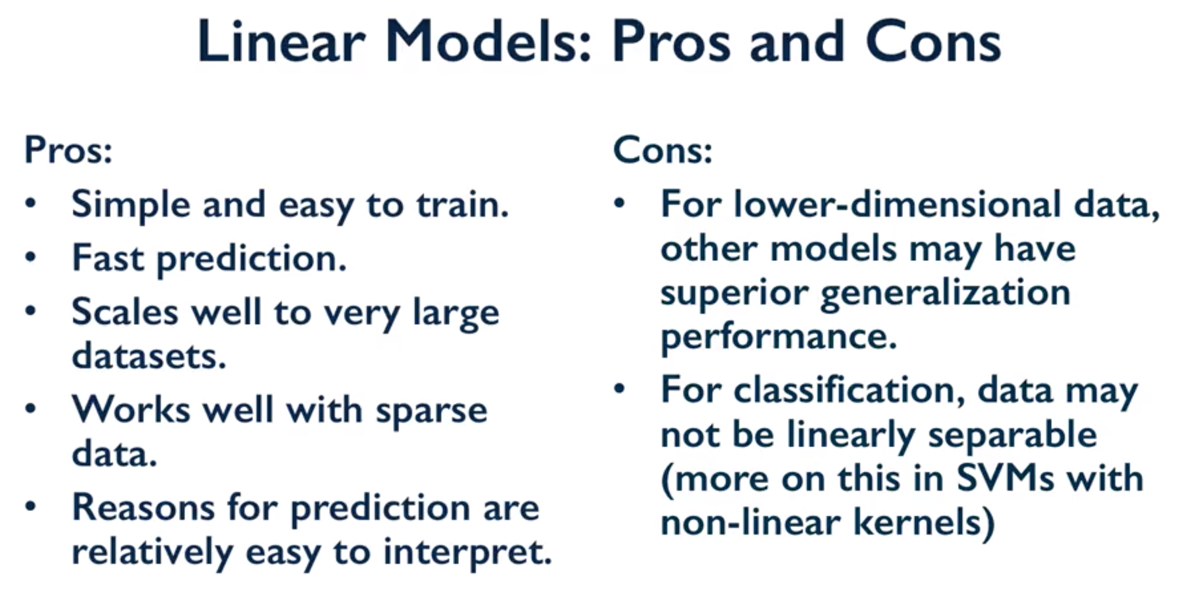
Course2

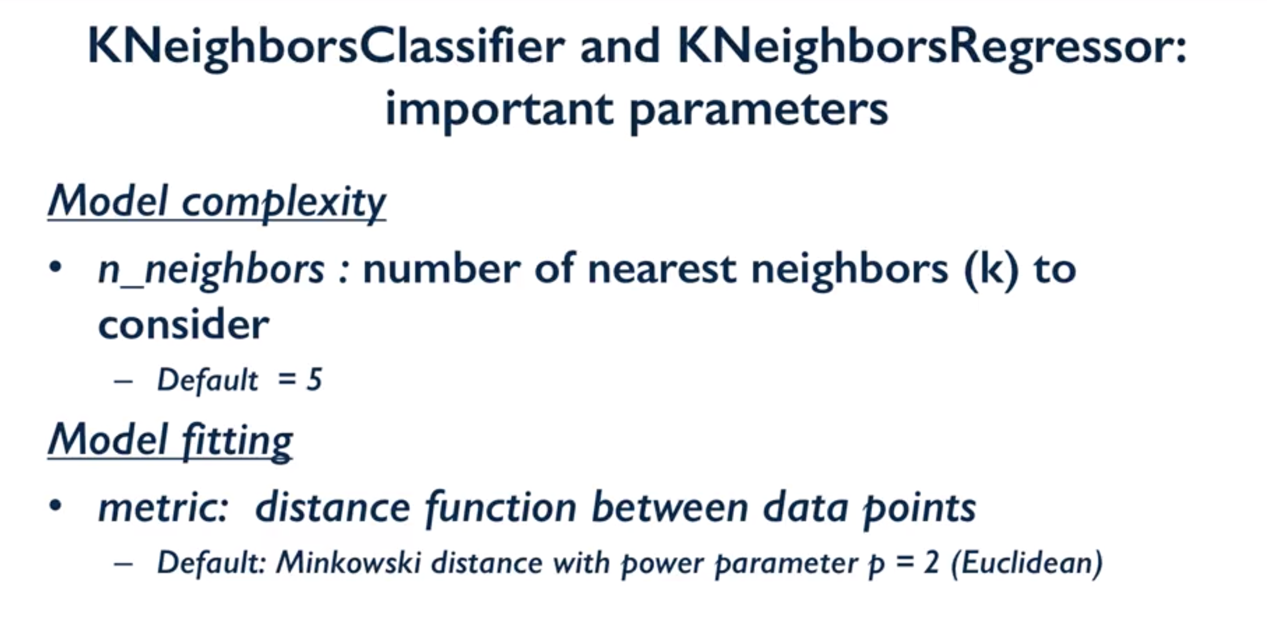
