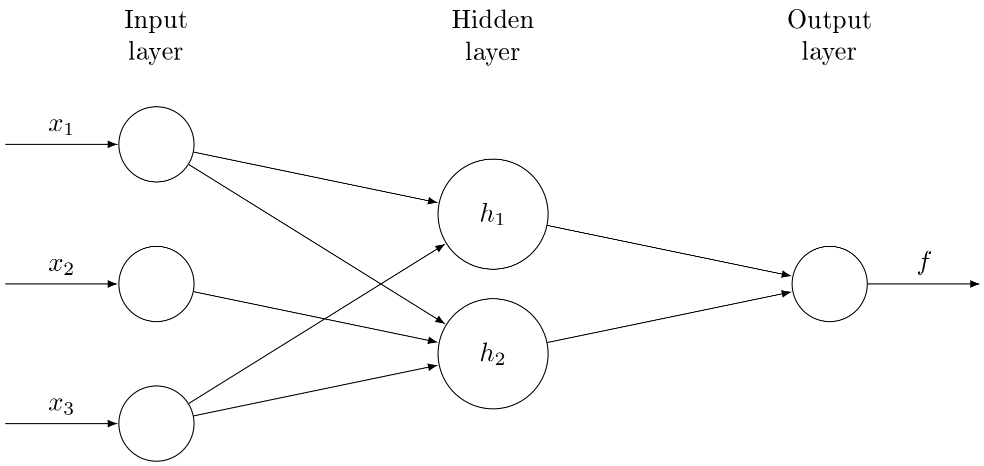
Задача 1.



x_1	x_2	x_3	$\overline{x}_1 \vee x_3$	$x_3 - x_1$	$x_1 \vee x_2 \vee x_3$	$x_1 + x_2 + x_3$
0	0	0	1	0	0	0
0	0	1	1	1	1	1
0	1	0	1	0	1	1
0	1	1	1	1	1	2
1	0	0	0	-1	1	1
1	0	1	1	0	1	2
1	1	0	0	-1	1	2
1	1	1	1	0	1	3

Таким образом, h_1 можно представить как $[x_3 - x_1 + \frac{1}{2} > 0]$, а h_2 - как $[x_1 + x_2 + x_3 - \frac{1}{2} > 0]$

h_1	h_2	$h_1 \& h_2$	$h_1 + h_2$
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	2

Таким образом, f можно представить как $[h_1 + h_2 - \frac{3}{2} > 0]$

Задача 2.

$L = -\sum_{j=1}^{m} y_j \log p_j = -\sum_{j=1}^{m} y_j \log \frac{e^{a_j}}{\sum_{k=1}^{m} e^{a_k}} =$
$= -\sum_{j=1}^{m} y_j (\log e^{a_j} - \log \sum_{k=1}^{m} e^{a_k}) = -\sum_{j=1}^{m} y_j (a_j - \log \sum_{k=1}^{m} e^{a_k}) =$
$= -\sum_{j=1}^{m} y_j a_j + \sum_{j=1}^{m} y_j \log \sum_{k=1}^{m} e^{a_k} = \log(\sum_{k=1}^{m} e^{a_k}) \sum_{j=1}^{m} y_j - \sum_{j=1}^{m} y_j a_j =$
$= \log \sum_{k=1}^{m} e^{a_k} - \sum_{j=1}^{m} y_j a_j$
$W = W - \eta L'_W$
$W' = W' - \eta L'_{W'}$
$\frac{\partial L}{\partial W'_{jk}} = \sum_{i=1}^{m} \frac{\partial L}{\partial a_i} \cdot \frac{\partial a_i}{\partial W'_{jk}} = \sum_{i=1}^{m} \left(\frac{e^{a_i}}{\sum_{k=1}^{m} e^{a_k}} - y_i \right) \delta_{ik} h_j = \sum_{i=1}^{m} (p_i - y_i) \delta_{ik} h_j = (p_k - y_k) h_j$
$\frac{\partial L}{\partial W_{jk}} = \sum_{i=1}^{d} \frac{\partial L}{\partial h_i} \cdot \frac{\partial h_i}{\partial W_{jk}} = \sum_{i=1}^{d} \left(\sum_{l=1}^{m} \frac{\partial L}{\partial a_l} \cdot \frac{\partial a_l}{\partial h_i} \right) \delta_{ik} x_j = \sum_{i=1}^{d} \left(\sum_{l=1}^{m} (p_l - y_l) W'_{il} \right) \delta_{ik} x_j = \sum_{l=1}^{m} (p_l - y_l) W'_{kl} x_j$
$W_{jk} = W_{jk} - \eta \sum_{l=1}^{m} (p_l - y_l) W'_{kl} x_j$
$W'_{jk} = W'_{jk} - \eta (p_k - y_k) h_j$

Задача 4.

