Hw2
qh2174

1. $y_0 = \arg\max_y P(y_0=y|\pi)\prod_{d=1}^{2} P_d(x_{0,d}|\theta_y^{(d)})$

$\Rightarrow \hat{\pi}, \hat{\theta}_y^{(1)}, \hat{\theta}_y^{(2)} = \arg\max_{\pi,\theta_y^{(1)},\theta_y^{(2)}} \sum_{i=1}^{n} \ln p(y_i|\pi) + \sum_{i=1}^{n} \ln p(x_{i,1}|\theta_{y_i}^{(1)}) + \sum_{i=1}^{n} \ln P(x_{i,2}|\theta_{y_i}^{(2)})$

and $\because P(y|\pi) \sim Bernoulli(y|\pi)$

$\therefore P(y_i|\pi) = \pi^{y_i}(1-\pi)^{(1-y_i)}$

$\because P_1 \sim Bernoulli(x_{0,1}|\theta_y^{(1)})$

$\therefore P(x_{i,1}|\theta_{y_i}^{(1)}) = (\theta_y^{(1)})^{x_{0,1}}(1-\theta_y^{(1)})^{1-x_{0,1}}$

$\because P_2 \sim Pareto(x_{0,2}|\theta_y^{(2)})$

$\therefore P(x_{i,2}|\theta_{y_i}^{(2)}) = \theta_y^{(2)}(x_{0,2})^{-(\theta_y^{(2)}+1)}$

Let $L = \sum_{i=1}^{n}[y_i\ln\pi + (1-y_i)\ln(1-\pi)] + \sum_{i=1}^{n}[x_{i,1}\ln\theta_y^{(1)} + (1-x_{i,1})\ln(1-\theta_y^{(1)})] + \sum_{i=1}^{n}[-(\theta_y^{(2)}+1)\ln x_{i,2} + \ln\theta_y^{(2)}]$

(a) $\nabla_\pi L = \sum_{i=1}^{n}\left[\frac{y_i}{\pi} - \frac{1-y_i}{1-\pi}\right]$

let $\nabla_\pi L = 0$, then $\sum_{i=1}^{n}\left[\frac{y_i}{\pi} - \frac{1-y_i}{1-\pi}\right] = 0$

$\Rightarrow \sum_{i=1}^{n}\{(1-\pi)y_i - \pi(1-y_i)\} = 0 \Rightarrow \pi = \frac{1}{n}\sum_{i=1}^{n} y_i$

$\therefore \hat{\pi} = \frac{1}{n}\sum_{i=1}^{n} y_i$

(b) $\nabla_{\theta_y^{(1)}} L = \sum_{i=1}^{n_y}\left[\frac{x_{i,1}}{\theta_y^{(1)}} + \frac{1-x_{i,1}}{1-\theta_y^{(1)}}\times(-1)\right]\mathbb{1}(y_i=y)$

let $\nabla_{\theta_y^{(1)}} L = 0$, then $\sum_{i=1}^{n_y}\left[\frac{x_{i,1}}{\theta_y^{(1)}} - \frac{1-x_{i,1}}{1-\theta_y^{(1)}}\right]\mathbb{1}(y_i=y) = 0$

$\Rightarrow \sum_{i=1}^{n_y}[(1-\theta_y^{(1)})x_{i,1} - (1-x_{i,1})\theta_y^{(1)}]\mathbb{1}(y_i=y) = 0 \Rightarrow \theta_y^{(1)} = \frac{1}{n_y}\sum_{i=1}^{n_y}\mathbb{1}(y_i=y)x_{i,1}$

$\therefore \hat{\theta}_y^{(1)} = \frac{1}{n_y}\sum_{i=1}^{n_y}\mathbb{1}(y_i=y)x_{i,1}$ where n_y is the number of $y_i = y$ $(y=0,1)$

(c) $\nabla_{\theta_y^{(2)}} L = \sum_{i=1}^{n_y}\left[\mathbb{1}(y_i=y)\cdot\left(\frac{1}{\theta_y^{(2)}} - \ln x_{i,2}\right)\right]$

Let $\nabla_{\theta_y^{(2)}} L = 0$ then $\theta_y^{(2)} = \frac{n_y}{\sum_{i=1}^{n_y}\mathbb{1}(y_i=y)\ln(x_{i,2})}$

$\therefore \hat{\theta}_y^{(2)} = \frac{n_y}{\sum_{i=1}^{n_y}\mathbb{1}(y_i=y)\ln(x_{i,2})}$

2.

