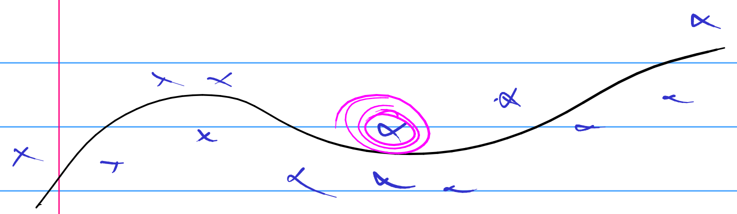
① Bagging - bootstrap aggregation

$\boxed{N}$ точек



- K раз
- выбрать N точек случайно с повторени[ем]
- обучить M_i
- $M(\bar{x}) = \frac{1}{K}\sum_{i=1}^{K} M_i(\bar{x})$

$$y_m(\bar{x}) = h(\bar{x}) + \varepsilon_m(\bar{x})$$

$$E_{\bar{x}}\left[\left(y_m(\bar{x}) - h(\bar{x})\right)^2\right] = E_x\left[\varepsilon_m^2\right]$$

$$\boxed{E_{Avg} = \frac{1}{M}\sum_{m=1}^{M} E_x\left[\varepsilon_m^2\right]}$$

$$E_x\left[\varepsilon_m(\bar{x})\right] = 0$$

$$E_x\left[\varepsilon_m \varepsilon_{m'}\right] = 0$$

$$E_{com} = E_x\left[\left(\frac{1}{M}\sum_{m=1}^{M} y_m(\bar{x}) - \underbrace{h(\bar{x})}_{\frac{1}{M}\sum h(\bar{x})}\right)^2\right] =$$

$$= E_x\left[\left(\frac{1}{M}\sum_m \left(y_m(\bar{x}) - h(\bar{x})\right)\right)^2\right] = \boxed{E_x\left[\frac{1}{M^2}\underbrace{\left(\sum_m \varepsilon_m(\bar{x})\right)^2}_{\sum \varepsilon_m^2}\right]}$$

$$E_{com} = \frac{1}{M} E_{Avg} \quad \leq \qquad E_{com} \leq E_{Avg}$$

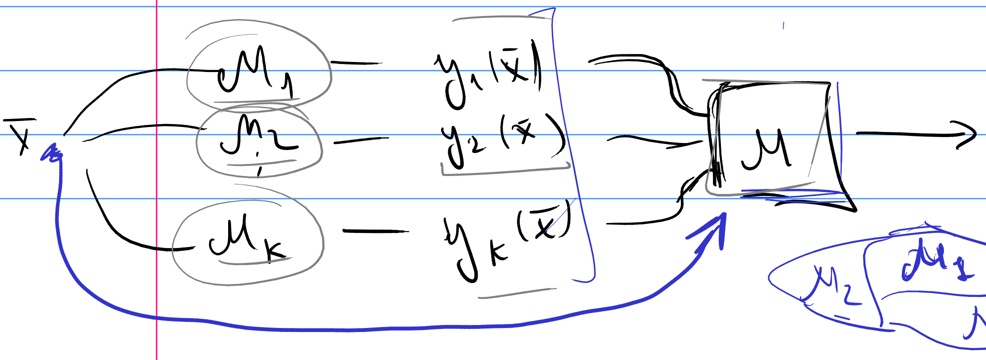
②

Model selection $\quad M_1, M_2, M_3 \quad p(D|M_i)$

Model composition $\quad M = M_1, M_2, M_3$

blending

stacking



$$y(\bar{x}) = \bar{w}^T \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix}$$

