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

Because, Since the dataset is imbalanced i.e. Either of class has too much data points.

(in binary classification) then even if the model gives single output as majority

class then also it will have very high accuracy.

**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?**

When we don’t want any positive point (actual) to be predicated as Negative point. i.e. Basically We want minimum FalsePositive value and High TruePoisitve value.

This many used when we want to predicate the patient has contagious disease or not

(like corona ) .

Recall = TP/(TP+FN) i.e. when we want all the predicated positive points should be true positive and we should not miss many positive points.

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

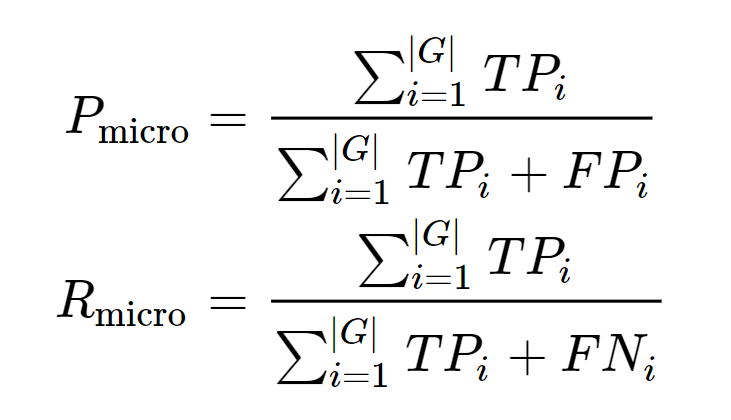
**What is macro averaged precision and micro averaged precision?**

In binary classification, we can easily define Confusion matrix . But in Multiclass classification Confusion Matrix can be build with one vs rest (all other).

With that strategy we will have K Confusion Matrices (considering k – class classification).

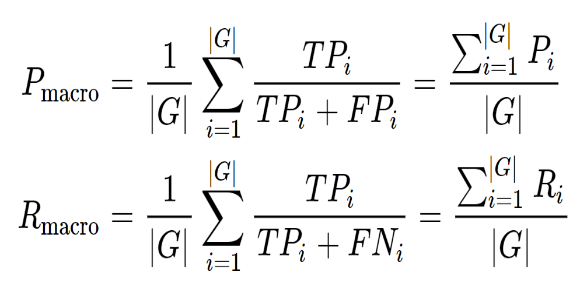
Then we can find Two types of Precision , Recall and F1-Score. Namely Macro averaged and another is Micro averaged.

Micro averaged Precision means here we are saying micro mean we consider each observation of all classes and then take average.

|G| = Number of classes

Macro averaged precision Means Here we are taking the average of precision of each class.

That’s why it is macro instead of focusing of observation of each class we are taking the matric of that class.



|G| = Number of classes

Reference : <https://www.datascienceblog.net/post/machine-learning/performance-measures-multi-class-problems/>

**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.**

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

**c) We can use AUC only for binary classification problems.**

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

Since AUC plots the TPR vs FPR it gives clear picture about positive class. So A is True

B is also True

C is False . Because We can draw AUC for multiclass with one vs all approach.

D is True

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

**Mean Absolute Percentage Error (MAPE) :**

it tells you by how many percentage points your forecasts are off, on average