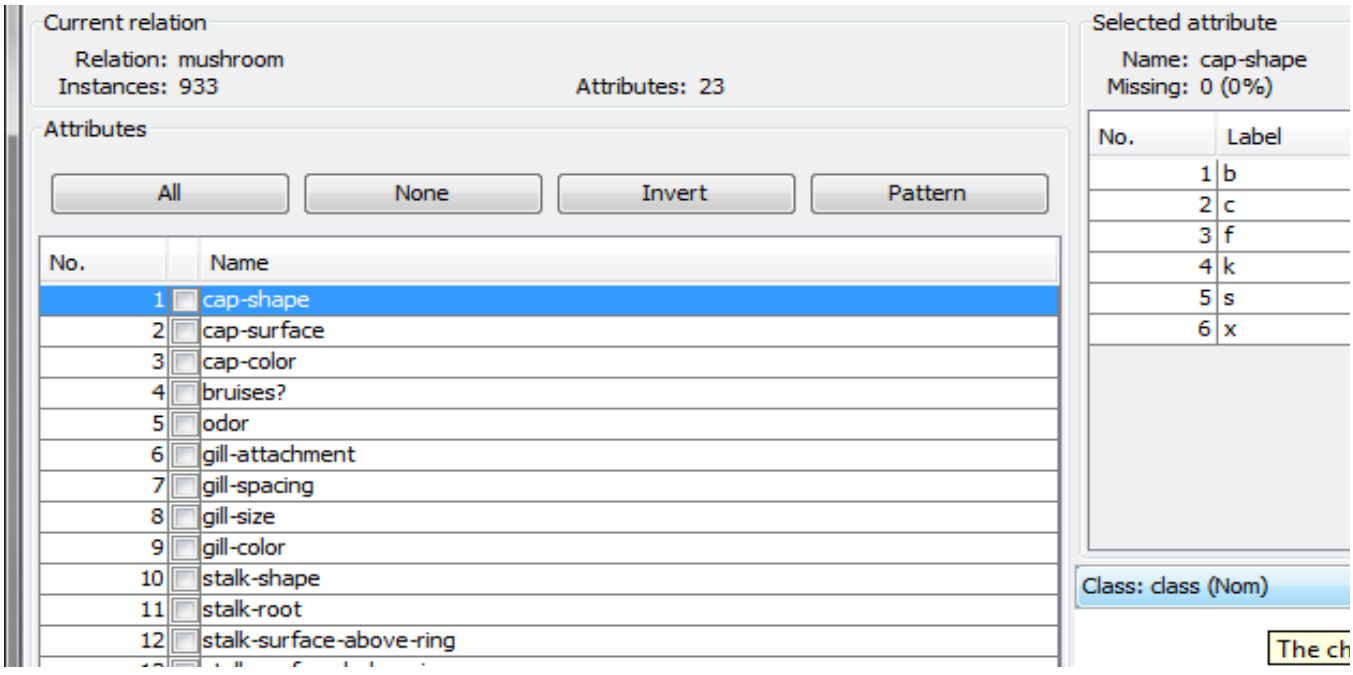
=== Classifier model (full training set) ===
J48 pruned tree
---

