1) Logistic Regression

$$\ell(w) = \sum_{i=1}^{N} y^{(i)} \log h(x^{(i)}) + (1-y^{(i)}) \log(1-h(x^{(i)}))$$

where $h(x) = \sigma(w^T x) = \frac{1}{1+\exp(-w^T x)}$

$$1-h(x) = \frac{\exp(-w^T x)}{1+\exp(-w^T x)}$$

(a) Hessian of $\ell(w)$

$\nabla_w h(x) = h(x)(1-h(x))x$

$\Rightarrow$ First calculate the gradient,

$$\nabla_w \ell(w) = \sum_{i=1}^{N} y^{(i)} \nabla_w \log h(x^{(i)}) + (1-y^{(i)}) \nabla_w \log(1-h(x^{(i)}))$$

$$= \sum_{i=1}^{N} y^{(i)} \frac{\nabla_w h(x^{(i)})}{h(x^{(i)})} + (1-y^{(i)}) \frac{-\nabla_w h(x^{(i)})}{1-h(x^{(i)})}$$

$$= \sum_{i=1}^{N} y^{(i)} \frac{h(x^{(i)})(1-h(x^{(i)}))x^{(i)}}{h(x^{(i)})} - \frac{(1-y^{(i)})(h(x^{(i)})(1-h(x^{(i)}))x^{(i)})}{1-h(x^{(i)})}$$

$$= \sum_{i=1}^{N} y^{(i)}(1-h(x^{(i)}))x^{(i)} - (1-y^{(i)})h(x^{(i)})x^{(i)}$$

$$= \sum_{i=1}^{N} \left[y^{(i)} - y^{(i)}h(x^{(i)}) - h(x^{(i)}) + y^{(i)}h(x^{(i)}) \right] x^{(i)}$$

$$\nabla_w \ell(w) = \sum_{i=1}^{N} \left[y^{(i)} - h(x^{(i)}) \right] x^{(i)}$$

Now, calculating the Hessian

$$\nabla_w^2 \ell(w) = \sum_{i=1}^{n} \left[-\nabla_w h(x^{(i)}) \right] x^{(i)}$$

$$\nabla_w^2 \ell(w) = -\sum_{i=1}^{n} h(x^{(i)}) \left(1 - h(x^{(i)})\right) x^{(i)} \left(x^{(i)}\right)^T$$

(b) show $z^T H z \leq 0$ | let $x^{(i)} \in \mathbb{R}^n$

$$H = -\sum_{i=1}^{n} h(x^{(i)}) \left(1 - h(x^{(i)})\right) x^{(i)} \left(x^{(i)}\right)^T$$

* $h(x^{(i)}) = \frac{1}{1 + \exp(-w^T x^{(i)})}$, where $0 < h(x^{(i)}) < 1$

so $h(x^{(i)}) \left(1 - h(x^{(i)})\right) > 0$

For any vector z, $z \in \mathbb{R}^{n \times 1}$ | $H \in \mathbb{R}^{n \times n}$

$$z^T H z = -\sum_{i=1}^{n} h(x^{(i)}) \left(1 - h(x^{(i)})\right) z^T x^{(i)} x^{(i)T} z$$

$$= -\sum_{i=1}^{n} h(x^{(i)}) \left(1 - h(x^{(i)})\right) \underbrace{z^T x^{(i)}}_{1 \times n \;\; n \times 1 \;\; 1 \times 1} \underbrace{\left(z^T x^{(i)}\right)^T}_{1 \times 1}$$

$$= \sum_{i=1}^{n} h(x^{(i)}) \left(1 - h(x^{(i)})\right) \underbrace{\left(z^T x^{(i)}\right)^2}_{\geq 0}$$

> 0

Hence, $z^T H z \leq 0$, hence H is semi-definite $\Rightarrow \ell$ is concave and has no local minima other than global one.

