CS542 Machine Learning

Homework 5

Chih Wei Tung (U17550453)

8.3

$p(a,b) \neq p(a)p(b)$

But $p(a,b|c) = p(a|c)p(b|c)$ for $c = 0$ and $c = 1$

By the summation rule: $p(a,b) = \sum_{c\in\{0,1\}} p(a,b,c)$

$\Rightarrow p(a) = \sum_{b\in\{0,1\}}\sum_{c\in\{0,1\}} p(a,b,c),\ \ p(b) = \sum_{a\in\{0,1\}}\sum_{c\in\{0,1\}} p(a,b,c)$

The rule of conditional probability applying the product: $p(c,a,b) = p(a,b|c)p(c)$

$\Rightarrow p(a,b|c) = \frac{p(a,b,c)}{\sum_{a\in\{0,1\}}\sum_{b\in\{0,1\}} p(a,b,c)}$

Use similar technique:

$p(a|c) = \frac{\sum_{b\in\{0,1\}} p(a,b,c)}{\sum_{a\in\{0,1\}}\sum_{b\in\{0,1\}} p(a,b,c)}$ and $p(b|c) = \frac{\sum_{a\in\{0,1\}} p(a,b,c)}{\sum_{a\in\{0,1\}}\sum_{b\in\{0,1\}} p(a,b,c)}$

8.4

$p(a,b,c) = p(a)p(c|a)p(b|c)$

The value of $p(a)$ and $p(b|c)$ were computed as follow:

a	$p(a)$
0	0.600
1	0.400

b	c	$p(b\|c)$
0	0	0.800
0	1	0.400
1	0	0.200
1	1	0.600

Use the rule of product:

$p(c|a) = \frac{\sum_{b\in\{0,1\}} p(a,b,c)}{\sum_{b\in\{0,1\}}\sum_{c\in\{0,1\}} p(a,b,c)}$

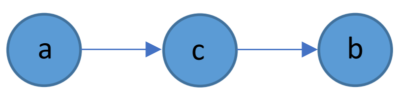
$\Rightarrow$

c	a	$p(c\|a)$
0	0	0.400
0	1	0.600
1	0	0.600
1	1	0.400

$\Rightarrow$ Multiply all the value

$\Rightarrow$ $p(a,b,c) = p(a)p(c|a)p(b|c)$

$\Rightarrow$ Corresponding directed graph:



8.11

$P(D = 1|G = 1) = 0.900$ and $P(D = 0|G = 0)$

Use Bayse Theorem and marginalizing:

$P(F = 0|D = 0) = \frac{P(D=0|F=0)P(F=0)}{P(D=0)}$

$P(D = 0|F = 0) = \sum_{B,G} P(D = 0|G)P(G|B, F = 0)P(B)$

$= 0.748$ (74.8%)

$P(D = 0) = \sum_{B,G,F} P(D = 0|G)P(G|B, F)P(B)P(F)$

$= 0.352$ (35.2%)

$\Rightarrow P(F = 0|D = 0) = 0.213$ (21.3%)

$\Rightarrow P(F = 0|D = 0, B = 0) = 0.11$ (11%)

8.14

The energy function is $E(x, y) = h \sum_i x_i - \beta \sum_{ij} x_i x_j - v \sum_i x_i y_i$ where $n, h, \beta \geq 0$ and $x_i, y_i \in \{-1, 1\}$

Set $h = \beta = 0 \Rightarrow E(x, y) = -N \sum_i x_i y_i$

When the energy is the lowest, the probable configuration will be the most.

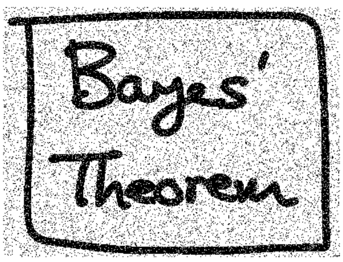
When the negative sign in form of N stays, can only happen if x_i and y_i are both 1 or both negative.

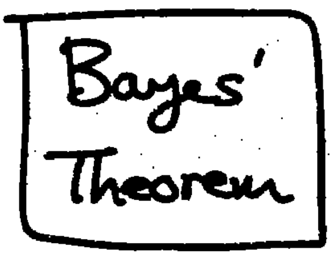
$\Rightarrow x_i = y_i$ for all i

Programming

a. First change the image to gray scale, then do the calculation.

Before: After:





When $h = -.01$, $\beta = 5$ and $\eta = 3$, the accuracy I got is 99.38%

b. For this part, I incremented the pixel, decremented the pixel, or left the pixel the same. And from the pictures below, it's obvious that the picture of the after one is much more better.

Before:

After: