

Maschinelles Lernen

Übungsblatt 9

Prof. Dr. Dr. Lars Schmidt-Thieme, Osman Akcatepe
Wirtschaftsinformatik und Maschinelles Lernen (ISMLL)
Universität Hildesheim

24. Januar 2012
Abgabe: 30. Januar 2012

Entscheidungsbäume (10 Punkte)

a) (2 Points)

Im einen Entscheidungsbaum, was ist die Aufgabe von einem internen Knoten auf einem einzelnen Attribut? Was repräsentieren ein Zweig und ein Blatt? Geben Sie einen kurzen Algorithmus zum Aufbau von einem Entscheidungsbaum.

b) (2 Points)

Wie überprüfen wir die Attribute, die als auf einem heuristischen oder statistischen Mass basiert ausgewählt sind? Was sind die Kriterien zum Stoppen Partitionierung.

c) (3 Points)

Extrahieren Sie die Klassifikationsregeln (z.B. if-then Regeln) und bauen Sie ein Entscheidungsbaum von Ihrem Extrahieren basierend auf folgenden Trainingsdaten auf:

<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)

Finden Sie die Informationsgewinne (Entropien der Basis 2) für die Klassen: age, income and creditRating.