

Derive variance of final prediction

$$\begin{aligned}\text{Var}(\bar{Y}) &= \text{Var}\left(\frac{1}{m}\sum_{i=1}^m Y_i\right) \\ &= \frac{1}{m^2}\sum_{i=1}^m \sum_{j=1}^m \text{Cov}(Y_i, Y_j)\end{aligned}$$

Since $\text{Var}(Y_i) = \sigma^2$ and $\text{Cov}(Y_i, Y_j) = \rho\sigma^2$ for $i \neq j$

$$\begin{aligned}\text{Var}(\bar{Y}) &= \frac{1}{m^2}(m\sigma^2 + m(m-1)\rho\sigma^2) \\ &= \frac{1}{m^2}(m\sigma^2 + m^2\rho\sigma^2 - m\rho\sigma^2) \\ &= \frac{\sigma^2}{m} + \rho\sigma^2\left(1 - \frac{1}{m}\right) \\ &= \frac{\sigma^2}{m} + \rho\sigma^2 - \frac{\rho\sigma^2}{m} \\ &= \rho\sigma^2 + \frac{(1-\rho)\sigma^2}{m}\end{aligned}$$

Thus, the variance of the final prediction is given by $\rho\sigma^2 + \frac{(1-\rho)\sigma^2}{m}$

The same logic applies to classification because the averaging of probabilities (or votes) from each base model reduces variance similarly to the averaging of continuous values.