* KNN
* Linear Regression – OLS

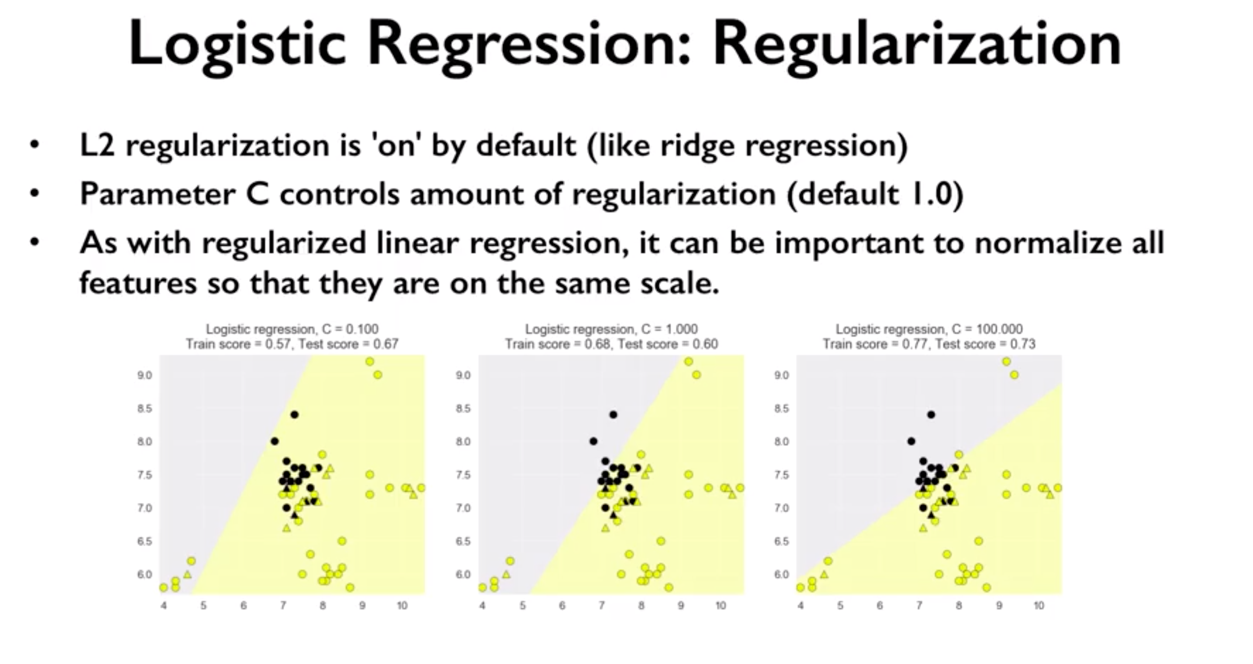
Ridge (L2 penalty)

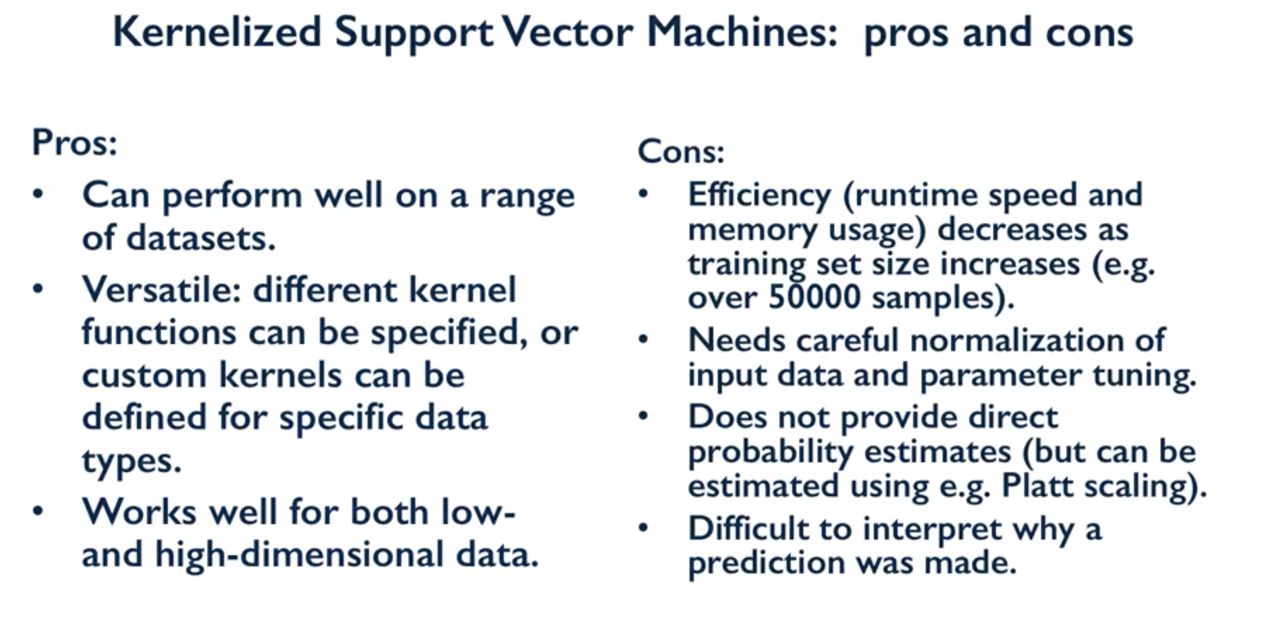
Lasso (L1 penalty) -> sparse model

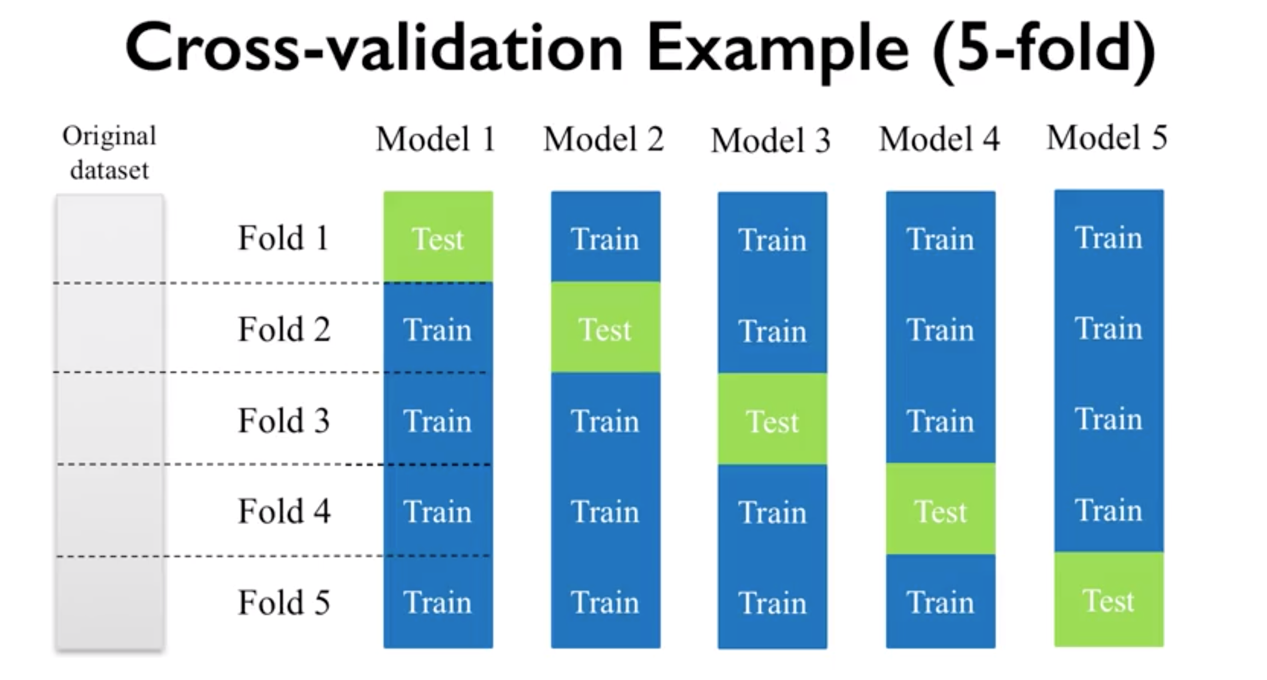
* Logistic Regression
* Cross validation
  + Stratified cross validation: The Stratified Cross-validation means that when splitting the data, the proportions of classes in each fold are made as close as possible to the actual proportions of the classes in the overall data set as shown here.
* Decision tree

