104.04.22

電子商務技術 期中考

*** 請按照題號順序作答，不會的題目也請寫上題號
所有的計算過程都應寫出，否則扣分

1. 右下圖是 contact-lenses 測試集使用 ID3 後產生的結果，試據此回答以下問題：

Relation: contact-lenses

No.	age Nominal	spectacle-prescrip Nominal	astigmatism Nominal	tear-prod-rate Nominal	contact-lenses Nominal
1	young	myope	no	reduced	none
2	young	myope	no	normal	soft
3	young	myope	yes	reduced	none
4	young	myope	yes	normal	hard
5	young	hypermetrope	no	reduced	none
6	young	hypermetrope	no	normal	soft
7	young	hypermetrope	yes	reduced	none
8	young	hypermetrope	yes	normal	hard
9	pre-presbyopic	myope	no	reduced	none
10	pre-presbyopic	myope	no	normal	soft
11	pre-presbyopic	myope	yes	reduced	none
12	pre-presbyopic	myope	yes	normal	hard
13	pre-presbyopic	hypermetrope	no	reduced	none
14	pre-presbyopic	hypermetrope	no	normal	soft
15	pre-presbyopic	hypermetrope	yes	reduced	none
16	pre-presbyopic	hypermetrope	yes	normal	none
17	presbyopic	myope	no	reduced	none
18	presbyopic	myope	no	normal	none
19	presbyopic	myope	yes	reduced	none
20	presbyopic	myope	yes	normal	hard
21	presbyopic	hypermetrope	no	reduced	none
22	presbyopic	hypermetrope	no	normal	soft
23	presbyopic	hypermetrope	yes	reduced	none
24	presbyopic	hypermetrope	yes	normal	none

```
tear-prod-rate = reduced: none
tear-prod-rate = normal
|  astigmatism = no
|  |  age = young: soft
|  |  age = pre-presbyopic: soft
|  |  age = presbyopic
|  |  |  spectacle-prescrip = myope: none
|  |  |  spectacle-prescrip = hypermetrope: soft
|  astigmatism = yes
|  |  spectacle-prescrip = myope: hard
|  |  spectacle-prescrip = hypermetrope
|  |  |  age = young: hard
|  |  |  age = pre-presbyopic: none
|  |  |  age = presbyopic: none
```

```
=== Predictions on training set ===

inst#,     actual, predicted, error, probability distribution
     1      3:none     3:none           0.13    0.044 *0.827
     2      1:soft     1:soft          *0.622   0.174   0.203
     3      3:none     3:none           0.018   0.186 *0.795
     4      2:hard     2:hard           0.086 *0.724   0.19
     5      3:none     3:none           0.154   0.019 *0.827
     6      1:soft     1:soft          *0.724   0.076   0.2
     7      3:none     3:none           0.024   0.092 *0.884
     8      2:hard     2:hard           0.166 *0.524   0.31
     9      3:none     3:none           0.113   0.025 *0.862
    10      1:soft     1:soft          *0.633   0.118   0.248
    11      3:none     3:none           0.017   0.113 *0.87
    12      2:hard     2:hard           0.108 *0.606   0.286
    13      3:none     3:none           0.133   0.011 *0.856
    14      1:soft     1:soft          *0.714   0.05    0.236
    15      3:none     3:none           0.021   0.054 *0.925
    16      3:none     3:none           0.187   0.394 *0.419
    17      3:none     3:none           0.068   0.023 *0.909
    18      3:none     1:soft       +  *0.509   0.142   0.349
    19      3:none     3:none           0.01    0.099 *0.891
    20      2:hard     2:hard           0.071 *0.599   0.33
    21      3:none     3:none           0.081   0.01  *0.909
    22      1:soft     1:soft          *0.594   0.062   0.344
    23      3:none     3:none           0.012   0.047 *0.941
    24      3:none     3:none           0.124   0.391 *0.485
```

(1.1) 請繪出 Decision Tree，並標註葉節點（leaf）上的 instance 個數。 (5%)

(1.2) 為何根節點（root）為 tear-prod-rate？(15%)

2. 下圖是使用 NaïveBayes 產生的預測結果，試據此回答以下問題：

(2.1) Confusion Matrix？ (5%)

(2.2) Kappa Statistic？(5%)

(2.3) True positive rate of the class “none” ？(5%)

(2.4) F-Measure of the class “hard” ？(5%)

(2.5) 參考第一題的原始資料 contact-lenses，採用 Laplace estimator 完成所有

屬性與類別屬性 contact-lenses 的 instance 個數分布表。(10%)

(2.6) 試用上表說明 instance #2 判斷為 soft 的計算過程。(10%)

3. 參考第一題的原始資料 contact-lenses，假設 support=30%, confidence=80%, 使用 Apriori Algorithm 產生關聯規則找出：

(3.1)所有 Frequent Itemsets。(10%)

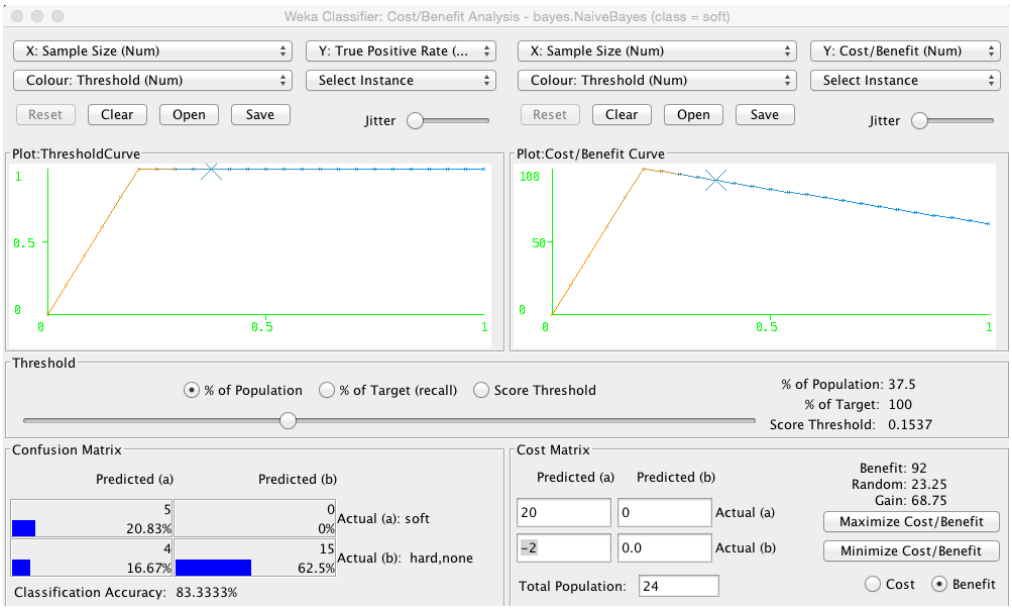
(3.2) 2 條關聯規則。(10%)

4. AND 邏輯運算表如下：

X	Y	AND	
0	0	0	C1
0	1	0	C1
1	0	0	C1
1	1	1	C2

在 Perceptron Learning Rule 模型中，自行設定 C_1 and C_2 以找出 v_0， v_1， v_2。必須寫出所有演算過程。(10%)

5. 底下是 contact-lenses 經 NaiveBayes 分類處理後的 Cost/Benefit Analysis 圖：



試問它的最佳取樣個數是多少？為什麼？(10%)