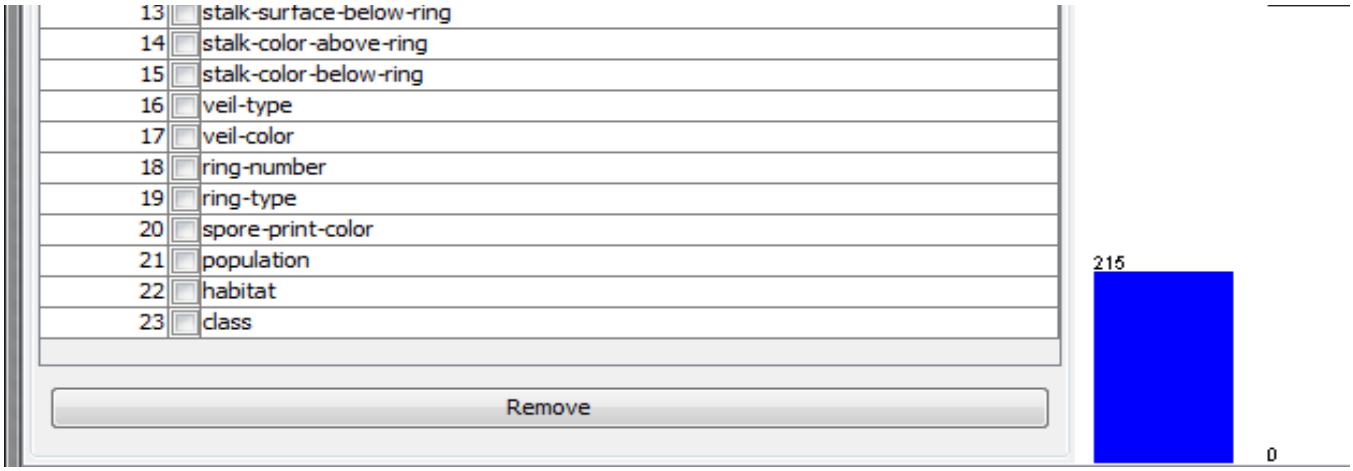
odor = a: e (332.0)
odor = c: e (0.0)
odor = f: e (0.0)
odor = l: e (323.0)
odor = m: e (0.0)
odor = n: e (183.0)
odor = p: p (95.0)
odor = s: e (0.0)
odor = y: e (0.0)

What do you think about this tree?
Are you satisfied with the result? Why
or why not?

Preliminarily, this tree looks good, but it appears to be over fit. Every possible category seems to have a node, which is why I am apprehensive to how this would perform on non-training data.

2. Remove the “odor” attribute and reconstruct the tree. What is this second tree?



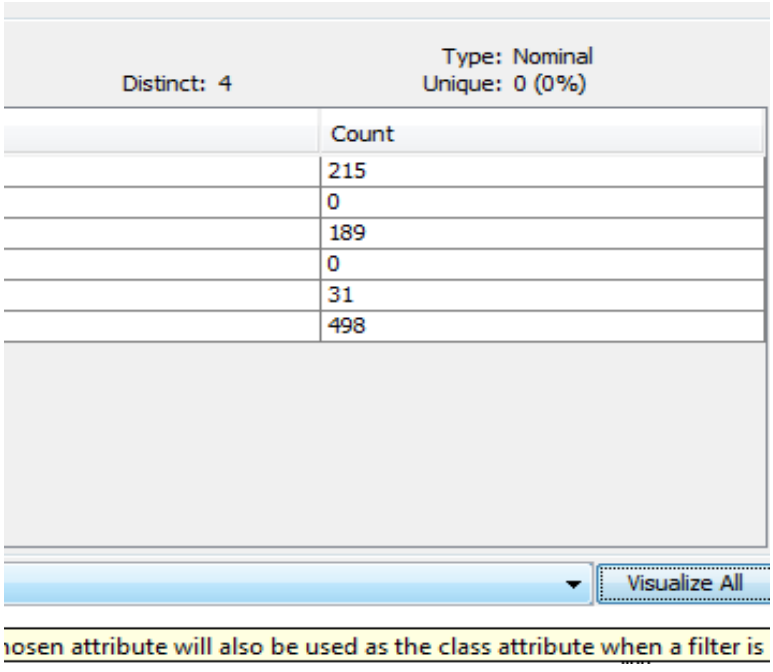


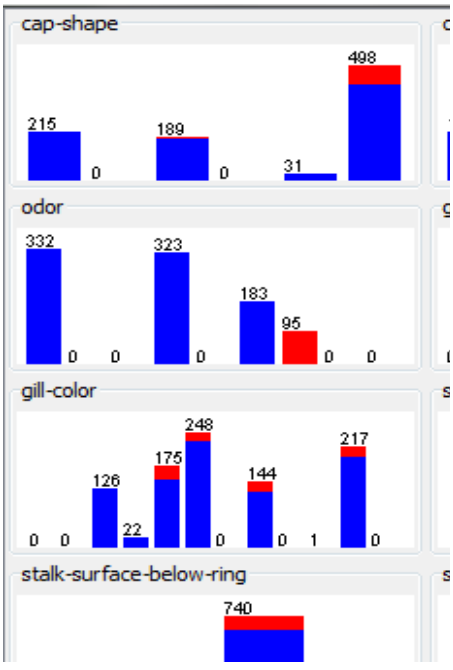
Preliminary Analysis: In this data set there are 933 examples represented in 23 attributes. Some attributes are

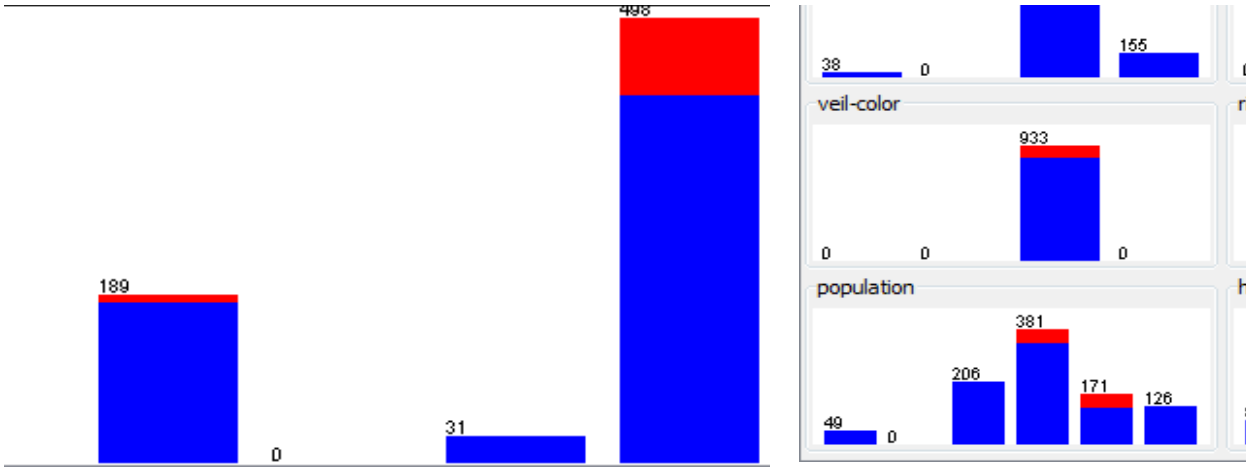
5/8)