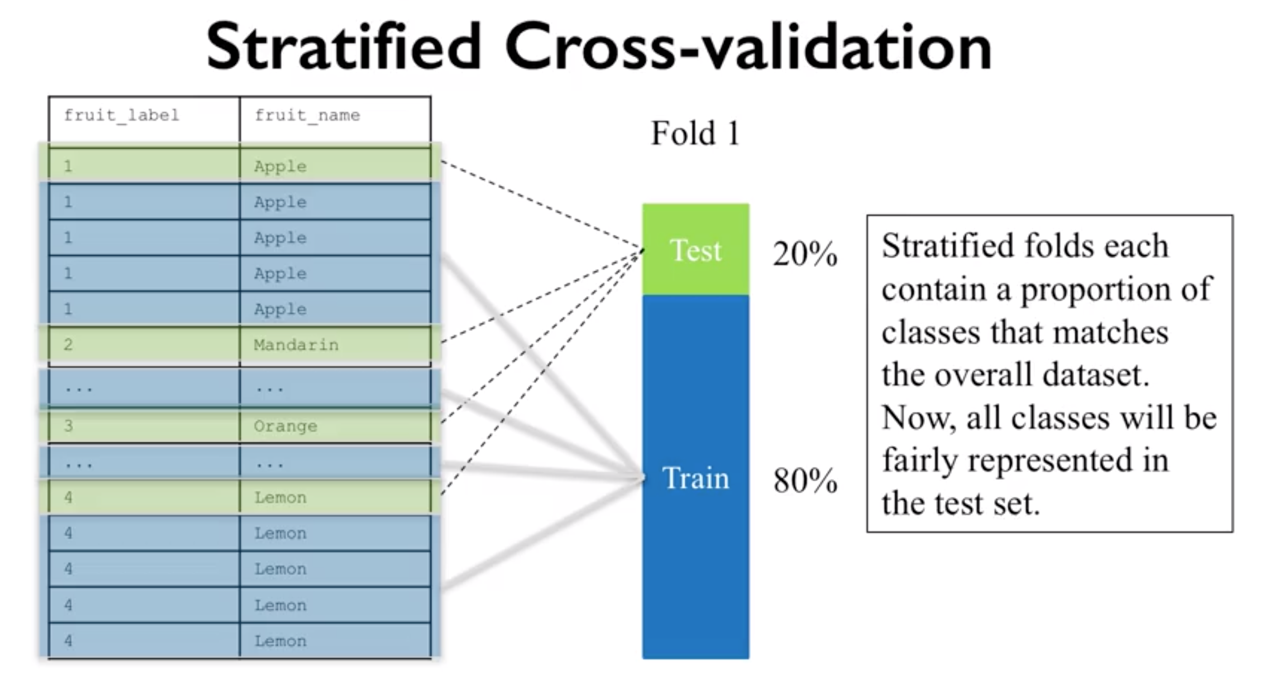


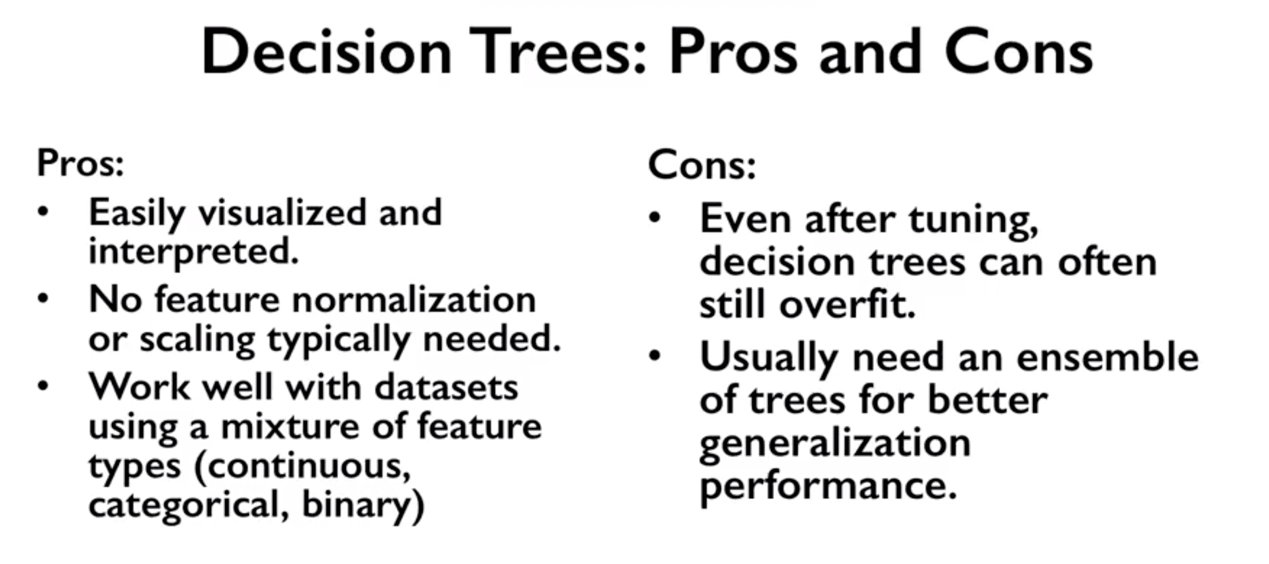






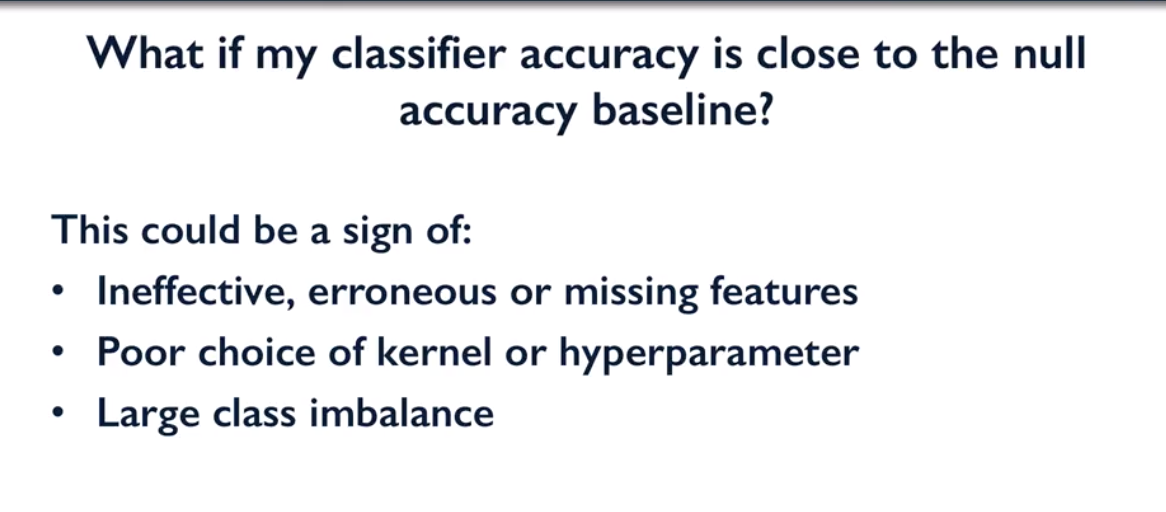




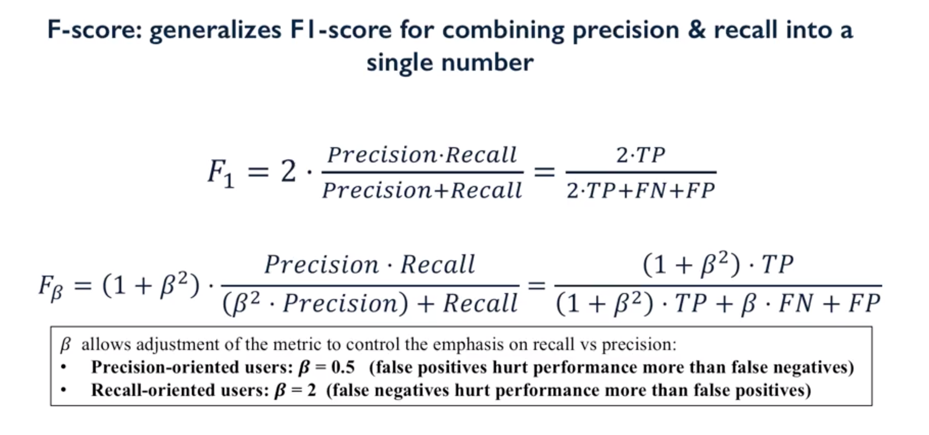


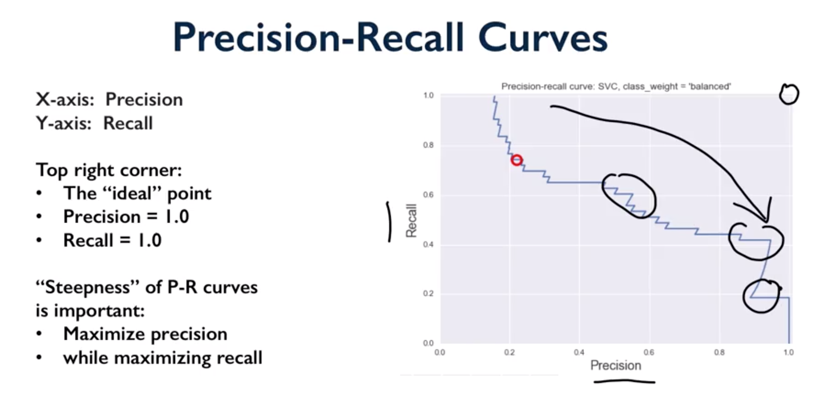
Course 3

