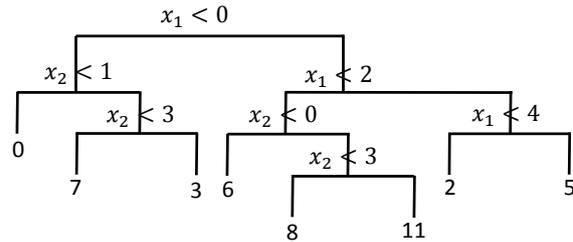
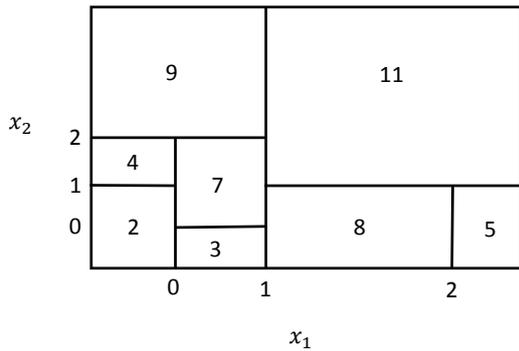


Entscheidungsbäume

Aufgabe 1

- Durch Anwendung des CART-Algorithmus kam es zu der unten links angegebenen Zerlegung des zweidimensionalen Inputraums. Skizzieren Sie den zu dieser Zerlegung zugehörigen Regressionsbaum. Zeichnen Sie dabei die Knoten, Schwellenwerte und die regionsspezifischen arithmetischen Mittel (d.h. die in den Regionen jeweils angegebenen Werte) in den Regressionsbaum ein.
- Bei der Anwendung des CART-Algorithmus wird der unten rechts angegebene Regressionsbaum erzeugt. Skizzieren und Beschriften Sie nun umgekehrt die zu diesem Baum zugehörige Zerlegung des zweidimensionalen Inputraums.



Aufgabe 2

Die folgende Tabelle gibt für 4 Kreditantragsteller an, ob sie den beantragten Kredit erhalten haben oder nicht und welche von 3 Risikomerkmale sie aufweisen.

Antragsteller	RM 1 (X_1)	RM 2 (X_2)	RM 3 (X_3)	Kredit erhalten (T)
1	Ja	Ja	Ja	Ja
2	Ja	Ja	Nein	Nein
3	Ja	Nein	Nein	Ja
4	Nein	Nein	Ja	Nein

- Konstruieren Sie einen Klassifikationsbaum. Verwenden Sie bei der Bestimmung der Partition \mathcal{P} die Fehlerrate $\varepsilon_1(R)$ und als Verlustfunktion die Fehlerfunktion

$$\frac{|R_0(i)|}{|R|} \cdot \varepsilon_1(R_0(i)) + \frac{|R_1(i)|}{|R|} \cdot \varepsilon_1(R_1(i)).$$

Zeichnen und beschriften Sie den resultierenden Klassifikationsbaum.

- Welche Prognose liefert der in Aufgabenteil a) resultierende Klassifikationsbaum für einen Kreditantragsteller, der das 1. und 3. Risikomerkmale, aber nicht das 2. Risikomerkmale aufweist?

Aufgabe 3

Die folgende Tabelle gibt für 14 Patienten an, ob sie mit einer bestimmten Krankheit infiziert sind und welche von 3 Krankheitssymptomen sie aufweisen.

- Konstruieren Sie einen Klassifikationsbaum. Verwenden Sie bei der Bestimmung der Partition \mathcal{P} die Entropie $\varepsilon_3(R)$ und als Verlustfunktion die Fehlerfunktion

$$\frac{|R_0(i)|}{|R|} \cdot \varepsilon_3(R_0(i)) + \frac{|R_1(i)|}{|R|} \cdot \varepsilon_3(R_1(i)).$$

Zeichnen und beschriften Sie den resultierenden Klassifikationsbaum.

Patient	Fieber (X_1)	Husten (X_2)	Kurzatmigkei (X_3)	infiziert (T)
1	Nein	Nein	Nein	Nein
2	Ja	Ja	Ja	Ja
3	Ja	Ja	Nein	Nein
4	Ja	Nein	Ja	Ja
5	Ja	Ja	Ja	Ja
6	Nein	Ja	Nein	Nein
7	Ja	Nein	Ja	Ja
8	Ja	Nein	Ja	Ja
9	Nein	Ja	Ja	Ja
10	Ja	Ja	Nein	Ja
11	Nein	Ja	Nein	Nein
12	Nein	Ja	Ja	Nein
13	Nein	Ja	Ja	Nein
14	Ja	Ja	Nein	Nein

- b) Geben Sie für den in Aufgabenteil a) resultierenden Klassifikationsbaum die Fehlklassifikationsrate an.
- c) Reduzieren Sie die Varianz des in Aufgabenteil a) erhaltenen Klassifikationsbaums durch Reduced error pruning. Verwenden Sie dabei die Fehl- klassifikationsrate als Pruning-Kriterium.
- d) Welche Prognose liefert der in Aufgabenteil c) resultierende Klassifikationsbaum für einen Patienten ohne Fieber und ohne Husten, aber mit Kurzatmigkei?