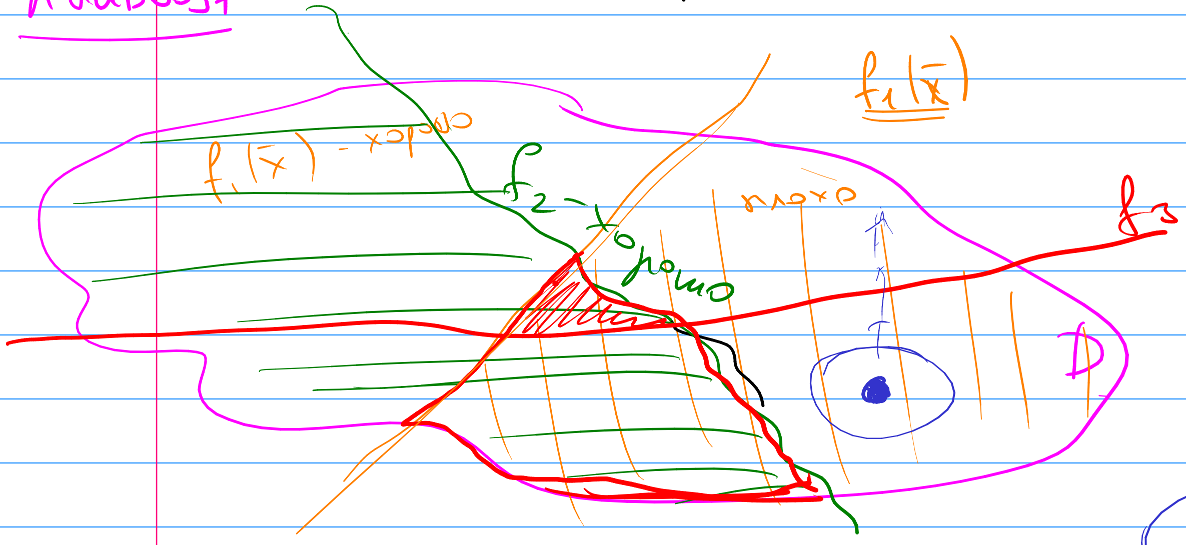
③ Boosting

$$F(\bar{x}) = \alpha_1 f_1(\bar{x}) + \alpha_2 \underbrace{f_2(\bar{x})} + \dots + \alpha_k f_k(\bar{x})$$

$$\boxed{\alpha_m, f_m(\bar{x})} \quad \underbrace{F_{m-1}(\bar{x})} = \sum_{i=1}^{m-1} \alpha_i f_i(\bar{x})$$

$$F_m(\bar{x})$$

AdaBoost



$$\varepsilon_m \le \frac{1}{2} - \gamma_m$$

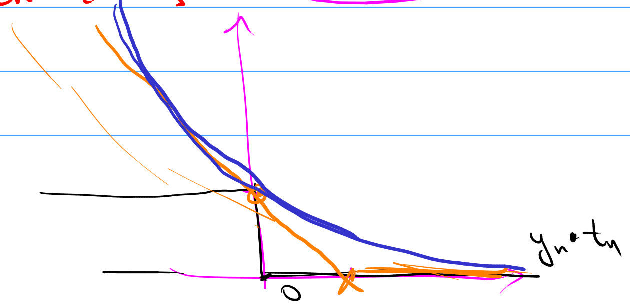
AdaBoost:

- $w_n^{(1)} = 1/N$ $\qquad D = \{(\bar{x}_n, t_n)\}_{n=1}^{N}$
- $\forall m = 1, \dots, M$:
  - → обучаем $y_m(\bar{x})$: $\quad J_m = \sum_{n=1}^{N} w_n^{(m)} \cdot [y_m(\bar{x}_n) \ne t_n]$
- $\alpha_m = \ln\left(\frac{1-\varepsilon_m}{\varepsilon_m}\right)$, $\quad \varepsilon_m = \frac{\sum_n w_n^{(m)} [y_m(\bar{x}_n) \ne t_n]}{\sum_n w_n^{(m)}}$
- $w_n^{(m+1)} = w_n^{(m)} \cdot \underbrace{e^{\alpha_m [y_m(\bar{x}_n) \ne t_n]}}$
- $y(\bar{x}) = \mathrm{sign}\left(\sum_m \alpha_m y_m(\bar{x})\right)$

$$\varepsilon \le \prod_m 2\sqrt{\varepsilon_m(1-\varepsilon_m)} \le e^{-2\sum \gamma_m^2}$$

