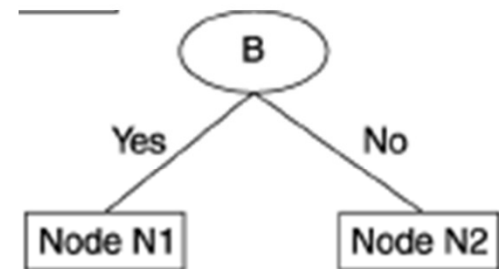


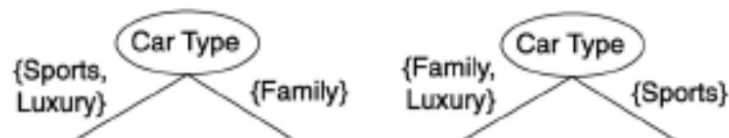
練習練習 2011.05.03

- 試計算各圖之Gini索引值運算過程

$$1 - \sum_{i=0}^{c-1} [p(i|t)]^2$$



	N1	N2
C0	1	5
C1	4	2
Gini = 0.375		



	Car Type	
	{Sports, Luxury}	{Family}
C0	9	1
C1	7	3
Gini	0.468	

	Car Type	
	{Sports}	{Family, Luxury}
C0	8	2
C1	0	10
Gini	0.167	

(a) 二元分割



	Car Type		
	Family	Sports	Luxury
C0	1	8	1
C1	3	0	7
Gini	0.163		

(b) 多元分割

練習 2011.05.03

- 請依表3.3並利用關聯法則推論出三個項目集合與四個項目集合

表 3.3 ☐ 信用卡促銷資料庫的子集合

雜誌促銷	手錶促銷	壽險促銷	信用卡保險	性別
是	否	否	否	男
是	是	是	否	女
否	否	否	否	男
是	是	是	是	男
是	否	是	否	女
否	否	否	否	女
是	否	是	是	男
否	是	否	否	男
是	否	否	否	男
是	是	是	否	女

練習 2011.05.03 (cont.)

- 試計算下列 rules 之 support 值與 Confidence 值

ID	年齡	性別	所在縣市
001	27	男	台北
002	28	女	高雄
003	33	女	台北
004	34	女	高雄
005	29	男	台北

Rule	Support	Confidence
(年齡：25~30) and (性別：男) → (所在縣市：台北)		
(所在縣市：台北) → (性別：女)		