(e) update rule implied by Newton's method

Upade rule:

$$w \leftarrow w - H^{-1} \nabla_w \ell(w)$$

where

$$H = \left[-\sum_{i=1}^{n} h(x^{(i)}) \left(1 - h(x^{(i)})\right) x^{(i)} (x^{(i)})^T \right]$$

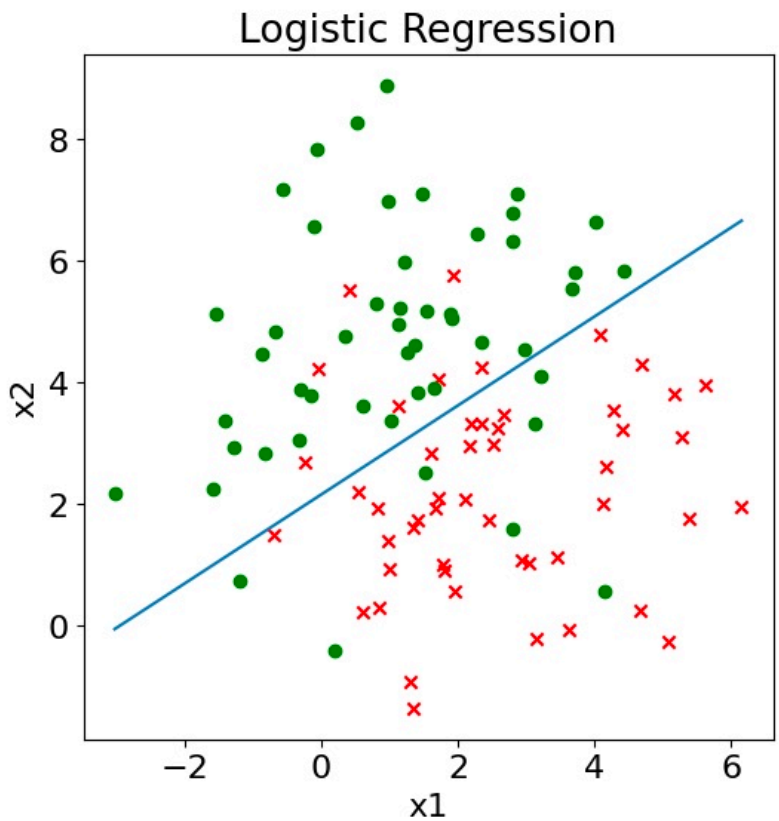
$$\nabla_w \ell(w) = \sum_{i=1}^{n} \left[y^{(i)} - h(x^{(i)}) \right] x^{(i)}$$

(d) Autograder

(e) $w =$ [-1.84922892, -0.62814188, 0.85846843]

Difference $= 1.6653345 \times 10^{-15}$

(f)



(2) Softmax Regression via Gradient Ascent.

(a)

$$\ell(\omega) = \log(L(\omega)) = \sum_{i=1}^{N} \sum_{k=1}^{K} \log\left(\left[p(y^{(i)} = k | x^{(i)}, \omega) \right]^{\mathbb{I}\{y^{(i)}=k\}} \right)$$

Case 1:

Considering $y^{(i)} = k = m$,

$$\nabla_{\omega_m} \ell(\omega) = \sum_{i=1}^{N} \mathbb{I}\{y^{(i)}=m\} \nabla_{\omega_m} \log\left(p(y^{(i)}=m | x^{(i)}, \omega) \right)$$

$$= \sum_{i=1}^{N} \mathbb{I}\{y^{(i)}=m\} \nabla_{\omega_m} \left(\log\left(\frac{\exp(\omega_m^T \phi(x^{(i)}))}{1 + \sum_{j=1}^{K-1} \exp(\omega_j^T \phi(x^{(i)}))} \right) \right)$$

