Теоретические задачи

2.1 Bias Varience decomposition

$$\mathbb{E}_{x,y}\mathbb{E}_{X^l}(y - a_{X^l}(x))^2 =$$
$$= \mathbb{E}_{x,y,X^l}(y - a_{X^l}(x))^2 =$$
$$= \mathbb{E}_{x,y,X^l}(\mathbb{E}_{x,y,X^l}(y - a_{X^l})^2|x)$$
$$= \mathbb{E}_{x,y,X^l}((y - a_{X^l})|x) =$$
$$= \mathbb{E}_{x,y,X^l}((a_{X^l}(x) - f(x) - \varepsilon)^2|x) =$$
$$= \mathbb{E}_{x,y,X^l}((a_{X^l(x)} - f(x))^2|x) =$$
$$= \mathbb{E}_{x,y,X^l}((a_X^l(x) - f(x))^2|x) + 2\mathbb{E}_{x,y,X^l}\varepsilon\mathbb{E}_{x,y,X^l}(a_{X^l} - f(x)|x) + \mathbb{E}_{x,y,X^l}(\varepsilon^2|x) =$$

Все без последнего слагаемого:

$$\mathbb{E}_{x,y,X^l}(a_{X^l} - \mathbb{E}_{x,y,X^l}a_{X^l}(x) + \mathbb{E}_{x,y,X^l}a_{X^l}(x) - f(x))^2 =$$
$$\mathbb{E}_{x,y,X^l}[(a_{X^l(x) - \mathbb{E}_{x,y,X^l}a_{X^l}(x)})^2|x] + \mathbb{E}_{x,y,X^l}[(E_{x,y,X^l}a_{X^l}(x) - f(x))^2|x] +$$
$$+2\mathbb{E}_{x,y,X^l}[(a_{X^l} - \mathbb{E}_{x,y,X^l}a_{X^l})(\mathbb{E}_{x,y,X^l}a_{X^l} - f(x))|x] =$$
$$\mathbb{E}_{x,y,X^l}[(a_{X^l(x) - \mathbb{E}_{x,y,X^l}a_{X^l}(x)})^2|x] + \mathbb{E}_{x,y,X^l}[(E_{x,y,X^l}a_{X^l}(x) - f(x))^2|x]$$

То гда все выражение равно:

$$\mathbb{E}_{x,y}(y - \mathbb{E}(y|x))^2 + \mathbb{E}_{x,y}(\mathbb{E}(y|x) - \mathbb{E}_{X^l}a_{X^l}(x))^2 + \mathbb{E}x, y\mathbb{E}X^l(a_{X^l}(X) - \mathbb{E}_{X^l}a_{X^l}(x))$$

2.2 Смещение и разброс в бэггинге

$$MSE = \mathbb{E}_{X,Y,x,y}[\frac{1}{n}\sum_{i=1}^{n} - y(x)]^2 =$$

$$= \mathbb{E}_{X,Y,x,y}(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i)^2 =$$

$$= \frac{1}{n^2}\mathbb{E}_{X,Y,x,y}(\sum_{i=1}^{n}\varepsilon_i^2(x) + \sum_{i \neq j}\varepsilon_i\varepsilon_j) =$$

В случае нескоррелированных ошибок базовых алгоритмов

$$= \frac{1}{n}(\frac{1}{n}\mathbb{E}_{X,Y,x,y}\varepsilon_i^2(x))$$

В случае несмещенной ошибки получаем, что мы уменьшили varience в n раз (понятно что в общем случае можно теоретически сместить оценкую и рассматривать опять несмещенную ошибку)

2.3 Корреляция ответов базовых алгоритмов

$$\mathbb{D}_{X,Y,x,y}(\frac{1}{M}\sum_i \xi_i) = \frac{1}{M^2}\mathbb{D}_{X,Y,x,y}(\sum_i \xi_i)$$

по формеле дисперсии суммы

$$= \frac{1}{M^2} M \mathbb{D}_{X,Y,x,y} \xi_1 + \frac{1}{M^2} \sum_{i \neq j} cov(\xi_i, \xi_j) =$$

$$= \frac{1}{M} \sigma^2 + \frac{1}{M^2} \sum_{i \neq j} cov(\xi_i, \xi_j) =$$

$$= \frac{1}{M} \sigma^2 + \frac{M(M-1)}{M^2} \rho \sigma^2 =$$

$$= \frac{1}{M} \sigma^2 + \frac{M-1}{M} \rho \sigma^2 =$$

$$= \rho \sigma^2 + \frac{\sigma^2}{M} (1 - \rho)$$

Что и требовалось