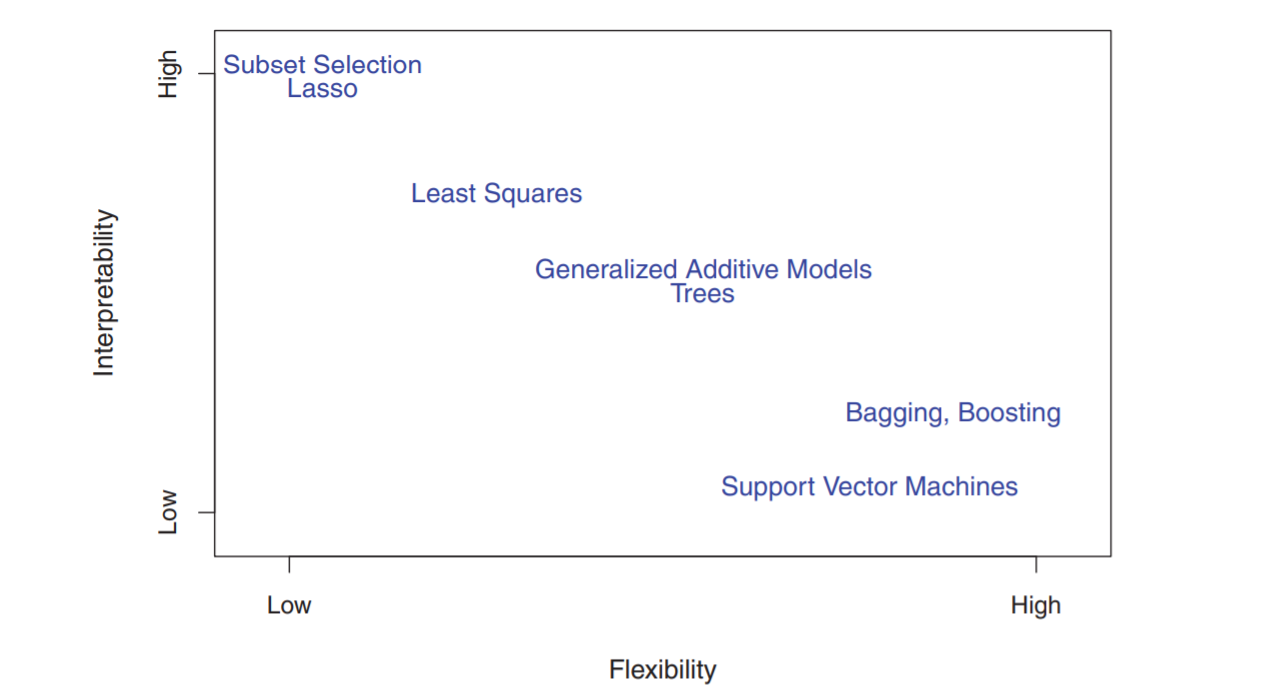
**Notes for Introduction to Statistical Learning**

**Chapter 2**

* We estimate *f* for
  + *prediction (for a given X what is Y)*
  + *inference (how does Y change with respect to X)*
* Accuracy of our model depends on *reducible error*  and *irreducible error(e)*
* **Parametric** Models – reduces the problem of estimating f down to one of estimating a set of parameters (coefficients)
  + First, make an assumption about the functional form ie linear Y = B0 + B1X1 + …+ BnXn
  + After model has been selected fit to training data ie estimate the coefficients B0 … Bn using ordinary least squares
* **Non-parametric** Models – do not make explicit assumptions about the functional form of *f*
  + Main advantage – by avoiding making assumptions have the potential to accurately fit a wider range of possible shapes for *f*
  + Main disadvantage – since we do not reduce problem to estimating parameters, need a large number of observations
  + Example – thin-plate spline
* Tradeoff Prediction Accuracy vs Model Interpetability
  + If we are mainly interested in inference, then restrictive models are muchFo more interpretable
* In unsupervised learning there is labeled training data we only see X vars but no outcome variable – ie cluster analysis like market segmentation
* Quantative Y var = regression, qualitative Y var = classification
* Measuring Model accuratcy
  + For regressions, mean squared error (MSE) = Ave (y – yhat)^2
  + As model flexibility increases training MSE will decrease and test MSE will increase
* Cross validation = method for estimating test MSE using the training data
* Bias-Variance Tradeoff =