ДЗ №3

Даниил Дмитриев, 494

27 марта 2017 г.

1 Задание 2.2

$$a(x) = \frac{1}{k}\sum_{i=1}^{k} a_i(x)$$

$$\mathbb{E}_{X,Y,x,y}(a(x)) = \frac{1}{k}\sum_{i=1}^{k}\mathbb{E}_{X,Y,x,y}(a_i(x)) = \mathbb{E}_{X,Y,x,y}(a_1(x)),$$

значит **Bias** не меняется.

$$\mathbf{Var}_{X,Y,x,y}(a(x)^2) = \frac{1}{k}\mathbf{Var}_{X,Y,x,y}(a_1(x)) + \frac{1}{k^2}\sum_{i\neq j}\mathbf{cov}(a_i(x), a_j(x))$$

Пусть коэффициент корреляции (a_i, a_j) равен ρ, тогда

$$\frac{1}{k}\mathbf{Var}_{X,Y,x,y}(a_1(x)) + \frac{1}{k^2}\sum_{i\neq j} cov(a_i(x), a_j(x)) =$$

$$= \frac{1}{k}\mathbf{Var}_{X,Y,x,y}(a_1(x)) + \frac{1}{k^2}\sum_{i\neq j}\rho\mathbf{Var}_{X,Y,x,y}(a_1(x)) =$$

$$= \left(\frac{1}{k} + \frac{\rho(k-1)}{k}\right)\mathbf{Var}_{X,Y,x,y}(a_1(x))$$

Значит, чем меньше корреляция между базовыми алгоритмами, тем меньше **Variance**.

2 Задача 2.3

Пусть $\xi_1, ..., \xi_M$ - одинаково распределенные величины, $D\xi_1 = \sigma^2, cor(\xi_1, \xi_2) = \rho$. Обозначим $\mathbb{E}\xi_1 = a$, тогда $\mathbb{E}\xi_1^2 = \sigma^2 + a^2$.

$$D\overline{\xi} = \frac{1}{M^2}D(\xi_1 + ... + \xi_M) = \frac{1}{M^2}\mathbb{E}(\xi_1 + ... + \xi_M)^2 - a^2.$$

$$\mathbb{E}(\xi_1 + ... + \xi_M)^2 = M(\sigma^2 + a^2) + \sum_{i \neq j} \mathbb{E}\xi_i\xi_j$$

Из определения корреляции следует, что $\mathbb{E}\xi_i\xi_j = \rho\sigma^2 + a^2$.
Значит,

$$\mathbb{E}(\xi_1+...+\xi_M)^2 = M(\sigma^2+a^2)+M(M-1)(\rho\sigma^2+a^2) = M^2(\rho\sigma^2+a^2)+M\sigma^2(1-\rho),$$

откуда,

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

3 Задание 3.1

В *sample_submission* для всех значений предсказывается среднее по всем таргетам из *train*