Core points, border points, anomalies

Precision = TP/(TP+FP) what percentage of patient are really 1 out of all predicted to be 1.

Recall = TN/(TN+FN) what percentage of patient are really 0 out of all predicted to be 0.

F1 = 2(precision \* recall)/(precision + recall)

Support = sum of patient in each category, number of occurrence in each category.

XGBoost is advanced version of gradient boosting, focused on efficiency and speed. GradientBoosting is very slow

Both xgboost and gbm follows the principle of gradient boosting. There are however, the difference in modeling details. Specifically, xgboost used a more regularized model formalization to control over-fitting, which gives it better performance. XGBoost uses 2nd order derivatives where GBM uses the loss function. XGBoost – Regularized boosting Dropout regularization for trees; GB focused on the variance but not the trade-off between bias, XGBoost focus on the regularization factors

Stacking is great for small or medium size datasets

Usually, XGBoost could be better than stacking

Stacking is less widely used as bagging and boosting:

1. Split the training set into two disjoint sets.

2. Train several base learners on the first part.

3. Test the base learners on the second part.

4. Using the predictions from 3) as the inputs, and the correct responses as the outputs, train a higher level learner.