Chapter 14. Combining Models

14.1. Bayesian Model Averaging

(pp. 654 - 655) Difference between ***Model Combination*** *methods* and ***Bayesian Model Averaging***

- In Model Combination, different data points withn the full set can potentially be generated from different values of the latent variable $\mathbf{z}$ and hence by different components/models

$$p(\mathbf{X}) = \prod_{n=1}^{N} p(\mathbf{x}_n) = \prod_{n=1}^{N} [\sum_{\mathbf{z}_n} p(\mathbf{x}_n, \mathbf{z}_n)]$$

  i.e., each individual data $\mathbf{x}_n$ has a corresponding latent variable $\mathbf{z}_n$
- In Bayesian Model Averaging,

$$p(\mathbf{X}) = \sum_{h=1}^{H} p(\mathbf{X}|h)p(h)$$

  Each model indexed by h contributes to the entire dataset $\mathbf{X}$ with prior probability $p(h)$. Of course as the size of the data set increases, this uncertainty reduces as the posterior probabilities $p(h|\mathbf{X})$ become more focussed on one of the models.

14.2. Committees

14.3. Boosting

14.4. Tree-based Models

14.5. Conditional Mixture Models