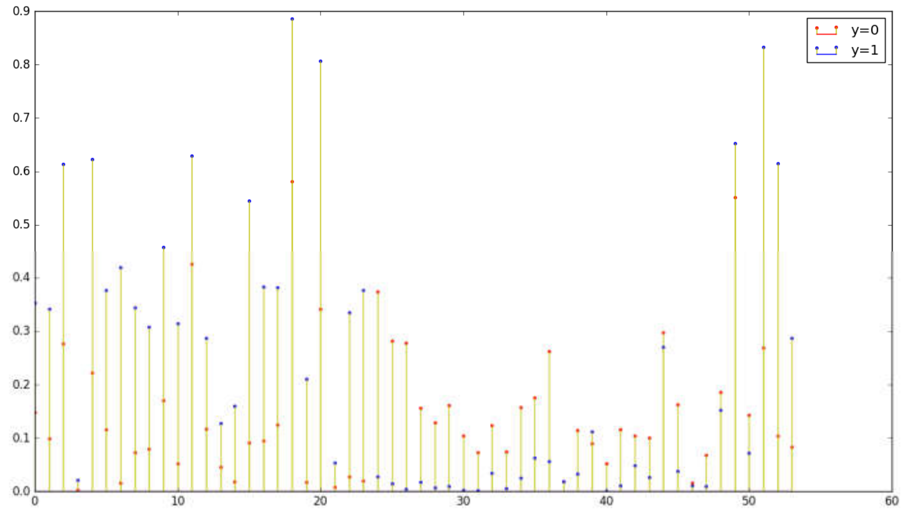
(a)

```
+---------------+------------+------------+
|       /       | actual y=0 | actual y=1 |
```