$\sum_{i=1}^{\ell} \tilde{w}_i^{(n+1)} [b_n(x_i) \neq y_i] = \frac{1}{2}$
$\tilde{w}_i^{(n+1)} = \frac{\tilde{w}_i^n e^{-\alpha_n b_n y_i}}{w_0}$
$\sum_{i=1}^{\ell} \tilde{w}_i^{(n+1)} [b_n(x_i) \neq y_i] = \frac{1}{w_0} \sum_{i=1}^{\ell} \tilde{w}_i^n e^{-\alpha_n b_n y_i} [b_n(x_i) \neq y_i] =$
$\frac{1}{w_0} \sum_{i=1}^{\ell} \tilde{w}_i^n e^{\alpha_n} [b_n(x_i) \neq y_i] =$
$\alpha_n = \frac{1}{2} \ln \frac{P}{N}$, где $P = \sum_{i=1}^{\ell} \tilde{w}_i^n [b_n(x_i) = y_i]$, $N = \sum_{i=1}^{\ell} \tilde{w}_i^n [b_n(x_i) \neq y_i]$
$= \frac{1}{w_0} \sum_{i=1}^{\ell} \tilde{w}_i^n \sqrt{\frac{P}{N}} [b_n(x_i) \neq y_i] = \frac{1}{w_0} \sqrt{\frac{P}{N}} \sum_{i=1}^{\ell} \tilde{w}_i^n [b_n(x_i) \neq y_i] = \frac{1}{w_0} \sqrt{\frac{P}{N}} N = \frac{\sqrt{PN}}{w_0}$

$w_0 = \sum_{i=1}^{\ell} \tilde{w}_i^n e^{-\alpha_n b_n y_i} = \sum_{i=1}^{\ell} \tilde{w}_i^n e^{-\alpha_n} [b_n(x_i) = y_i] + \sum_{i=1}^{\ell} \tilde{w}_i^n e^{\alpha_n} [b_n(x_i) \neq y_i] = \sqrt{\frac{N}{P}} P + \sqrt{\frac{P}{N}} N = 2\sqrt{NP}$
$\sum_{i=1}^{\ell} \tilde{w}_i^{(n+1)} [b_n(x_i) \neq y_i] = \frac{\sqrt{PN}}{2\sqrt{PN}} = \frac{1}{2}$
Задача 5.
1. Градиентный бустинг на каждой итерации настраивается на антиградиент, поэтому:
$-L' = y - \tilde{y}$
$L' = \tilde{y} - y$
$L = \frac{(\tilde{y}-y)^2}{2} + C$
2. $L = (\tilde{y} - y)^4$
$-L' = -4(\tilde{y} - y)^3$, $y = (6, 8, 6, 4, 1)$, $\tilde{y} = (5, 10, 6, 3, 0)$
$-L' = -4(-1, 2, 0, -1, -1)^3 = -4(-1, 8, 0, -1, -1) = (4, -32, 0, 4, 4)$
3. $L = -(y \log z + (1 - y) \log(1 - z))$
$-L' = -\frac{y}{z} + \frac{1-y}{1-z}$
$b_n = \underset{b}{\operatorname{argmin}} \sum_{i=1}^{\ell} (b(x_i) + L'(F_{n-1}(x_i), y_i))^2 = \underset{b}{\operatorname{argmin}} \sum_{i=1}^{\ell} (b(x_i) + \frac{y_i}{F_{n-1}(x_i)} - \frac{1-y_i}{1-F_{n-1}(x_i)})^2$
$\gamma_n = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{\ell} L(F_{n-1}(x_i) + \gamma b_n(x_i), y_i) =$
$= \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{\ell} (-y_i \log(F_{n-1}(x_i) + \gamma b_n(x_i)) - (1 - y_i) \log(1 - F_{n-1(x_i)} - \gamma b_n(x_i)))$
Задача 6.
1. $L(M) = e^{-M}$
$w_i = -L'(M_i) = -e^{-M_i}$
Возьмем большое отрицательное значение M_{noise}, тогда $\lim_{M_{\text{noise}} \to -\infty} e^{-M_{\text{noise}}} = \infty$
Возьмем близкое к 0 пороговое значение $M_{\text{threshold}}$, тогда $\lim_{M_{\text{threshold}} \to 0} e^{-M_{\text{threshold}}} = 1$
То есть $\forall C > 0 \exists M_{\text{noise}} \frac{w_{\text{nosie}}}{w_{\text{threshold}}} > C$
2. $L(M) = \log(1 + e^{-M})$
$w_i = -L'(M_i) = \frac{e^{-M_i}}{1+e^{-M_i}}$
$\lim_{M_{\text{noise}} \to -\infty} \frac{e^{-M_{\text{noise}}}}{1+e^{-M_{\text{noise}}}} = 1$
$\lim_{M_{\text{threshold}} \to 0} \frac{e^{-M_{\text{threshold}}}}{1+e^{-M_{\text{threshold}}}} = \frac{1}{2}$
Видно, что во втором случае веса будут лежать в между $\frac{1}{2}$ и 1, и бустинг будет устойчив к шуму