108.04.17

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

1. 下圖為資料集與利用 NaiveBayes 訓練 contact-lenses 後的預測結果，請依序回答以下問題：

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

```
=== 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) Confusion matrix? (5%)

(1.2) TP rate of class none? (5%)

(1.3) F-measure of class hard? (5%)

(1.4) Kappa statistic? (5%)

(1.5) 參考資料集，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%)

(1.6) 試用上表說明 instance #4 被預測為 hard 的計算過程。(10%)

(1.7) 試繪出 Lift chart (sample size rate vs. TP rate) of class soft。(10%)

2. 某訓練出來的預測模型於 2000 筆資料的測試後得到 80%的正確率，試問它在 99%信心度條件下的真實正確率為何？(10%)

Table 5.1 Confidence Limits for the Normal Distribution

Pr[$X \geq z$]	z
0.1%	3.09
0.5%	2.58
1%	2.33
5%	1.65
10%	1.28
20%	0.84
40%	0.25

3. 使用 complete-linkage agglomerative clustering 將以下 9 點分為三群: (15%)
$p_1(1,0)$, $p_2(1,1)$, $p_3(2,2)$, $p_4(4,1)$, $p_5(5,2)$, $p_6(6,4)$, $p_7(4,5)$, $p_8(4,4)$, $p_9(3,4)$

4. AND 邏輯運算表如下：

X	Y	AND
0	0	0
0	1	0
1	0	0
1	1	1

在 Perceptron Learning Rule 模型中，設定 AND=1 為 C_1；AND=0 為 C_2，試找出 $\boldsymbol{v}_0$，$\boldsymbol{v}_1$，$\boldsymbol{v}_2$。演算過程必須詳述。(15%)

5. 試問何謂 kernel function?為何它對於 SVM 模型那麼重要？(10%)