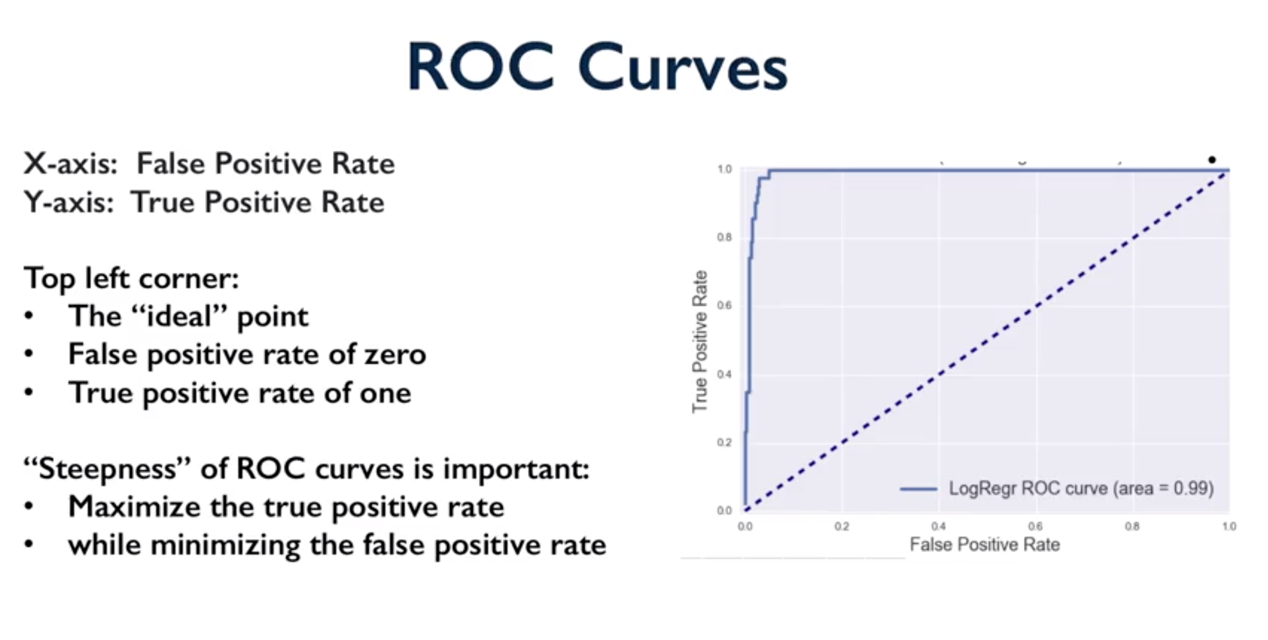
* Model evaluation
  + Dummy classifier



* Confusion matrix
* Precision, Recall, ROC curve
  + Trade-off between Precision(TP/TP+FP) and Recall(TP/TP+FN)
    - Recall oriented: tumor detection, legal discovery
    - Precision oriented: search engine ranking, query suggestion, customer facing
    - F1 score = 2\*Precision\*Recall/(precision+recall)



* + 



* Multi-class evaluation
  + Confusion matrix
* Regression evaluation
* Model Selection

