

Inicie $f_0(x)$ para ser uma constante, $f_0(x) = \operatorname{argmin}_p \sum_{i=1}^N L(y_i, p)$

Para $m = 1$ até M **faça**

Para $i = 1$ até n **faça**

Calcule o gradiente negativo

$$z_{im} = - \left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f=f_{m-1}}$$

fim

Ajuste uma árvore de regressão $g_m(x)$ para prever os destinos z_{im} das covariáveis x_i para todos os dados de treinamento.

Calcule um tamanho da etapa de descida de gradiente como:

$$\rho_m = \operatorname{argmin}_\rho \sum_{i=1}^n L(y_i, f_{m-1}(x_i) + \rho g_m(x_i))$$

Atualize o modelo com

$$f_m(x) = f_{m-1}(x) + \rho_m g_m(x)$$

fim

Saída: $f_M(x)$