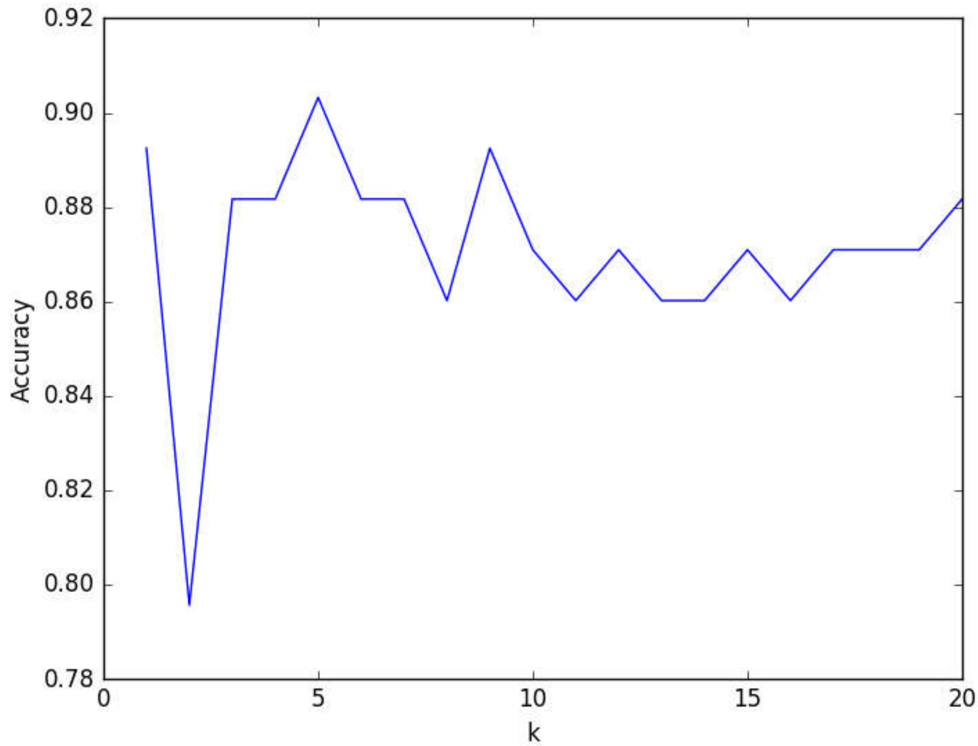
predicted y=0	54	6
predicted y=1	2	31

0.913978494624

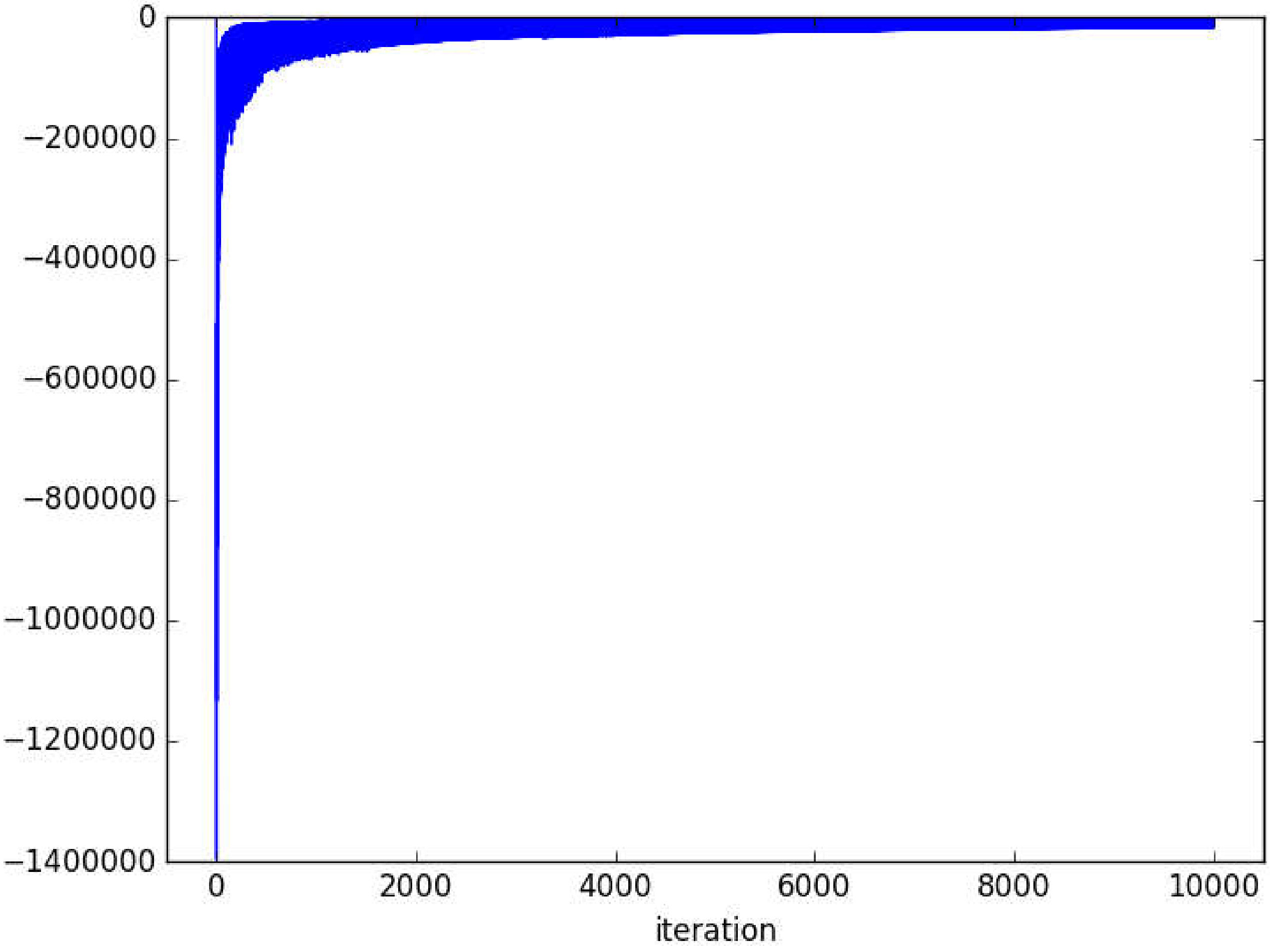
(b)



(c)



(d)



0
−200000
−400000
−600000
−800000
−1000000
−1200000
−1400000
0
2000
4000
6000
8000
10000
iteration