Deniz ALKIŞLAR / 21502930

1.1

P (M = Machine #2 | G = Loss) =
[P (L | Machine #2) P(Machine #2)] /
[P (L | Machine #2) P(Machine #2) + P (L | not(Machine #2)) P(not(Machine #2))]

P (L | Machine #2) = [(828 / 72)/ (828 + 212 + 18 + 72)]
P (Machine # 2) = 1130 / 1330
P (L | not(Machine #2)) = 135 / 200
P (not(Machine #2)) = 1 – P (Machine #2)

= 0.86

1.2

Machine #1 → Me = 3 / 5 → My friend is more likely to lose.
Friend = 3 / 4

Machine #2 → Me = 828 / 1040 → My friend is more likely to lose.
Friend = 72 / 90

1.3

Me → (40 + 212) / (252 + 888)
Friend → (25 + 18) / 190

In this case, my friend is more likely to lose.

1.4

Lose(me), Win(friend), W(me), L(friend)

(3 / 5) * (1 / 4) * (2 / 5) * (3 / 4) = 0.045

2.1

P (b = 5, r = 5 | C) = ?

b = 2, 3, 4 , 5
r = 3, 4, 5, 6

P (b = 5, r = 5 | C) = (1 / 4) * (1 / 4) = 1 / 16

2.2

P (b = 5, r = 5 | D) = ?

Since product of b and r is an odd number, b or r cannot be an even number.
b = 1 , 3 , 5
r = 1 , 3 , 5

P (b = 5, r = 5 | C) = (1 / 3) * (1 / 3) = 1 / 9

2.3

Condition D is a joint probability whereas condition C is not.

3.1

$$P(X=x)=\frac{\lambda^{x}*e^{-\lambda}}{(x!)}$$

$$\log(\lambda)=\sum_{i=1}^{n}\log(Xi*\log(\lambda)-\lambda-\log(Xi)!)$$

$$\log(\lambda)=\log(\lambda)*\sum_{i=1}^{n}\log(Xi-n(\lambda)-(\sum_{i=1}^{n}\log(Xi)!))$$

$$\frac{d}{d\lambda}\log(\lambda)=\frac{1}{\lambda}\sum_{i=1}^{n}Xi-n=0$$

$$MLE=\frac{1}{n}\sum_{i=1}^{n}(Xi)$$

3.2

Pareto (λ | k ,1) = k * 1^{k} * λ^{-k-1} = k * λ^{-k-1}

Bayes → P(λ|D) = [P(D|λ) * P(λ)] / P(D)

α * P(D|λ) * P(λ)

λ_{MAP} = $\arg_{\lambda}\max P(D|\lambda)*P(\lambda)$

$$(D|\lambda)*P(\lambda)=\prod_{i=1}^{n}\frac{\lambda^{Xi}e^{-\lambda}}{Xi!}k\lambda^{(-k-1)}$$

$$\ln((D|\lambda)*P(\lambda))=\sum_{i=1}^{n}\ln(\lambda)*Xi-\lambda-\ln(Xi!)+lnk-(k+1)ln\lambda$$

$$\ln((D|\lambda)*P(\lambda))=\sum_{i=1}^{n}\frac{d}{d\lambda ln}(\lambda)*Xi-\frac{d}{d\lambda}\lambda-\frac{d}{d\lambda}\ln(Xi!)+\frac{d}{d\lambda}lnk-\frac{d}{d\lambda}(k+1)ln\lambda$$

$$\sum_{i=1}^{n}(\frac{1}{\lambda}*Xi)-\sum_{i=1}^{n}(1)+\frac{(-k-1)}{\lambda}$$

$$\sum_{i=1}^{n}(\frac{1}{\lambda}*Xi)-n+\frac{(-k-1)}{\lambda}$$

$$\lambda MAP\geq 1\rightarrow\frac{1}{n}(\sum_{i=1}^{n}(Xi)-k-1)\geq 1$$

3.3

$$(D|\lambda)*P(\lambda)=\prod_{i=1}^{n}(\frac{\lambda^{Xi}e^{-\lambda}}{Xi!})*U(a,b)$$

$$\ln((D|\lambda)*P(\lambda))=\sum_{i=1}^{n}(\ln(\lambda)*Xi-\lambda-\ln(Xi!))+lnU(a,b)$$

$$\frac{d}{d\lambda}\ln((D|\lambda)*P(\lambda))=\sum_{i=1}^{n}(\frac{d}{d\lambda}\ln(\lambda)*Xi-\frac{d}{d\lambda}\lambda-\frac{d}{d\lambda}\ln(Xi!))+\frac{d}{d\lambda}lnU(a,b)$$

$$\rightarrow \frac{d}{d\lambda} lnU(a,b)=0$$

$$\frac{d}{d\lambda}\ln((D|\lambda)*P(\lambda))=\sum_{i=1}^{n}(\frac{1}{\lambda}*Xi)-\sum_{i=1}^{n}(1)$$

$$\sum_{i=1}^{n}(\frac{1}{\lambda}*Xi)-n=0$$

$$MAP=\frac{1}{n}\sum_{i=1}^{n}(Xi)=MLE$$

4.1

We calculate the probabilities of an e-mail belong to the space or medical group. Both probabilities have a common term in their denominators. So we simply ignore the denominator since it does not change anything. We only consider the greater probability and denominator does not affect the comparison.

4.2

We are given 1600 mails in the training set. They are separated equally so the set is balanced and the percentage of space emails is 50%.

4.3

We have to estimate 2n + 1 parameters where n is the length of the vocabulary (26507).

4.4

The accuracy is 51.5%. The model ended up predicting most of the emails as a medical email. MLE is a bad choice because most of the values in the train feature set is 0. This makes the probabilities -inf and in the case of equality, we predict the test feature as medical. So approximately half of the emails were predicted wrong.

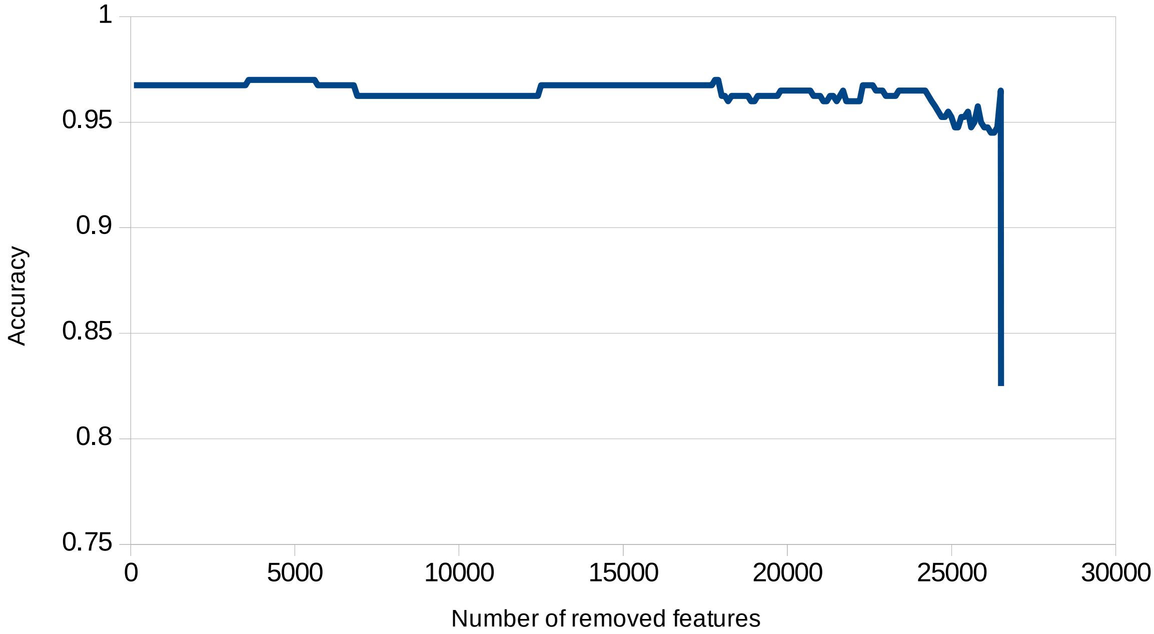
4.5

The accuracy is 96.75% in the case of MAP estimate.

4.6

(index, score)
(25773, 0.23782571866866448)
(11999, 0.10029385007743857)
(13288, 0.09405009311847834)
(2848, 0.08291043624604728)
(14702, 0.0803577299678568)
(15990, 0.07889640202950099)
(614, 0.07685905382356567)
(3300, 0.07369272356385102)
(5749, 0.07039092388813102)
(12807, 0.06950051718036071)

4.7



Max accuracy = 97%