$$= \sum_{i=1}^{N} \mathbb{I}\{y^{(i)}=m\} \frac{1 + \sum_{j=1}^{K-1} \exp(\omega_j^T \phi(x^{(i)}))}{\exp(\omega_m^T \phi(x^{(i)}))} \times \left(\frac{\phi(x^{(i)}) \exp(\omega_m^T \phi(x^{(i)})) \left(1 + \sum_{j=1}^{K-1} \exp(\omega_j^T \phi(x^{(i)}))\right) - \exp(\omega_m^T \phi(x^{(i)})) \phi(x^{(i)}) \exp(\omega_m^T \phi(x^{(i)}))}{\left(1 + \sum_{j=1}^{K-1} \exp(\omega_j^T \phi(x^{(i)}))\right)^2} \right)$$

$$= \sum_{i=1}^{N} \mathbb{I}\{y^{(i)}=m\} \phi(x^{(i)}) \left(\frac{1 + \sum_{j=1}^{K-1} \exp(\omega_j^T \phi(x^{(i)})) - \exp(\omega_m^T \phi(x^{(i)}))}{1 + \sum_{j=1}^{K-1} \exp(\omega_j^T \phi(x^{(i)}))} \right)$$

$$= \sum_{i=1}^{N} \phi(x^{(i)}) \left(\mathbb{I}\{y^{(i)}=m\} - \frac{\mathbb{I}\{y^{(i)}=m\} \exp(\omega_m^T \phi(x^{(i)}))}{1 + \sum_{j=1}^{K-1} \exp(\omega_j^T \phi(x^{(i)}))} \right)$$

$$\nabla_{\omega_m} \ell(\omega) = \sum_{i=1}^{N} \phi(x^{(i)}) \left(\mathbb{I}\{y^{(i)}=m\} - \frac{\exp(\omega_m^T \phi(x^{(i)}))}{1 + \sum_{j=1}^{K-1} \exp(\omega_j^T \phi(x^{(i)}))} \right)$$

$$\nabla_{\omega_m} l(\omega) = \sum_{i=1}^{N} \phi(x^{(i)}) \left[\mathbb{1}\{y^{(i)} = m\} - p(y^{(i)} = m \mid x^{(i)}, \omega) \right]$$

Case 2: considering $y^{(i)} = k \neq m$,

$$\nabla_{\omega_k} l(\omega) = \nabla_{\omega_k} \left[\sum_{i=1}^{N} \mathbb{1}(y^{(i)} = k) \log \left(p(y^{(i)} = k \mid x^{(i)}, \omega) \right) \right]$$

$$= \nabla_{\omega_k} \left[\sum_{i=1}^{N} \mathbb{1}(y^{(i)} = k) \times \log \left(\frac{1}{1 + \sum_{j=1}^{K-1} \exp(\omega_j^T \phi(x))} \right) \right]$$

no ω_k term so, 0

$$\nabla_{\omega_k} l(\omega) = 0$$

Hence,

$$\nabla_{\omega_m} l(\omega) = \sum_{i=1}^{N} \phi(x^{(i)}) \left[\mathbb{1}\{y^{(i)} = m\} - p(y^{(i)} = m \mid x^{(i)}, \omega) \right]$$

for $m = 1, \ldots, K-1$

(b) Auto grader

(c) Accuracy on the test data is 94 %

3 Gaussian Discriminate Analysis

(a)

$$p(y=1|x) = \frac{p(x|y=1)\,p(y=1)}{p(x)}$$

$$= \frac{p(x|y=1)\,p(y=1)}{p(x|y=0)\,p(y=0) + p(x|y=1)\,p(y=1)}$$

$$= \frac{\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x^{(i)}-\mu_1)^T\Sigma^{-1}(x^{(i)}-\mu_1)\right)\phi}{\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x^{(i)}-\mu_0)^T\Sigma^{-1}(x^{(i)}-\mu_0)\right)(1-\phi) + \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x^{(i)}-\mu_1)^T\Sigma^{-1}(x^{(i)}-\mu_1)\right)(\phi)}$$

$$= \frac{\exp\left(-\frac{1}{2}(x^{(i)}-\mu_1)^T\Sigma^{-1}(x^{(i)}-\mu_1)\right)\phi}{(1-\phi)\exp\left(-\frac{1}{2}(x^{(i)}-\mu_0)^T\Sigma^{-1}(x^{(i)}-\mu_0)\right) + \phi\exp\left(-\frac{1}{2}(x^{(i)}-\mu_1)^T\Sigma^{-1}(x^{(i)}-\mu_1)\right)}$$

