CS 6140: Machine Learning　　　Spring 2018

Assignment 3

1.3

1.

Spambase

Fold	Training Accuracy	Training Precision	Training Recall	Test Accuracy	Test Precision	Test Recall
1	0.92557	0.92693	0.88270	0.92399	0.92593	0.88235
2	0.92531	0.92852	0.88142	0.91686	0.89375	0.88820
3	0.92452	0.92324	0.88411	0.92162	0.92500	0.87574
4	0.92452	0.92561	0.87992	0.92399	0.92045	0.90000
5	0.92188	0.92127	0.87995	0.94537	0.93750	0.92025
6	0.92478	0.92452	0.88352	0.92874	0.93671	0.88095
7	0.92689	0.92907	0.88259	0.91924	0.91954	0.88889
8	0.93059	0.92625	0.89888	0.92874	0.90385	0.90385
9	0.93032	0.93115	0.89029	0.89549	0.88757	0.85714
10	0.92689	0.92725	0.88765	0.91686	0.94203	0.82803
Mean	0.92613	0.92638	0.88510	0.92209	0.91923	0.88254
Standard Deviation	0.00254	0.00275	0.00553	0.01187	0.01768	0.02433

Breast Cancer

Fold	Training Accuracy	Training Precision	Training Recall	Test Accuracy	Test Precision	Test Recall
1	0.98828	0.99468	0.97396	0.98246	1.00000	0.95000
2	0.99219	1.00000	0.97849	0.96491	0.96154	0.96154
3	0.99023	0.99476	0.97938	0.96491	0.94444	0.94444
4	0.98828	0.99471	0.97409	1.00000	1.00000	1.00000
5	0.98828	0.99492	0.97512	0.98246	0.91667	1.00000
6	0.98633	0.98889	0.97268	1.00000	1.00000	1.00000
7	0.99023	1.00000	0.97312	0.98246	1.00000	0.96154
8	0.98633	0.98925	0.97354	0.96491	0.95652	0.95652
9	0.98633	0.98947	0.97409	1.00000	1.00000	1.00000

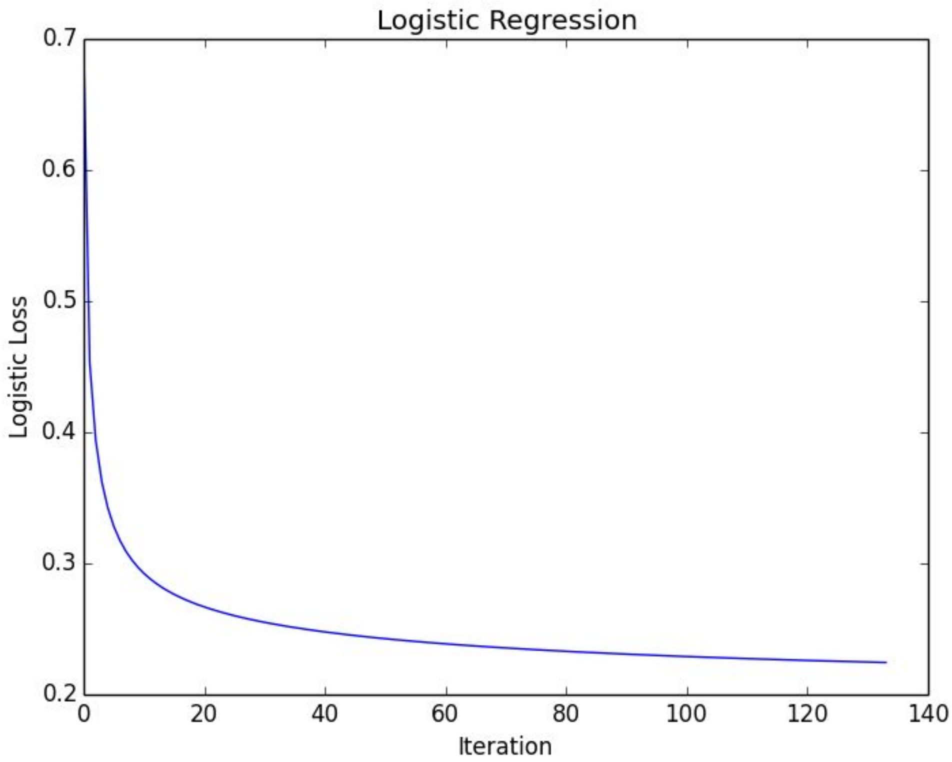
10	0.99220	0.98953	0.98953	0.92857	0.94737	0.85714
Mean	0.98887	0.99362	0.97640	0.97707	0.97265	0.96312
Standard Deviation	0.00215	0.00403	0.00486	0.02112	0.02948	0.04152

Pima Indian Diabetes

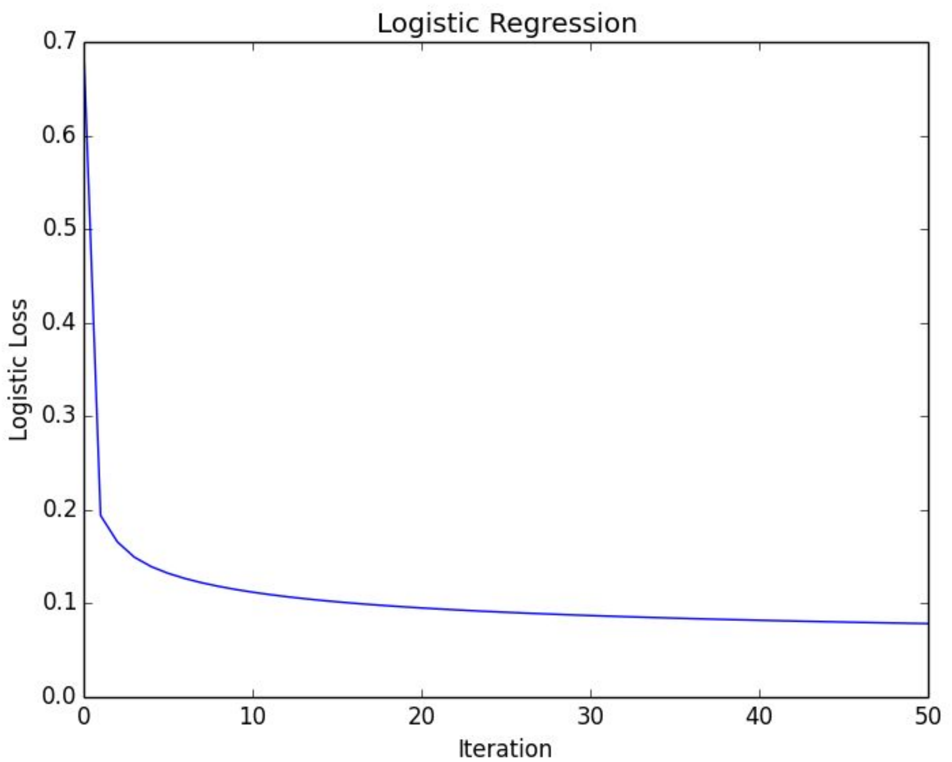
Fold	Training Accuracy	Training Precision	Training Recall	Test Accuracy	Test Precision	Test Recall
1	0.78582	0.74093	0.59336	0.72727	0.65000	0.48148
2	0.77424	0.73000	0.58871	0.83117	0.65217	0.75000
3	0.78871	0.74317	0.57872	0.74026	0.78261	0.54545
4	0.78292	0.73626	0.56780	0.71429	0.70833	0.53125
5	0.77713	0.72775	0.57676	0.81818	0.84211	0.59259
6	0.77713	0.73262	0.56846	0.84416	0.80000	0.74074
7	0.78148	0.73370	0.56962	0.75325	0.77273	0.54839
8	0.77713	0.72487	0.57322	0.81818	0.82609	0.65517
9	0.78035	0.73846	0.58776	0.72368	0.56250	0.39130
10	0.77746	0.73632	0.59438	0.80263	0.62500	0.52632
Mean	0.78024	0.73441	0.57988	0.77731	0.72215	0.57627
Standard Deviation	0.00429	0.00549	0.00984	0.04760	0.09093	0.10638

2.

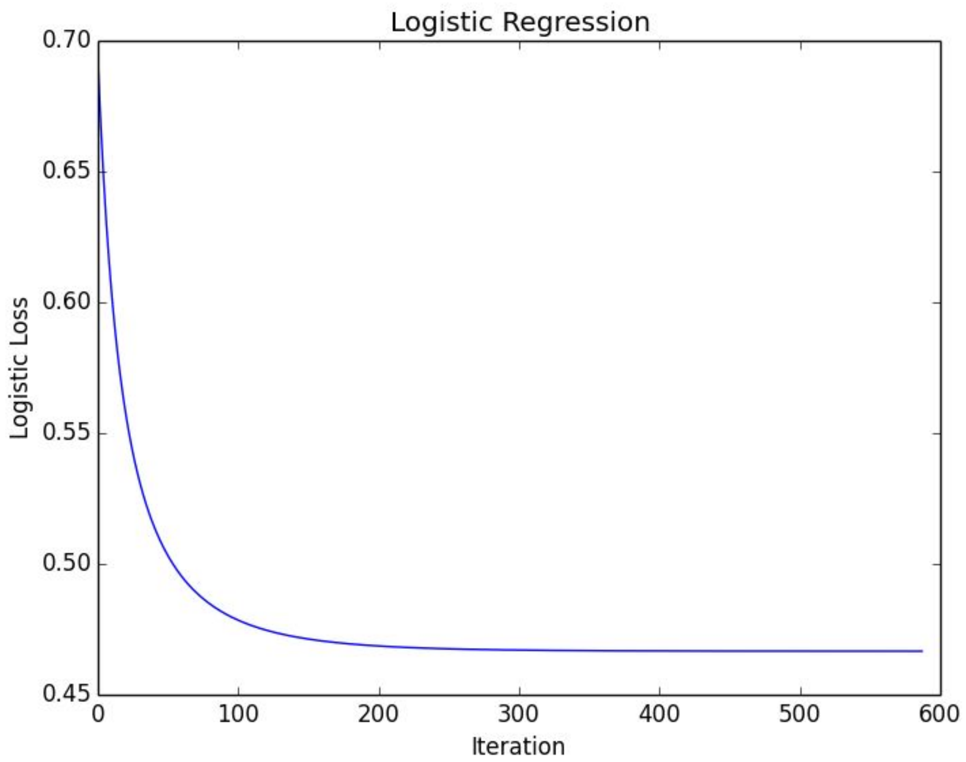
Spambase



Breast Cancer



Pima Indian Diabetes



3.

I set the initial iterations to 1000 and ran the regression with tolerance 0.1. Gradually increasing the tolerance by 10 percent checked if the mean accuracy increases. If so continued to test for more other tolerance values and if not used the previous tolerance value. I printed for average learning loops for each fold and found that the maximum loops never reach more than 750, so left it unchanged at 1000.

1.4

1.

1. $\sigma(a) = \dfrac{1}{1+e^{-a}}$

$\dfrac{d\sigma(x)}{dx}$ when $x = \omega^T x$

$$\frac{d}{dx}\left(\frac{1}{1+e^{-\omega^T x}}\right) \Rightarrow - \frac{\frac{d}{dx}\left(1+e^{-\omega^T x}\right)}{\left(1+e^{-\omega^T x}\right)^2}$$

$$\Rightarrow - \frac{\frac{d}{dx}(1) + \frac{d}{dx}\left(e^{-\omega^T x}\right)}{\left(1+e^{-\omega^T x}\right)^2} = - \frac{e^{-\omega^T x} \cdot \frac{d}{dx}\left[-\omega^T x\right]}{\left(1+e^{-\omega^T x}\right)^2}$$

$$\Rightarrow \frac{\omega^T e^{-\omega^T x}}{\left(1+e^{-\omega^T x}\right)^2}$$

2.

Posterior probability $\sigma(a) = \frac{1}{1+e^{-a}}$

$$P(y=1|x,w) = \sigma(w^T x) = \frac{1}{1+e^{-w^T x}}$$

$y \in \{-1, 1\}$

$$P(y=-1|x,w) = 1 - P(y=1|x,w)$$

$$= 1 - \frac{1}{1+e^{-w^T x}}$$

$$P(y=-1|x,w) = \frac{1}{1+e^{w^T x}}$$

$$P(y=\pm 1|x,w) = \frac{1}{1+e^{-yw^T x}}$$

3.

$y_i \in \{-1, 1\}$

$$P(y_i|x_i) = \frac{1}{1+e^{-y_i f(x_i)}}$$

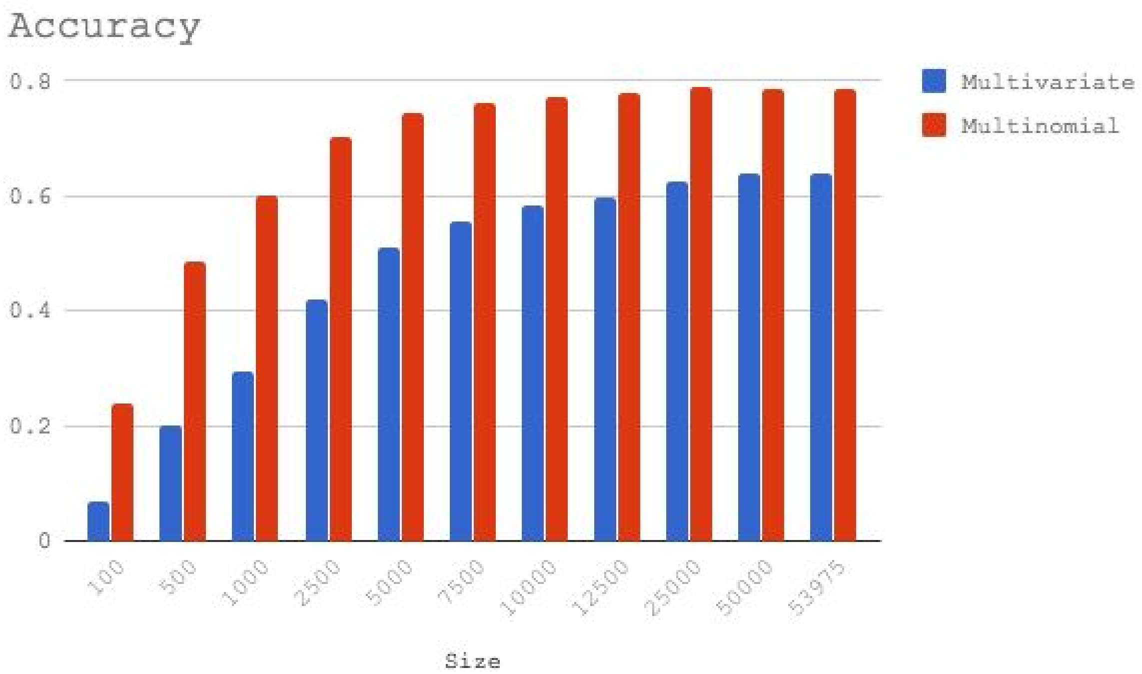
Likelihood is $\prod_{i}^{N} \frac{1}{1+e^{-y_i f(x_i)}}$