$t_n \in \{\pm 1\}$

$$E = \sum_{n=1}^{N} e^{-\frac{1}{2} t_n f(\bar{x}_n)}$$



$$F_m(\bar{x}) = \sum_{l=1}^{m} \alpha_l y_l(\bar{x}) =$$

$$= \underbrace{F_{m-1}(\bar{x})} + \underbrace{\alpha_m y_m(\bar{x})}$$

$$E[F_m(\bar{x})] \xrightarrow[\alpha_m,\, y_m(\bar{x})]{} \min$$

$$E[F_m] = \sum_{n=1}^{N} e^{-\frac{1}{2} t_n (F_{m-1}(\bar{x}) + \alpha_m y_m(\bar{x}))} =$$

$$= \sum_{n=1}^{N} e^{-\frac{1}{2} t_n F_{m-1}(\bar{x})} \cdot \underbrace{e^{-\frac{1}{2} t_n \alpha_m y_m(\bar{x}_n)}}$$

$$-\frac{1}{2} \alpha_m t_n y_m(\bar{x}_n) = \frac{\alpha_m}{2} \cdot [t_n \neq y_m(\bar{x}_n)] -$$

$$- \frac{\alpha_m}{2} \underbrace{[t_n = y_m(\bar{x}_n)]} =$$

$$- \frac{\alpha_m}{2} \left(1 - [t_n \neq y_m(\bar{x}_n)] \right)$$

$$= \alpha_m [t_n \neq y_m(\bar{x}_n)] - \frac{\alpha_m}{2}$$

$$e^{-\frac{1}{2} \alpha_m t_n y_m(\bar{x}_n)} = e^{-\frac{\alpha_m}{2}} \cdot e^{\alpha_m [t_n \neq y_m(\bar{x}_n)]}$$

$$E[F_m] = \sum_{n=1}^{N} \underbrace{e^{-\frac{1}{2} t_n F_{m-1}(\bar{x})}} \cdot e^{-\frac{\alpha_m}{2}} \cdot \underbrace{e^{\alpha_m [t_n \neq y_m(\bar{x}_n)]}} =$$

$$\sum_{n:\, y_m(\bar{x}_n) = t_n} e^{-\frac{1}{2} t_n F_{m-1}(\bar{x})} e^{-\frac{\alpha_m}{2}} + \sum_{n:\text{ wrong}} e^{-\frac{1}{2} t_n F_{m-1}} \underbrace{e^{\frac{\alpha_m}{2}}}$$

$$e^{-\frac{\alpha_m}{2}} + \left(e^{\frac{\alpha_m}{2}} - e^{-\frac{\alpha_m}{2}} \right)$$

$$= \sum_n \overbrace{e^{-\frac{1}{2} t_n F_{m-1}(\bar{x}_n)}}^{a_n} e^{-\frac{\alpha_m}{2}} +$$

$$+ \left(e^{\frac{\alpha_m}{2}} - e^{-\frac{\alpha_m}{2}}\right) \sum_{n:\text{wrong}} e^{-\frac{1}{2} t_n F_{m-1}(\bar{x}_n)}$$

$$\sum_n a_n \cdot \begin{cases} e^{-\frac{\alpha_m}{2}}, & \text{correct} \\ e^{\alpha_m/2}, & \text{wrong} \end{cases}$$

$$\sum_{\text{correct}} a_n e^{-\frac{\alpha_m}{2}} + \sum_{\text{wrong}} a_n \cdot e^{-\frac{\alpha_m}{2}} + \sum_{\text{wrong}} a_n \cdot \left(e^{\frac{\alpha_m}{2}} - e^{-\frac{\alpha_m}{2}}\right)$$

$$E[F_m] = \text{const} + \text{const} \cdot \sum_{n: t_n \neq y_m(\bar{x}_n)} \underbrace{e^{-\frac{1}{2} t_n F_{m-1}(\bar{x}_n)}}_{= \, \sigma_n^{(m)}}$$

$$y_m(\bar{x}_n): \quad \sum_n \sigma_n^{(m)} [t_n \neq y_m(\bar{x}_n)] \to \min$$

$$\alpha_m: \quad \ldots \xrightarrow[\alpha_m]{} \min$$

$$\sigma_n^{(m+1)} = e^{-\frac{1}{2} t_n F_m(\bar{x}_n)} = e^{-\frac{1}{2} t_n F_{m-1}(\bar{x}_n)} \cdot e^{-\frac{1}{2} t_n \alpha_m y_m(\bar{x}_n)}$$

$$= \underbrace{e^{-\frac{1}{2} t_n F_{m-1}(\bar{x}_n)}}_{\sigma_n^{(m)}} \cdot e^{-\frac{\alpha_m}{2}} \cdot e^{\alpha_m [t_n \neq y_m(\bar{x}_n)]}$$