

Machine Learning - Lecture 10: Knowledge test

Chris Thornton

November 3, 2011

Deriving decision-trees for numeric data, there is the option of treating all numeric values as discrete, i.e., proceeding exactly as we do with categorical data. List all the problems that may then arise.

Question

Estimate the amount of information typically conveyed by a single SMS txt message.

Question

A particular body of data can provide different amounts of information in different contexts. Explain this effect.

Question

Say we decide to use the decision tree method in a particular machine learning application. List the ways in which this project might fail.

Question

Explain why maximum uncertainty should be associated with a set of equal probabilities.

Question

Imagine that a certain course has one lecture per week and you attend every lecture throughout the term. Would you expect the information you obtain from the lecture slides to increase or decrease as you go through the term. Explain why.

Question

Applying the decision tree algorithm to a certain training set, you obtain a tree with a single branch containing a huge number of cases. How would you characterize this outcome? What would you expect the generalization error of the tree to be?