ML Theortical questions

1.Accuracy is calculated as [(TP+TN)/(TP+TN+FP+FN)], and is also a part of the performance's statistics as well as- F1 score, sensitivity, specificity, TP, FP, etc...

Similarly to the example precented in lecture 8, Integrating a large population sample can be can be misleading considering the accuracy only. If we take a 900 patients population-100 of them have AF and 800 are non-AF, We can get very imbalanced data so there will be bad learning even though the accuracy is 89%. It is common in medicine. So to conclude, other stats other than Accuracy provide better insights when classes are skewed.

4. Logistic regression is great in a low number of dimensions and when the predictors don't suffice to give more than a probabilistic estimate of the response. SVMs do better when there's a higher number of dimensions, and especially on problems where the predictors do certainly (or near-certainly) determine the responses. Additionally, SVMs only consider points near the margin (support vectors). Logistic regression considers all the points in the data set. Practically speaking, when training SVM , using RBF kernel there will be two hyperparameters to search for- the capacity and the gamma.

להרחיב עוד על ההיפר פרמטרים