**Question-1: Why can’t we use accuracy as a metric for an imbalanced dataset?**

In case of binary classification, we know that imbalance dataset means positive points>> negative points or negative points>>positive points

Let suppose we have 1000 point in which 300 is negative and 700 is positive and we do classification using k-nn where is k is not 1, as we have imbalance dataset, so majority vote is more in positive side. So, it is more possible we get positive class on negative class.

TP is 700 and TN is 200

Where accuracy is (true positive + true negative)/ len(dataset)

(700+200)/1000 = 900/1000 = 0.9

According to the accuracy it is 90% but model is dumb as it predicts more positive over the negative class, doing wrong predictions

**Question-2: In which scenario we will prefer recall over precision? What are examples from the real world where a high precision model is desirable and when a high recall model is desirable?**

Precision means ratio of (True Positive/ True positive +True Negative)

Here we want to know about the percentage of the actual positive among the predicted positive

Recall means the ratio of True Positive / Positive)

Here we want to know actual True percentage which belong to the True

In medical application we need to use the recall, because we need to focus on the true positive (this should be high) and the false positive (very low)

In recommendation system we need to use the precision because we need low value of false positive i.e. we need not to recommended the false item , it is ok to recommended less true item than the false item

**Question-3: What are the different performance metrics that can be used for Multiclass classification problems?**

Precession, recall and F-1 score

**What is macro averaged precision and micro averaged precision?**

Macro averaged precision and micro average precision is used as the multiclass performance metric,

* in **macro averaged** precision we compute the precision for every class independently respectively and then compute the average precision
* in **Micro averaged** precision we compute the aggregate precision of the class to find the average precision

**Question -4 Which of the following statements is/are correct about AUC metric ?**

**a) It tells how much the model is capable of distinguishing between classes.** True

**b) The AUC of a random model is 0.5.** True

**c) We can use AUC only for binary classification problems.** True / (but we can use it for multiclass like we have 4 class now for 2 class, 1,0 if class 2 then 1 else 0)

**d) Mathematically,it is the expectation that a uniformly drawn random positive is ranked before a uniformly drawn random negative.** True

**Question-5 : What is the most common metric used for Forecast Accuracy(Future prediction on Stock Market, Future Sale in Business)?**

Mean percentage absolute error