Assignment 2

Q3 (a)

Ans For one vs all multi-class classification using logistic regression :-

N class instances need n binary classifiers.

Because in each instance we are trying to classify one class against the rest.

For One Vs One multiclass classification using logistic regression :

N class instances need $\boxed{\frac{N \times (N-1)}{2}}$ binary classification.

Because in each instance we will try to classify one class against every other class individually so for 1st class we need (N-1) classifiers, for the next class we would need (N-2) classifiers ....

$$= 1+2+3 \cdots (n-1)$$

$$\boxed{\Rightarrow \frac{(n-1)(n)}{2}}$$

Q4 To prove : Gamma Distribution belongs to the same family of curves as poisson distribution

proof : we are going to prove this by proving that both the distributions belong to exponential family

Expression for Exponential Distribution :-

$$\boxed{P(y:\vec{n}) = b(y)\,\exp(\vec{n}\,T(y) - a(\vec{n}))}$$

where,

$y \rightarrow$ data

$\vec{n} \rightarrow$ vector of natural parameters

$b(y) \rightarrow$ base measure

$T(y) \rightarrow$ Sufficient statistic

$a(\vec{n}) \rightarrow$ log partition

poisson distribution

$$\boxed{P(y;\lambda) = \frac{\lambda^y e^{-\lambda}}{y!}} \quad \lambda > 0, \; y = 0, 1, 2 \cdots$$

taking log on both sides :-

$$\log(P(y;\lambda)) = y\log\lambda - \lambda - \log(y!)$$

taking exp both sides

$$P(y;\lambda) = \exp(y\log\lambda - \lambda - \log(y!))$$

$$= \exp(-\log(y!))\exp(y\log\lambda - \lambda)$$

let, $b(y) = \exp(-\log(y!))$, $T(y) = y$, $\eta = \log\lambda$

$$\lambda = e^{\eta}$$

substituting these values

$$P(y;\eta) = b(y)\cdot\exp(\eta T(y) - e^{\eta})$$

let, $a(\eta) = e^{\eta}$

Hence: $\boxed{P(y:\eta) = b(y)\exp(\eta T(y) - a(\eta))}$

Poisson Distribution belongs to exponential family

(ii) Gamma Distribution

$$\boxed{P(y;\beta,\alpha) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y}} \quad y,\alpha,\beta > 0$$

log both sides

$$\log P(y;\beta,\alpha) = \alpha\log(\beta) - \log(\Gamma(\alpha)) + (\alpha-1)\log(y) - \beta y$$

take exp on both sides

$$P(y;\beta,\alpha) = \exp((\alpha-1)\log(y) - \beta y + \alpha\log\beta - \log(\Gamma(\alpha))$$

$$= \exp\left([(\alpha-1), -\beta]\begin{bmatrix}\log(y) \\ y\end{bmatrix} - (\log(\Gamma(\alpha) - \alpha\log\beta)\right) \quad \text{—} ①$$

$$\vec{\eta} \rightarrow [\underset{\eta_1}{\underset{\downarrow}{(\alpha-1)}}, \underset{\eta_2}{\underset{\downarrow}{-\beta}}], \quad T(y) = \begin{bmatrix}\log(y) \\ y\end{bmatrix}, \quad b(y) = 1$$

substituting these values in eq I, we get

$$b(y)\cdot exp(\vec{\eta}^T T(y) - (\log(\Gamma(\eta+1)) - (\eta_1+1)\underbrace{\log(-\eta_2))}_{a(\eta_1,\eta_2)}$$

$$a(\vec{\eta}) = a(\eta_1, \eta_2)$$

$$\boxed{P(y;\eta) = b(y)\cdot exp(\vec{\eta}^T T(y) - a(\vec{\eta}))}$$

$\therefore$ Gamma Distribution belongs to exponential family.

Since both poisson & Gamma belong to exponential family of curves, hence both belong to the same family of curves.

Q7 (a) Derieve: F_5 score in terms of precision & recall

FB. Score is an adjustable single score metric used in machine learning for evaluating binary classification model using precision & recall values for the class

The general formula for Fβ score is :-

$$F\beta\text{-score} = \frac{(1+\beta^2)(\text{Precision} \times \text{recall})}{\beta^2 \text{precision} + \text{recall}}$$

For F5 score, $\beta = 5$, we have

$$F5 \text{ score} = \frac{(1+5^2)(\text{Precision} \times \text{recall})}{5^2 \text{precision} + \text{recall}}$$

$$F_5\text{-Score} = \frac{26(\text{Precision} \times \text{recall})}{25(\text{precicision} + \text{recall})}$$

Now, for value of α, we know that

$$\alpha = \frac{1}{1+\beta^2} \quad (\because \beta = 5)$$

$$= \frac{1}{1+5^2} = \frac{1}{1+25}$$

$$\boxed{\alpha = 1/26 = .0385}$$

Q7(b)

Ans

Precision

→ Precision deals with total correct +ve classifications out of all the +ve classif.

→ Precision is a measure of how many of the +ve predictions made are correct (True +ve)

→ we try to minimize false +ves in precision

→ Formula :-

$$\boxed{\text{precision} = \frac{TP}{TP + FP}}$$

TP → True +ve

FP → False +ve

Recall

→ Recall is a measure of how many of the +ve cases the classifier correctly predicted, overall the +ve cases

→ In Recall, we try to minimize false -ves

→ $$\boxed{\text{Recall} = \frac{TP}{TP + FP}}$$

F_N → False -ve

TP → True +ve

Observations

→ when we, decrease false +ve then false negatives increase, when we increase precision, recall decreases

→ when we decrease false negative then false positives increases (ie) when we increase recall, precision decreases

→ In F5-score, the value of β is higher so, it giving less weight to precision then recall, whereas in F_1 score equal weight is given to both precision & recall

Question 8 :

Ans : Leave P out Cross Validation is best suited for given dataset as our data set is already balanced. Our data is already balanced, leave P out will everytime leave different P data for test and perform training on remaining and find the best possible combination of train & test such that our model can neither be overfit or underfit

If data would have been unbalanced stratif-ied 3 fold Cross validation could have been a better idea

As data is already balanced we need not to worry about ratio of classes in train & test.