

2016 電子商務技術 期末考

** 考試時間共 3 小時

** 若遇計算，請四捨五入取至小數點第 3 位

1. 參考下圖 weather 資料回答以下問題：

(a) 請以 equal-width binning 方式將 temperature 的值轉成 nominal 型態；(10%)

(b) 請改以 equal-frequency binning 方式轉換 humidity 的值。(10%)

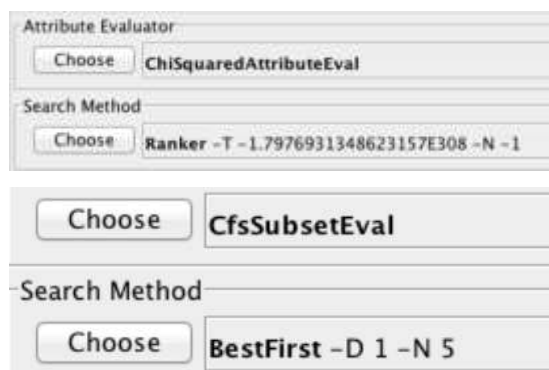
Relation: weather

No.	outlook Nominal	temperature Numeric	humidity Numeric	windy Nominal	play Nominal
1	sunny	85.0	85.0	FALSE	no
2	sunny	80.0	90.0	TRUE	no
3	overcast	83.0	86.0	FALSE	yes
4	rainy	70.0	96.0	FALSE	yes
5	rainy	68.0	80.0	FALSE	yes
6	rainy	65.0	70.0	TRUE	no
7	overcast	64.0	65.0	TRUE	yes
8	sunny	72.0	95.0	FALSE	no
9	sunny	69.0	70.0	FALSE	yes
10	rainy	75.0	80.0	FALSE	yes
11	sunny	75.0	70.0	TRUE	yes
12	overcast	72.0	90.0	TRUE	yes
13	overcast	81.0	75.0	FALSE	yes
14	rainy	71.0	91.0	TRUE	no

2. 參考下圖回答問題：

(a) CfsSubsetEval 的功能是什麼？(5%)

(b) 這兩種挑選屬性的方法的差異在那裡？(5%)



3. 某 SOM 模型有三個 input nodes、三個 output nodes A, B, C，而連線之間的權重如下，請根據此模型回答問題。

$W_{1A} = 0.83$	$W_{2A} = 0.77$	$W_{3A} = 0.81$
$W_{1B} = 0.47$	$W_{2B} = 0.23$	$W_{3B} = 0.3$
$W_{1C} = 0.61$	$W_{2C} = 0.95$	$W_{3C} = 1.34$

(a) 輸入以下 6 個 instances 後，此 SOM 模型的結果為何？(15%)

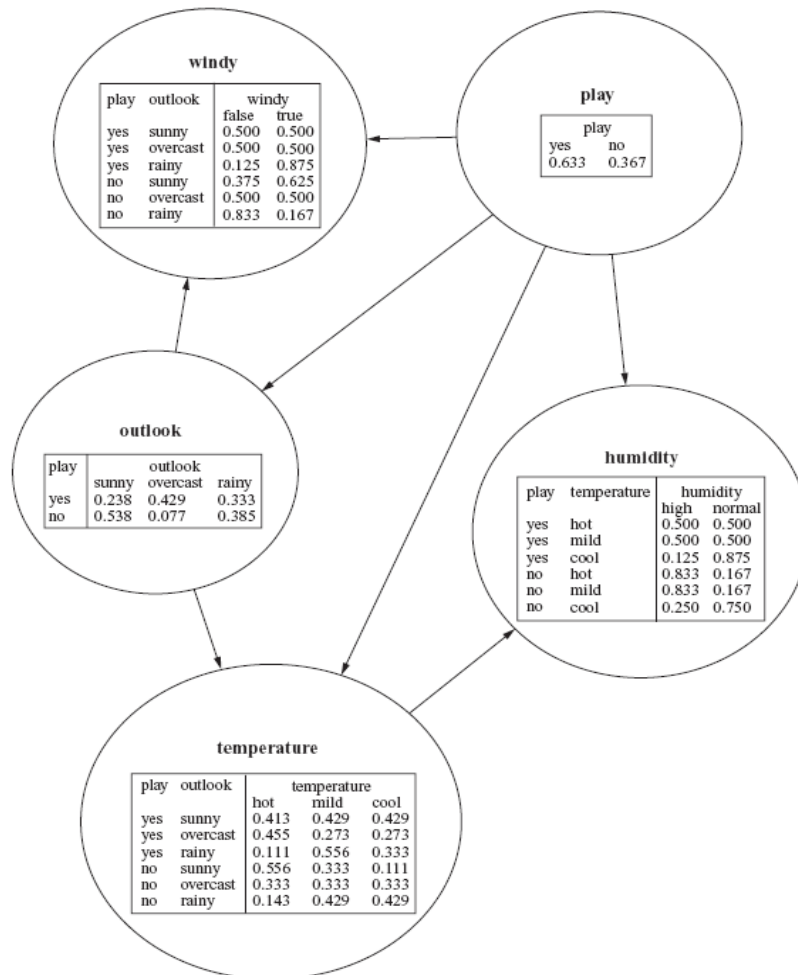
$X_1(1.1, 1.7, 1.8)$ 、 $X_2(0, 0, 0)$ 、 $X_3(0, 0.5, 1.5)$ 、 $X_4(1, 0, 0)$ 、 $X_5(0.5, 0.5, 0.5)$ 、 $X_6(1, 1, 1)$

其中： $r(t) = 0.5$

$d(t) = 1$

(b) 利用(a)的結果判斷 $X_i(1.1, 1.7, 1.8)$ 的類別。(5%)

4. 試依底下的貝氏網路圖預測範例：outlook=rainy, temperature=hot, humidity=normal, windy=false 的 play 值（需算出 play=yes 及 play=no 的機率各為多少）。(15%)



5. 利用 Agglomerative clustering 將以下七個點分成兩群，兩群間的相似度計算採 single-linkage。(15%)

	P1	P2	P3	P4	P5	P6	P7
x	0.40	0.22	0.35	0.26	0.08	0.45	0.50
y	0.53	0.38	0.32	0.19	0.41	0.30	0.50

6. 就 Adaboost 而言，假設訓練資料有 15 筆： x_0, x_1, \dots, x_{14} ，訓練第一個分類器 M1 時使用抽樣出來的 10 筆資料： x_0, x_1, \dots, x_9 ，測試時得到 2 筆錯誤： x_3, x_6 ，試問訓練下一個分類器在抽樣時所有訓練資料的權重為何。(10%)
7. 試比較 Random Forest 與 Bagging。(10%)