

#### Quiz 4

1. (1 pt.) What is the primary splitting criterion in the famous C4.5 decision tree algorithm?

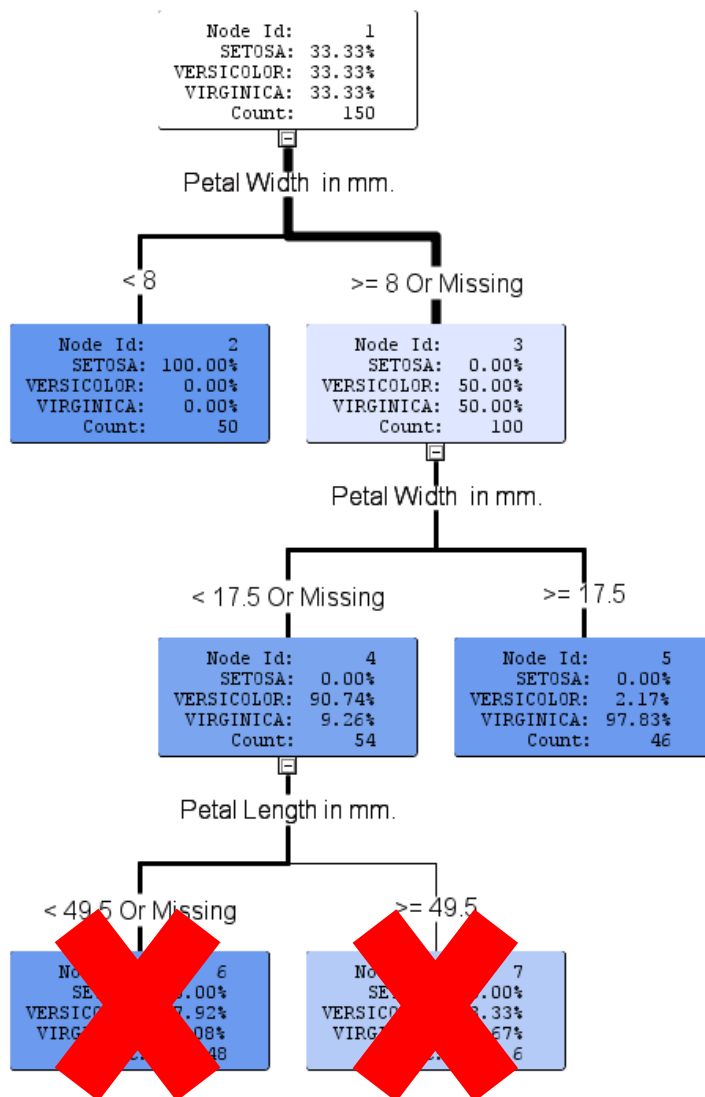
Information Gain or Entropy

2. (1 pt.) What is the primary splitting criterion in SAS Enterprise Miner?

Logworth

---

Use the trained decision tree below to answer the remaining questions.



3. (1 pt.) Write out the English language rule defined by the top, left branch of this tree.

If an input's petal width is less than 8 mm, then it is classified as Setosa.

#### Quiz 4

4. (1 pt.) What is the depth of this decision tree? (Do not include the root node in your count.)

3

5. (1 pt.) How many leaf nodes does it contain?

4

6. (1 pt.) State a tuning parameter setting that was obviously used in the training of this tree.

Maximum splits in a branch is 2.

Maximum depth is 3.

7. (1 pt.) What is the most important input variable to this tree?

Petal Width. (It accounts for the top 2 splits.)

8. (1 pt.) If a new record of data has the following attributes:

Petal Length (mm)	Petal Width (mm)	Sepal Length (mm)	Sepal Width (mm)
20	15	14	18

How would it be classified by this tree?

Versicolor or Versicolor 97.92% and Virginica 2.08%

9. (2 pts.) The presented tree is a maximal tree. Prune this tree manually given the leaf plot below. Draw an X over the nodes that should be removed.