$$= \frac{1}{1 + \frac{(1-\phi)\exp\left(-\frac{1}{2}(x^{(i)}-\mu_0)^T\Sigma^{-1}(x^{(i)}-\mu_0)\right)}{(\phi)\exp\left(-\frac{1}{2}(x^{(i)}-\mu_1)^T\Sigma^{-1}(x^{(i)}-\mu_1)\right)}}$$

$$= \frac{1}{1 + \exp\left(\log\left(\frac{1-\phi}{\phi}\right) - \frac{1}{2}(x^{(i)}-\mu_0)^T\Sigma^{-1}(x^{(i)}-\mu_0) + \frac{1}{2}(x^{(i)}-\mu_1)^T\Sigma^{-1}(x^{(i)}-\mu_1)\right)}$$

$$= \frac{1}{1+\exp\left(\log\left(\frac{1-\phi}{\phi}\right) - \frac{1}{2}\left[-2x^{(i)T}\Sigma^{-1}\mu_0 + \mu_0^T\Sigma^{-1}\mu_0 - \mu_1\Sigma^{-1}\mu_1 + 2x^{(i)T}\Sigma^{-1}\mu_1\right]\right.}$$

$$= \frac{1}{1 + \exp\left(\log\left(\frac{\phi}{1-\phi}\right) + x^{(i)T}\Sigma^{-1}(\mu_0 - \mu_1) - \frac{1}{2}\mu_0^T\Sigma^{-1}\mu_0 + \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1\right)}$$

$$p(y=1 \mid x; \phi, \Sigma, \mu_0, \mu_1) = \frac{1}{1 + \exp(-\hat{x}^T\omega)}$$

where

$$\hat{x} = \begin{bmatrix} 1 \\ x^{(i)T} \end{bmatrix}_{(n+1 \times 1)} \quad \text{and}$$

$$\omega = \begin{bmatrix} \log\left(\frac{\phi}{1-\phi}\right) + \frac{1}{2}\mu_0^T\Sigma^{-1}\mu_0 - \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1 \\ \Sigma^{-1}(\mu_0 - \mu_1) \end{bmatrix}$$

(5)

$$\ell(\phi, \mu_0, \mu_1, \Sigma) = \log \prod_{i=1}^{m} p(x^{(i)}, y^{(i)}; \phi, \mu_0, \mu_1, \Sigma)$$

$$= \log \prod_{i=1}^{m} p(x^{(i)} \mid y^{(i)}; \phi, \mu_0, \mu_1, \Sigma) p(y^{(i)}; \phi)$$

$$= \sum_{i=1}^{m} \log p(x^{(i)} \mid y^{(i)}; \phi, \mu_0, \mu_1, \Sigma) + \log p(y^{(i)}; \phi)$$

$$= \sum_{i=1}^{n} \left[{}_{(i)}\mu^T \Sigma^{-1} {}_{(i)}\mu + {}_{(i)}x^T \Sigma^{-1} {}_{(i)}x - 2\, {}_{(i)}\mu^T \Sigma^{-1} {}_{(i)}x \right]$$

$$\nabla_{\mu} \frac{1}{2} \sum_{i=1}^{n} \left({}_{(i)}x - \mu \right)^T \Sigma^{-1} \left({}_{(i)}x - \mu \right) = \left(\Sigma^{-1} \sum_{i=1}^{n} \left(\mu - {}_{(i)}x \right) \right)^T$$

$$\nabla_{\phi} \left(\Sigma^{-1} \sum_{i=1}^{n} \left(\mu - {}_{(i)}x \right) \right)^T$$

$$\Rightarrow \quad \phi = \frac{1}{n} \sum_{i=1}^{n} {}_{(i)}x \qquad \text{Hence proved}$$

$$\text{If} \quad 0 = \nabla_{\phi} \sum_{i=1}^{n} \left(\phi^{-1} \sum_{j=1}^{2} {}_{(i)}x_j \Pi + \left(\phi^{-1} \right) \sum_{j=1}^{2} {}_{(i)}x_j \Pi \right)$$

$$\phi^2 - \left(\phi \sum_{j=1}^{2} {}_{(i)}x_j \Pi + \left(\phi^{-1} \right) \sum_{j=1}^{2} {}_{(i)}x_j \Pi \right)$$

$$\frac{\phi^{-1}}{2} - \frac{\phi^{-1}}{\sum_{j=1}^{2} {}_{(i)}x_j \Pi} + \frac{\phi}{\sum_{j=1}^{2} {}_{(i)}x_j \Pi} = 0$$

$$\frac{\phi^{-1}}{2} - \frac{\phi^{-1}}{\sum_{i=1}^{n} \left(\phi + \frac{\phi^{-1}}{{}_{(i)}x} \right)} + \frac{\phi}{\sum_{i=1}^{n} {}_{(i)}x} = 0$$

$$\nabla_{\phi} \left(\Sigma^{-1}, \sigma^{-1}, \phi \right) = \frac{1}{2} \sum_{i=1}^{n} \left({}_{(i)}x - \mu \right)^T \Sigma^{-1} \left({}_{(i)}x - \mu \right) - \frac{n}{2} \log \left| \Sigma \right| - \frac{n}{2} \log \left(2\pi \right)$$

$$\left(\phi^{-1} \right) \log \left({}_{(i)}x - 1 \right) + \left(\phi \right) \log {}_{(i)}x +$$

$$\left({}_{(i)}x - \mu \right)^T \Sigma^{-1} \left({}_{(i)}x - \mu \right) - \left(\left| \Sigma \right| \right) \log \frac{1}{2} - \left(\frac{n}{2} \right) \log \frac{1}{2\pi} = \frac{1}{2} \sum_{i=1}^{n} \left(\Sigma^{-1}, \sigma^{-1}, \phi \right)$$

$$\left(\phi^{-1} \right) \log \left({}_{(i)}x - 1 \right) + \left(\phi \right) \log {}_{(i)}x +$$

$$\left({}_{(i)}x - \mu \right)^T \Sigma^{-1} \left({}_{(i)}x - \mu \right) + \left(\left| \Sigma \right| \right) \log \frac{1}{2} + \left(\frac{n}{2} \right) \log \left(2\pi \right) = \frac{1}{2} \sum_{i=1}^{n}$$

$$\left(\left({}_{(i)}x - \mu \right)^T \Sigma^{-1} \left({}_{(i)}x - \mu \right) \right) \log + \left(\left({}_{(i)}x - \mu \right)^T \Sigma^{-1} \left({}_{(i)}x - \mu \right) \right) \log \left(\frac{1}{2\pi |\Sigma|^{\frac{1}{2}}} \right) dx \Big) \log \frac{1}{2} \sum_{i=1}^{n}$$

$$\mu_1 = \frac{\sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=1\}\, x^{(i)}}{\sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=1\}}$$

$$\nabla_{\mu_1} \ell(\phi, \mu_0, \mu_1, \Sigma) = \sum_{i=1}^{m} \nabla_{\mu_1} \left((x^{(i)} - \mu_1)^T \Sigma^{-1} (x^{(i)} - \mu_1) \right) \mathbb{1}\{y^{(i)}=1\}$$

$$= \sum_{i=1}^{m} \nabla_{\mu_1} \left[x^{(i)T} \Sigma^{-1} x^{(i)} - 2\mu_1^T \Sigma^{-1} x^{(i)} + \mu_1^T \Sigma^{-1} \mu_1 \right] \mathbb{1}\{y^{(i)}=1\}$$

$$= \sum_{i=1}^{m} \left[-2\Sigma^{-1} x^{(i)} + 2\Sigma^{-1}\mu_1 \right] \mathbb{1}\{y^{(i)}=1\}$$

$$= \left(\frac{1}{m}\sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=1\}\right) \left[-2\Sigma^{-1} \sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=1\}\, x^{(i)} + 2\Sigma^{-1}\mu_1 \sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=1\} \right]$$

$$0 = \sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=1\}\, x^{(i)} - \mu_1 \sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=1\}$$

$$\mu_1 = \frac{\sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=1\}\, x^{(i)}}{\sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=1\}}$$

