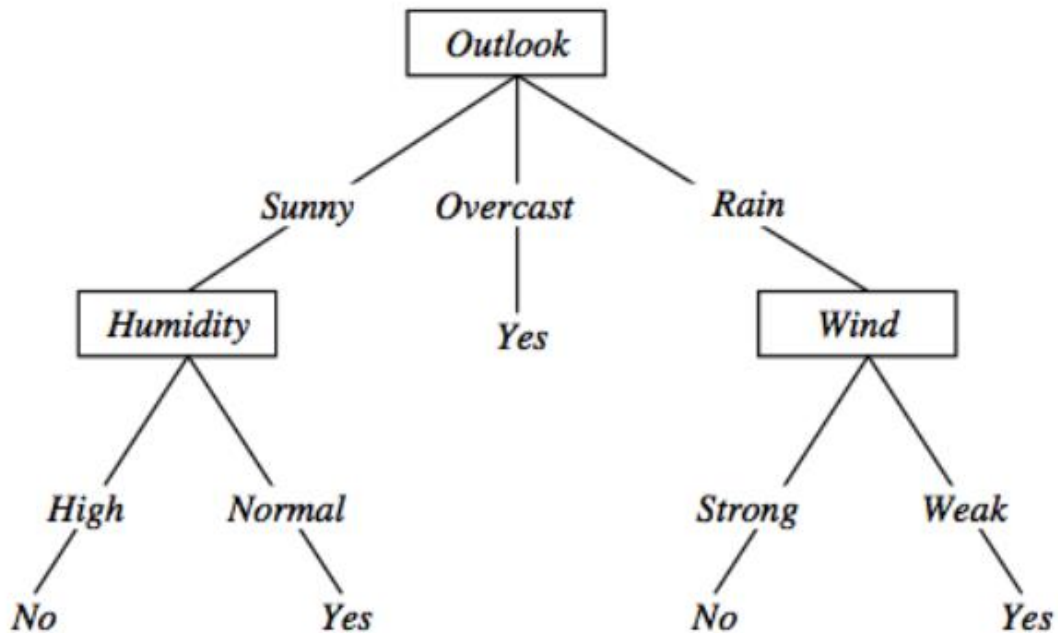


Aufgabe: Tabelle für Entscheidungsbaum.

Decision tree representation (PlayTennis)



<Outlook=Sunny, Temp=Hot, Humidity=High, Wind=Strong> No

DAY	OUTLOOK	HUMIDITY	WIND	PLAY
D1	Rain	High	Weak	Yes
D2	Sunny	High	Weak	No
D3	Sunny	High	Strong	No
D4	Overcast	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	Normal	Weak	Yes
D9	Sunny	Normal	Strong	Yes
D10	Rain	High	Strong	No
D11	Overcast	High	Strong	Yes
D12	Overcast	Normal	Weak	Yes

Aufgabe:

Erörtern Sie das Problem der Modellierung von Zeit in Entscheidungsbäumen. z.B. Michael schaltet am Montag um 8:05 das Radio ein, am Dienstag um 8:02, am Mittwoch um 7:58, etc. Was müsste man tun, um Zeit in Entscheidungsbäumen modellieren zu können.

Lösung:

Um die Modellierung der Zeit in Entscheidungsbäumen zu behandeln, verwenden wir die Technik der Diskretisierung. Die Datendiskretisierung wird als eine Methode zum Übersetzen von Attributwerten kontinuierlicher Daten in einen endlichen Satz von Intervallen mit minimalem Informationsverlust charakterisiert. Wir wählen Schnittpunkte, z.B. 7:00 Uhr, 8:00 Uhr, 9:00 Uhr, um fortlaufende Attribute zu teilen. Diese Schnittpunkte liegen im Allgemeinen in einer Teilmenge von Grenzpunkten, sodass an einem Grenzpunkt zwei benachbarte Instanzen in einer sortierten Liste unterschiedliche Zielwertattribute haben.

Aufgabe:

Gegeben ist eine Tabelle: "Geht Maria wandern?" Nutzen Sie den ID3-Algorithmus, um einen optimalen Entscheidungsbaum daraus zu generieren.

Erfassung von Daten über 14 Tage

Tag	Wetter	Pulsrate	Knie-schmerzen	Ging sie wandern
1	sonnig	hoch	schwach	nein
2	sonnig	hoch	stark	nein
3	bewölkt	hoch	schwach	ja
4	regnet	hoch	schwach	ja
5	regnet	normal	schwach	ja
6	regnet	normal	stark	nein
7	bewölkt	normal	stark	ja
8	sonnig	hoch	schwach	nein
9	sonnig	normal	schwach	ja
10	regnet	normal	schwach	ja
11	sonnig	normal	stark	ja
12	bewölkt	hoch	stark	ja
13	bewölkt	normal	schwach	ja
14	regnet	hoch	stark	nein

Lösung:

