Q3) a) One vs all multi class classification using logistic regression splits the dataset into ~~bina~~ multiple binary classification problem; where in each classification one class is considered separately vs. ~~#~~ all the rest of the classes.
For example if we have 3 class then,

C1 vs [C2, C3]
C2 vs [C3, C1]
C3 vs [C1, C2] will be the classifiers.

Thus for ~~N classifiers~~ N-class instances, we will have N binary classifiers.

The prediction of classes are made with the model which is most confident.

b) One vs One multi class classification using logistic regression splits dataset into ~~multiple~~ binary classification problems; where in each classification one class is considered vs ~~the~~ another class and ~~si~~ thus all pairs of classes are generated.
For example, if we have 3 classes then,

C1 vs C2, ignore C3
C2 vs C3, ignore C1
C1 vs C3, ignore C2 will be the classifiers.

Thus for N-class instances, we will have $\frac{N(N-1)}{2}$ binary classifiers.

Q4) We need to prove that gamma distribution belongs to same family of curve as poisson distribution.

Now, if we can show that both gamma and poisson distribution belong to exponential family then the above fact will be proved true.

① proving that gamma distribution belongs to exponential family.

natural exponential family: $\exp\left\{ \frac{\theta x - b(\theta)}{a(\phi)} + c(x,\phi) \right\}$

gamma distribution: $f(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ where

$x, \alpha, \beta > 0$

taking log on both sides,

$$\log(f(x)) = \log\left(\frac{\beta^{\alpha}}{\Gamma(\alpha)}\right) + \log\left(x^{\alpha-1}\right) + \log\left(e^{-\beta x}\right)$$

$$\Rightarrow \alpha \log(\beta) - \log(\Gamma(\alpha)) + (\alpha-1)\log(x) - \beta x \log(e)$$

$$\Rightarrow \alpha \log(\beta) - \log(\Gamma(\alpha)) + (\alpha-1)\log(x) - \beta x$$

taking exponents on both sides,

$$f(x) = \exp\{-\beta x + \alpha \log \beta + (\alpha - 1)\log x$$

$$- \log \Gamma(\alpha)\}$$

$$\Rightarrow \exp\left\{\frac{-\frac{\beta}{\alpha}x + \frac{\alpha}{\alpha}\log\beta}{1/\alpha} + \right.$$

$$\left.(\alpha - 1)\log x - \log(\Gamma(\alpha))\right\}$$

(dividing the 1st 2 terms by α)

$$\Rightarrow \exp\left\{\frac{\frac{\beta}{\alpha}x - \log\beta}{-1/\alpha} + (\alpha-1)\log(x) - \right.$$

$$\left.\log\Gamma(\alpha)\right\} \quad (1)$$

(taking -1 common in 1st 2 terms)

comparing the above eqn with equation of exponential family —

$$\theta = \frac{\beta}{\alpha} \qquad \phi = \frac{1}{\alpha} \qquad a(\phi) = \frac{-1}{\alpha}$$

The only term which is not matching is $\log\beta$

So, let $\beta = \frac{\theta}{\phi}$

$$\log\beta = \log\theta - \log\phi \quad \text{(taking log on b/s)}$$

substitute in (1) —

$$= \exp\left\{ \frac{\theta x - \log\theta}{-\phi} + \frac{\log\phi}{\phi} + \left(\frac{1}{\phi} - 1\right)\log x - \log\left(\Gamma\left(\frac{1}{\phi}\right)\right)\right\}$$

comparing again with normal equation we get,

$$b(\theta) = \log\theta$$

$$a(\phi) = -\phi$$

$$c(x, \phi) = \left(\frac{1}{\phi} - 1\right)\log(x) - \log\left(\Gamma\left(\frac{1}{\phi}\right)\right)$$

Thus gamma distribution belongs to exponential family

② Proving that poisson distribution belongs to exponential family.

Poisson distribution : $f(x) = \frac{\lambda^x e^{-x}}{x!}$, $\lambda > 0$, $x = 0, 1, 2 \ldots$

taking log on b/s,

$$\log(f(x)) = x\log\lambda - \lambda - \log(x!)$$

taking exponential on both sides,

$$f(x) = \exp\{x \log \lambda - \lambda - \log(x!)\}$$

comparing with natural exponential equation we get,

$$\theta = \log \lambda$$
$$\lambda = e^{\theta}$$
$$b\, \cancel{\log}(\theta) = e^{\theta}$$
$$a(\phi) = 1$$
$$c(x, \phi) = -\log(x!)$$

$$f(x) = \exp\{x\theta - b(\theta) + c(x, \phi)\}$$

Thus poisson distribution belongs to exponential family as well.

Thus, we can say from the above 2 proofs that gamma and poison distribution belongs to same family of curves.

Q7) a) F score is used for testing accuracy of binary classification. It is calculated using precision and recall.

$$\text{Precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}$$

$$\text{Recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}$$

F_1 score is the harmonic mean of precision and recall.

$$F_1 = \frac{2}{\frac{1}{\text{recall}} + \frac{1}{\text{precision}}}$$

F_β is a more general F score $\rightarrow$ $F = \frac{1}{\frac{\alpha}{P} + \frac{(1-\alpha)}{R}}$

$$\Rightarrow F_\beta = \frac{(\beta^2+1)PR}{\beta^2 P + R}$$ where; P = precision, R = recall, $\beta = \frac{1-\alpha}{\alpha}$

where $\alpha \in [0,1]$ and $\beta^2 \in [0 \cdot \infty]$

putting ~~$\beta = 5$~~ $\beta = 5$, we get

$$F_5 = \frac{(25+1)PR}{25P+R} = \frac{26PR}{25P+R}$$

corresponding value of α =

$$25 = \frac{1-\alpha}{\alpha}$$

$\Rightarrow \quad 25\alpha = 1-\alpha$

$\Rightarrow \quad 26\alpha = 1$

$\Rightarrow \quad \underline{\alpha = 0.04}$

Q7) b) $$F_\beta = \frac{(1+\beta^2) \times \text{precision} \times \text{recall}}{(\beta^2 \times \text{precision}) + \text{recall}}$$

we can rewrite the above as —

$$F_\beta = \frac{(\beta^2 + 1)}{\frac{\beta^2}{\cancel{\text{precision}}\ \text{recall}} + \frac{1}{\text{precision}}}$$

now we can see if $\beta > 1$ then F_β will give more weight to recall. Thus we can say that when β is changed from 1 to 5 then recall will be given more emphasis classmate

Date
Page

Q5) a) K means will perform a good job when clusters are spherical in shape and also that the data can be clustered i.e. in the data the points must be close to each other of similar clusters 2 points must be far away in case of different clusters. ~~Confusi~~ Covariance matrix will give us information about individual clusters and not the dissimilarity ~~of~~ between 2 clusters. Also covariance matrix gives no information about similarity of data points. Thus the key points for which K means would have performed well are not satisfied.