Hence Proved

$$\mu_0 = \frac{\sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=0\}\, x^{(i)}}{\sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=0\}}$$

$$0 = \sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=0\}\, x^{(i)} - \mu_0 \sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=0\}$$

$$\sigma_0 \sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=0\}$$

$$= \left(\frac{1}{m}\sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=0\}\right) \left[-2\Sigma^{-1} \sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=0\}\, x^{(i)} + 2\Sigma^{-1}\mu_0 \sum_{i=1}^{m} \mathbb{1}\{y^{(i)}=0\} \right]$$

(G) For $n=1$, $\Sigma = [\sigma^2]$, $|\Sigma| = \sigma^2$

$$\ell(\phi, \mu_0, \mu_1, \Sigma) = \sum_{i=1}^{m} -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log(|\Sigma|) - \frac{1}{2}(x^{(i)} - \mu_{y^{(i)}})^T \Sigma^{-1} (x^{(i)} - \mu_{y^{(i)}}) + (1-\phi)\log(1-\phi) + \phi \log(\phi)$$

$$\nabla_{\Sigma} \ell(\phi, \mu_0, \mu_1, \Sigma) = \sum_{i=1}^{m} \nabla_{\Sigma} \left(-\frac{1}{2}\log(|\Sigma|) - \frac{1}{2}(x^{(i)} - \mu_{y^{(i)}})^T \Sigma^{-1} (x^{(i)} - \mu_{y^{(i)}}) \right)$$

$$0 = \sum_{i=1}^{m} -\frac{1}{2}\frac{1}{|\Sigma|}|\Sigma| \Sigma^{-1} - \frac{1}{2}(x^{(i)} - \mu_{y^{(i)}})(x^{(i)} - \mu_{y^{(i)}})^T$$

$$0 = \sum_{i=1}^{m} -\frac{1}{|\Sigma|}|\Sigma| \Sigma^{-1} + \frac{1}{2}\sum_{i=1}^{m} (x^{(i)} - \mu_{y^{(i)}})^T (x^{(i)} - \mu_{y^{(i)}})$$

$$\frac{m}{|\Sigma|}|\Sigma| \Sigma^{-1} = \sum_{i=1}^{m} (x^{(i)} - \mu_{y^{(i)}})^T (x^{(i)} - \mu_{y^{(i)}})$$

$$\boxed{\Sigma = \frac{1}{m}\sum_{i=1}^{m} (x^{(i)} - \mu_{y^{(i)}})^T (x^{(i)} - \mu_{y^{(i)}})}$$

Hence Proved

4) Naive Bayes for classifying SPAM

(a) Objective is to compute, $P(\mu | X, Y)$

For MAP, we want Posterior distribution $P(\mu | X, Y)$

using Bayes formula, $P(\mu | X, Y) = \frac{P(X, Y | \mu) P(\mu)}{P(X, Y)}$ — (i)

$P(X, Y | \mu)$ is a likelihood,

$$P(X, Y | \mu) = \prod_{i=1}^{K} \prod_{j=1}^{M} (\phi_i)^{N_i} (\mu_j^i)^{N_j^i} \quad \text{— (ii)}$$

$P(\mu)$ is a prior,

$$P(\mu) = \frac{1}{Z} \prod_{i=1}^{K} \prod_{j=1}^{M} (\mu_j^i)^{\alpha} \quad \text{— (iii)}$$

