Eidgenössische Technische Hochschule Zürich
Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Dr. Jakob Bernasconi

Informationsverarbeitung in neuronalen Netzwerken

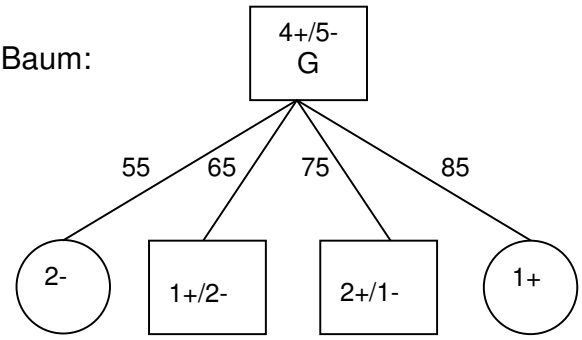
Frühjahrssemester 2009

Musterlösung Übung 1

Aufgabe 1

Man berechne die Entropieabnahmen für die beiden möglichen Bäume (1. Entscheidungsschritt). Jede neue Entropie wird berechnet, indem man die Restentropie der neuen Knoten mit der Wahrscheinlichkeit, in diesem Knoten zu sein, multipliziert und über alle Knoten summiert:

G-Baum:

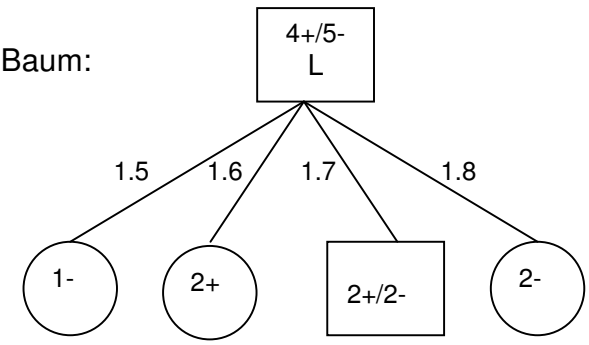


$$E_0 = -\frac{4}{9}\ln\frac{4}{9} - \frac{5}{9}\ln\frac{5}{9} = 0.687$$

$$E_1 = 2 \cdot \frac{3}{9}\left[-\frac{1}{3}\ln\frac{1}{3} - \frac{2}{3}\ln\frac{2}{3}\right] = 0.424$$

$$\Delta E_G = E_0 - E_1 = 0.263$$

L-Baum:

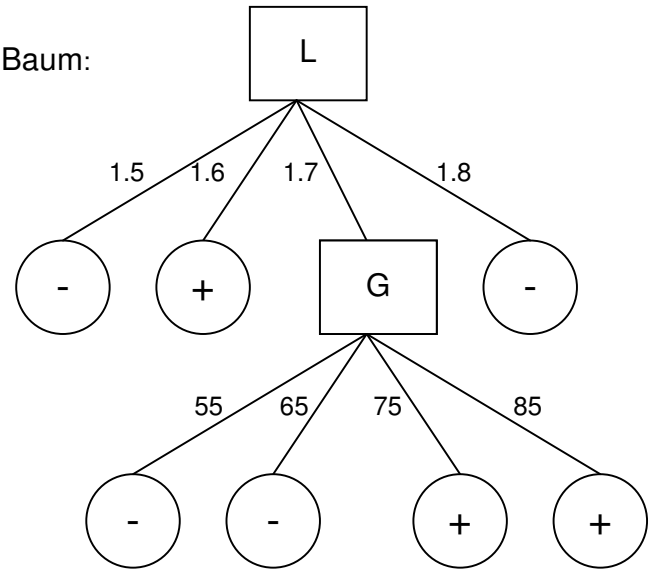


$$E_1 = \frac{4}{9}\left[-\frac{1}{2}\ln\frac{1}{2} - \frac{1}{2}\ln\frac{1}{2}\right] = 0.308$$

$$\Delta E_L = E_0 - E_1 = 0.379$$

Da $\Delta E_L > \Delta E_G$, erreicht man also die höchste Entropieabnahme, wenn man zuerst nach L unterscheidet. Der ID3-Baum verwendet also L als erstes Kriterium und hat folgende Form:

ID3-Baum:

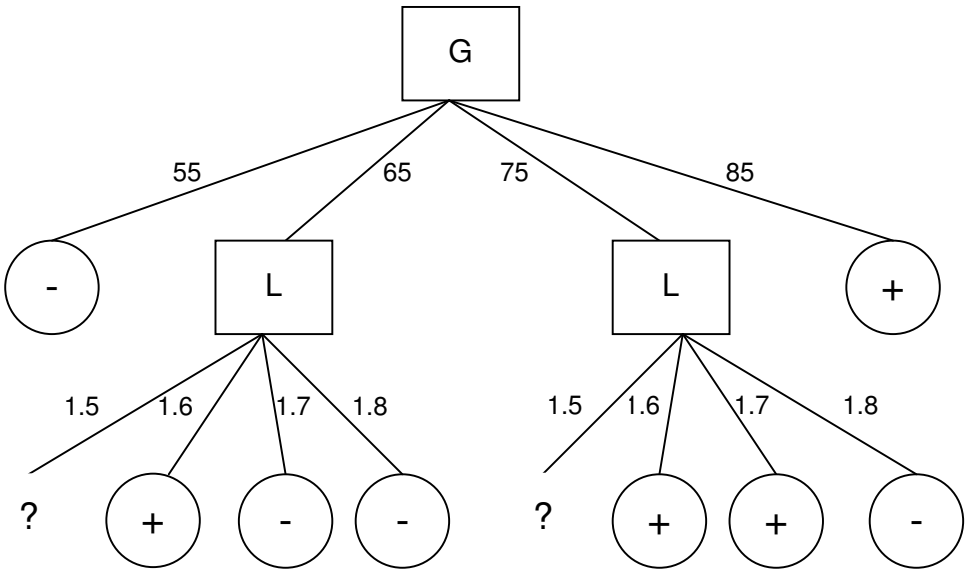


Die Regeln, die durch den ID3 Baum definiert werden, lauten:

"Übergewicht" if [L=1.6] or [L=1.7 and (G=75 or G=85)]

"kein Übergewicht" if [L=1.5] or [L=1.8] or [L=1.7 and (G=55 or G=65)]

Der alternative Entscheidungsbaum hat folgende Form und Regeln:



"Übergewicht" if [G=85] or [G=65 and L=1.6] or [G=75 and (L=1.6 or L=1.7)]

"kein Übergewicht" if [G=55] or [G=65 and (L=1.7 or L=1.8)] or [G=75 and L=1.8]

[keine Aussage für (G,L) = (75, 1.5) und (65, 1.5) !!]

Aufgabe 2

Perceptron-Entscheid:

Übergewicht falls	$W_0 + W_1 G + W_2 L \geq 0$
kein Übergewicht falls	$W_0 + W_1 G + W_2 L < 0$

Jedes Lernbeispiel ergibt eine Ungleichung für W_0, W_1 und W_2 (total 9 Ungleichungen).

Mit $W_0 = 50$ und $W_1 = 1$:

(1) $115 + 1.7\ W_2 < 0 \quad \rightarrow \quad W_2 < -67.447$

(2) $115 + 1.8\ W_2 < 0 \quad \rightarrow \quad W_2 < -63.889$

(3) $125 + 1.6\ W_2 \geq 0 \quad \rightarrow \quad W_2 \geq -78.125$

(4) $135 + 1.7\ W_2 \geq 0 \quad \rightarrow \quad W_2 \geq -79.412$

(5) $105 + 1.5\ W_2 < 0 \quad \rightarrow \quad W_2 < -70$

(6) $125 + 1.8\ W_2 < 0 \quad \rightarrow \quad W_2 < -69.444$

(7) $115 + 1.6\ W_2 \geq 0 \quad \rightarrow \quad W_2 \geq -71.875$

(8) $105 + 1.7\ W_2 < 0 \quad \rightarrow \quad W_2 < -61.765$

(9) $125 + 1.7\ W_2 \geq 0 \quad \rightarrow \quad W_2 \geq -73.529$

Es folgt, dass alle Ungleichungen erfüllt sind (d.h. alle Testresultate korrekt klassifiziert werden), falls:

$$-71.875 \leq W_2 < -70 .$$

Mit der folgenden Wahl für die Gewichte

$$W_0 = 50 , \quad W_1 = 1 , \quad W_2 = -71$$

ergibt sich dann z.B. die folgende Regel:

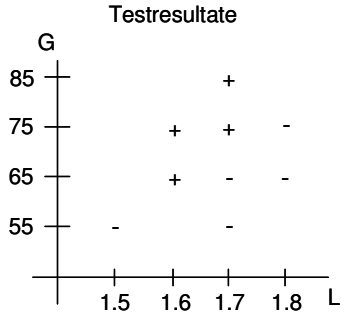
"Übergewicht" if $50 + G - 71\ L \geq 0$,

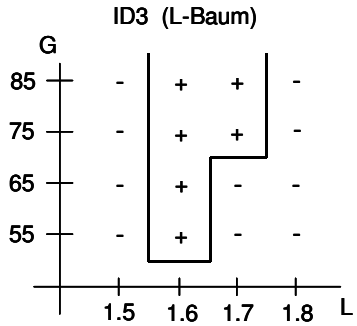
oder

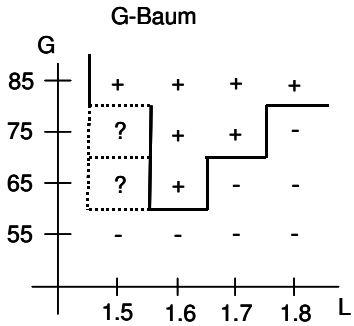
"Übergewicht" if $G \geq 71\ L - 50$.

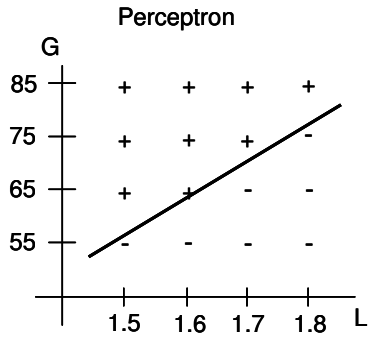
Aufgabe 3

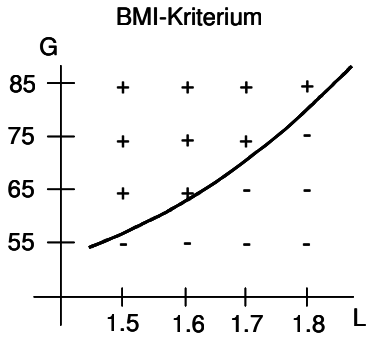
Graphische Darstellung der Lernbeispiele und der verschiedenen Regeln:











Bemerkungen:

- Das ID3-Verfahren führt zu einer “unsinnigen“ Regel, weil für L = 1.5, 1.6 und 1.8 nur Testresultate mit dem gleichen Entscheid vorliegen und diese dann durch ID3 auf alle Gewichtsklassen extrapoliert werden.
- Der G-Baum hat den Nachteil, dass er für (G, L) = (75, 1.5) und (65, 1.5) keine Aussage macht.
- Das Perceptron kann nur Geraden als Trennkurven realisieren. Das entsprechende Übergewichts-Kriterium stimmt im betrachteten Bereich aber recht gut mit dem BMI-Kriterium überein.