What do you think about this tree

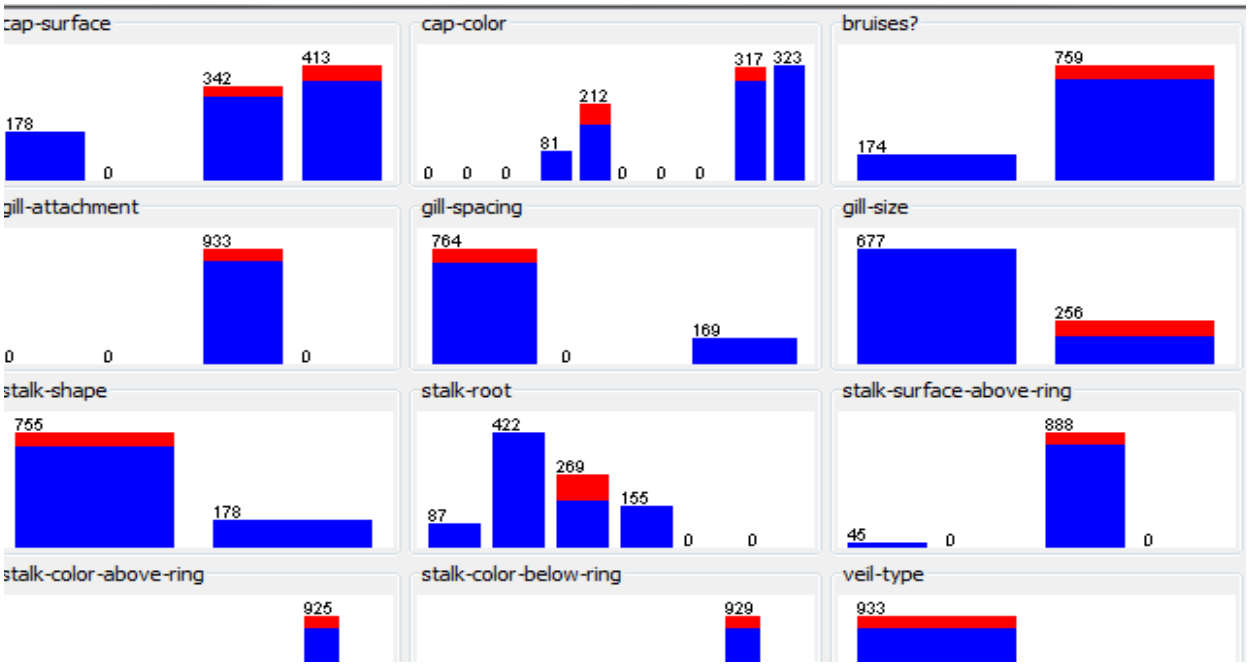
result? Why

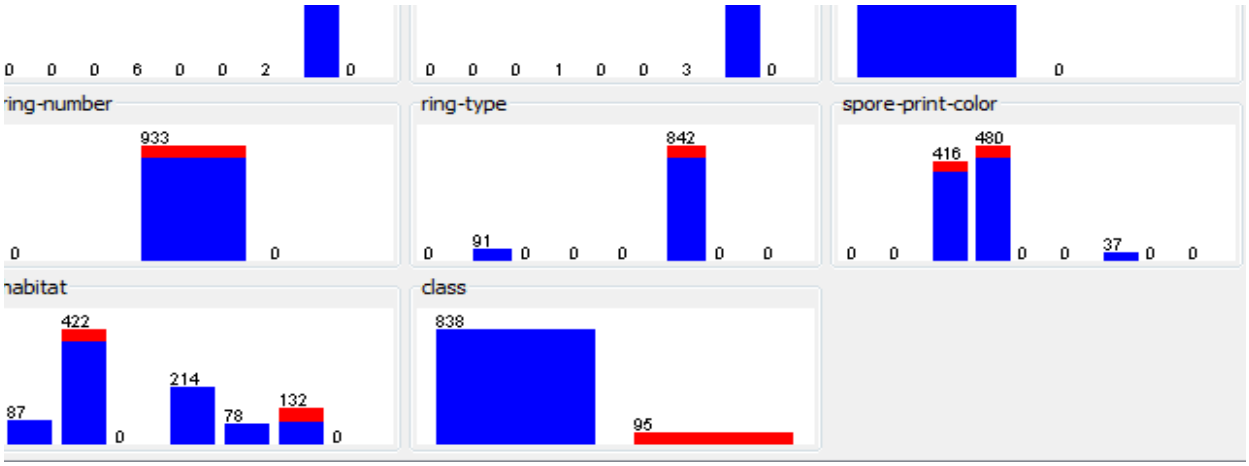

Distinct: 4
Type: Nominal
Unique: 0 (0%)
Count
215
0
189
0
31
498
Visualize All
hosen attribute will also be used as the class attribute when a filter is


cap-shape
odor
gill-color
stalk-surface-below-ring



binary, like bruises, gill-size, veil - type, stalk shape, and class, ye other attributes are categorical or con


cap-surface
178
0
342
413
cap-color
0 0 0
81
212
0 0 0
317
323
bruises?
174
759
gill-attachment
0
0
933
0
gill-spacing
764
0
169
gill-size
677
256
stalk-shape
755
178
stalk-root
87
422
269
155
0
0
stalk-surface-above-ring
45
0
888
0
stalk-color-above-ring
925
stalk-color-below-ring
929
veil-type
933



tinuous. I assume that the class attribute is the identifier of poisonous and non-poisonous. Visually, it