Taking log of eq (i) and substituting (ii) and (iii)

$$\log(P(\mu | X, Y)) = \sum_{i=1}^{K} \sum_{j=1}^{M} N_i \log(\phi_i) + N_j^i \log(\mu_j^i) + \sum_{i=1}^{K} \sum_{j=1}^{M} \alpha \log(\mu_j^i) - \log(P(X, Y)) - \log(Z)$$

$$\log\left(P(\mu|X,Y)\right) = \sum_{i=1}^{K}\sum_{j=1}^{M} N_i \log(\phi_i) + \left(N_j^i + \alpha\right)\log(\mu_j^i) - \log\left(P(X,Y)\right) - \log(Z)$$

↳ This is similar to log-likelihood derived in the lecture notes.

This can be further simplify by removing dependent terms (as done in the lecture),

$$\log\left(P(\mu|X,Y)\right) = \sum_{i=1}^{K}\sum_{j=1}^{M} N_i \log(\phi_i) + \sum_{i=1}^{K}\left[\sum_{j=1}^{M-1}\left(N_j^i+\alpha\right)\log(\mu_j^i) + \left(N_M^i+\alpha\right)\log\left(1-\sum_{j=1}^{M-1}\mu_j^i\right)\right] - \log\left(P(X,Y)\right) - \log(Z)$$

Taking partial derivative w.r.t μ_j^i, and equating to zero.

$$0 = \frac{N_j^i+\alpha}{\mu_j^i} - \frac{N_M^i+\alpha}{1-\sum_{j=1}^{M-1}\mu_j^i}$$

$$\Rightarrow \quad \frac{N_j^i+\alpha}{\mu_j^i} = \text{constant} \quad \forall\, j$$

$$\Rightarrow \quad \mu_j^i = \frac{N_j^i+\alpha}{\sum_{j=1}^{M} N_j^i + \alpha M}$$

Hence Proved

(b)

(i) Autograder

(ii) Top 5 most indicative tokens are: ['valet' 'ebai' 'unsubscrib' 'spam' 'httpaddr'].

(iii)

Accuracy for 50 mail data : 96.125%

Accuracy for 100 mail data : 97.375%

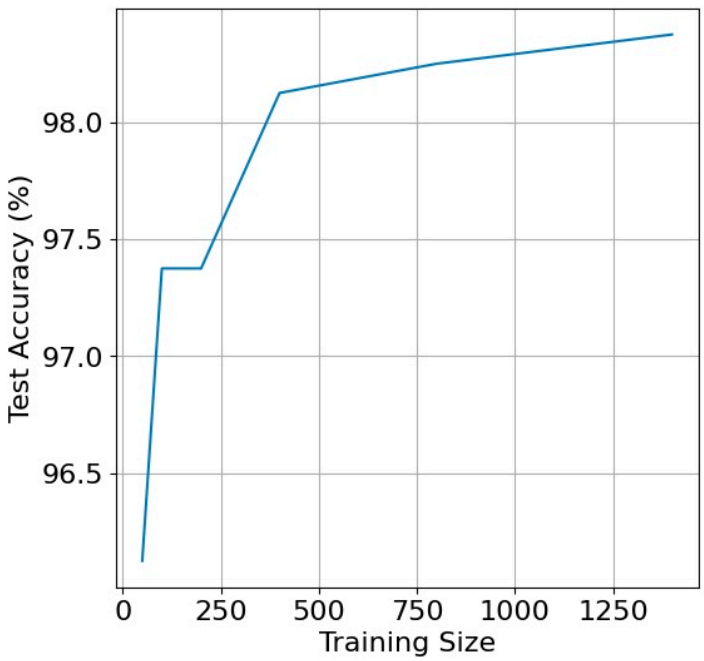
Accuracy for 200 mail data : 97.375%

Accuracy for 400 mail data : 98.125%

Accuracy for 800 mail data : 98.25%

Accuracy for 1400 mail data : 98.375%

(iv)



(v)

1400 Training set size gives the best accuracy.

Study Group Members: Eddie Aljamal, Alex Janosi, Sisir Potluri, and Shashank Verma