

Machine Learning

Exercise Sheet 9

Prof. Dr. Dr. Lars Schmidt-Thieme, Osman Akcatepe
Information Systems and Machine Learning Lab (ISMLL)
University of Hildesheim

24. January 2012
Deadline 30. January bis 14.15

Decision Trees (10 Points)

a) (2 Points)

In a decision tree, what does an internal node denote on a single attribute? What does a branch and a leaf node represent? Give a short algorithm for a decision tree construction.

b) (2 Points)

How do we test the attributes are selected on the basis of a heuristic or statistical measure? What are the conditions for stopping partitioning?

c) (3 Points)

Extract classification rules (e.g. if-then rules) and construct a decision tree from your extraction basing on given the following training set:

<i>age</i>	<i>income</i>	<i>student</i>	<i>creditRating</i>	<i>buysComputer</i>
<=30	high	no	fair	no
<=30	high	no	excellent	no
31-40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31-40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31-40	medium	no	fair	yes
31-40	high	yes	fair	yes
>40	medium	no	excellent	no

d) (3 Points)

Find the information gains (entropies for basis 2) for the classes: age, income and creditRating.