taking negative log, likelihood function is

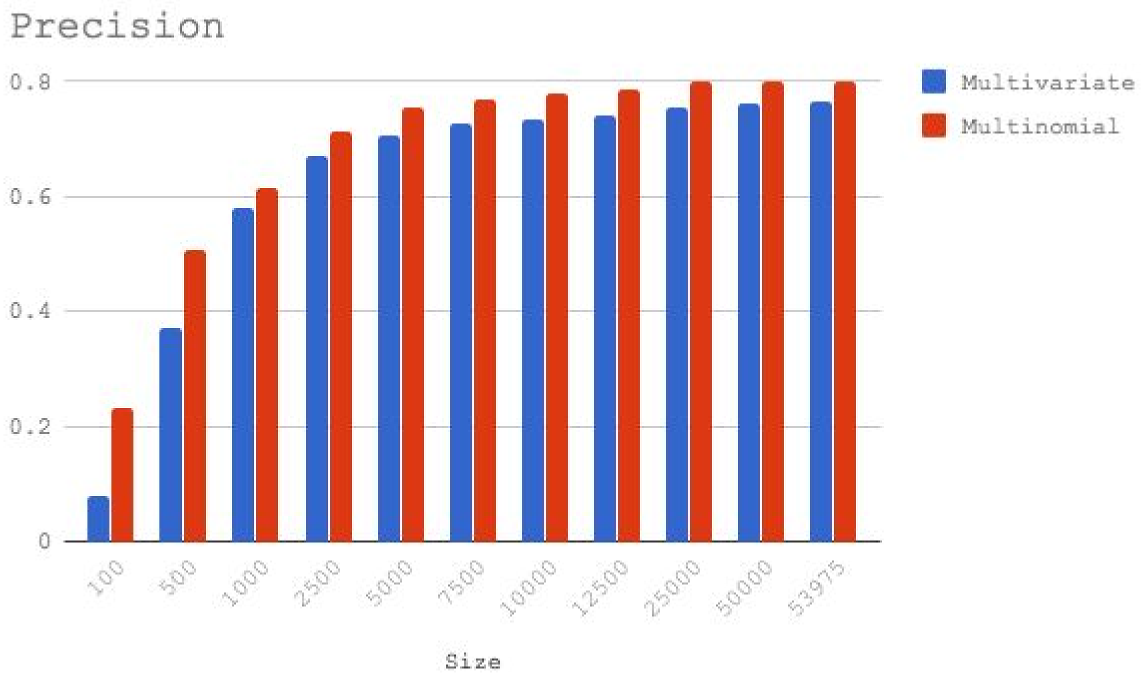
$$\sum_{i}^{N} \log(1+e^{-y_i f(x_i)})$$

2.5

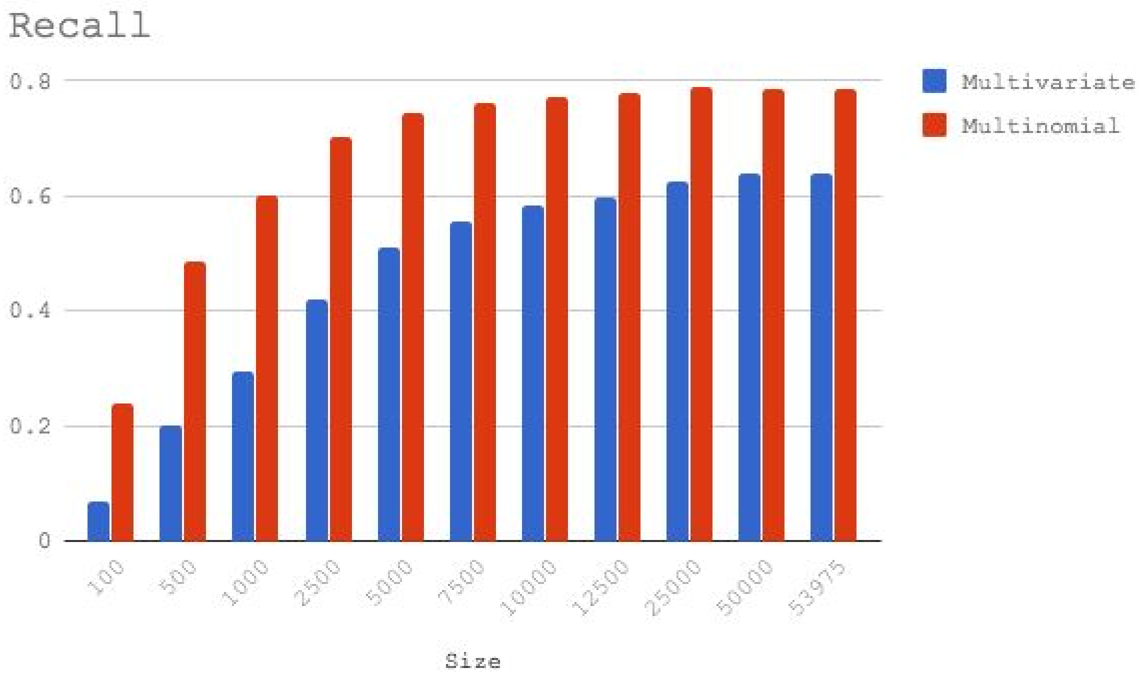
1.



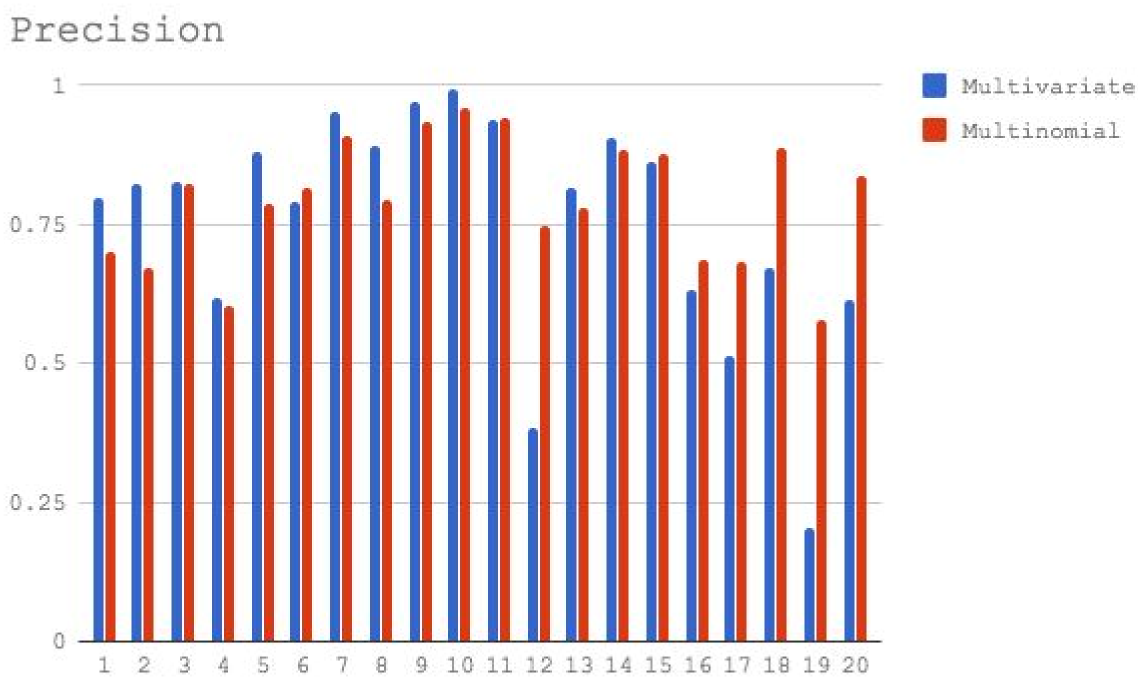
Accuracy
0.8
0.6
0.4
0.2
0
Multivariate
Multinomial
100
500
1000
2500
5000
7500
10000
12500
25000
50000
53975
Size

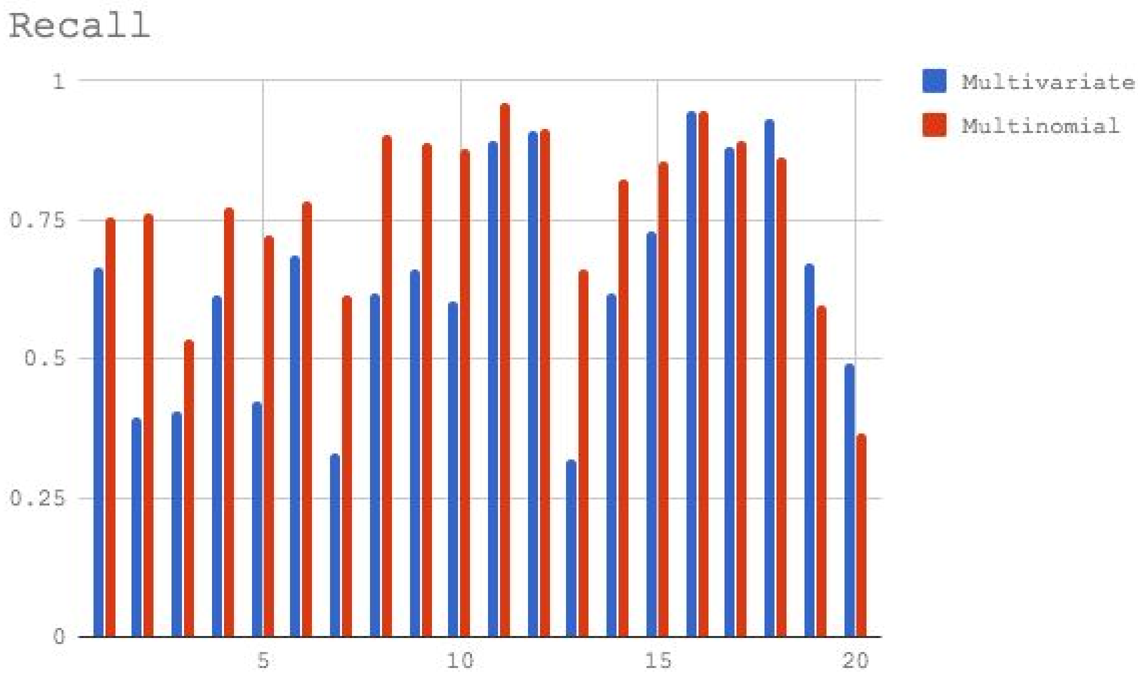


Precision
0.8
0.6
0.4
0.2
0
Multivariate
Multinomial
100
500
1000
2500
5000
7500
10000
12500
25000
50000
53975
Size


Recall
0.8
0.6
0.4
0.2
0
100
500
1000
2500
5000
7500
10000
12500
25000
50000
53975
Size
Multivariate
Multinomial

2


Precision
1
0.75
0.5
0.25
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Multivariate
Multinomial



2.6

1.

$$
\begin{aligned}
\hat{\omega}_{MAP} &= \underset{\omega}{\arg\max}\ P(\omega \mid C) \\
&= \underset{\omega}{\arg\max}\ \frac{P(C \mid \omega) \cdot P(\omega)}{P(C)} \\
&= \underset{\omega}{\arg\max}\ P(C \mid \omega) \cdot P(\omega) \\
&= \underset{\omega}{\arg\max}\ \prod_{C_i \in C} P(C_i \mid \omega) \cdot P(\omega)
\end{aligned}
$$

$$
\hat{\omega}_{MAP} = \underset{\omega}{\arg\max} \left(\sum_{C_i \in C} \log P(C_i \mid \omega) + \log P(\omega) \right)
$$

$$
\hat{\omega}_{MAP} = \underset{\omega}{\arg\max} \left(n_d \cdot \log \omega + (N - n_d) \cdot \log(1-\omega) + \log P(\omega) \right)
$$

2.

$$P(x|\alpha) = \frac{\Gamma\left(\sum_{i=1}^{k} \alpha_i\right)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \prod_{j=1}^{k} x_j^{\alpha_j - 1}$$

$$P(\omega|C) \propto P(C|\omega).P(\omega)$$

$$\propto \left(\prod_{j=1}^{m} \omega_j^{N_j}\right)\left(\prod_{j=1}^{m} \omega_j^{\alpha_j - 1}\right)$$

$$= \prod_{j=1}^{m} \omega_j^{N_j + \alpha_j - 1}$$

$$= P(\omega | N + \alpha)$$

$$\hat{\omega}_{MAP} = \frac{N_j + \alpha_j - 1}{n + \sum_{j=1}^{m} (\alpha_j - 1)}$$

3.

$$\widehat{\theta}_{MLE} = argmax\, P(D|\theta) \text{ and } \widehat{\theta}_{MAP} = argmax\, P(D|\theta) P(\theta)$$

In both cases the MAP is the product of MLE and Prior probability