Functional Gradient Descent

[Mason et al.][Friedman]

- want to minimize

$$\mathcal{L}(F) = \mathcal{L}(F(x_1), \ldots, F(x_m)) = \sum_i \exp(-y_i F(x_i))$$

- say have current estimate F and want to improve
- to do gradient descent, would like update

$$F \leftarrow F - \alpha \nabla_F \mathcal{L}(F)$$

- but update restricted in class of weak classifiers

$$F \leftarrow F + \alpha h_t$$

- so choose h_t "closest" to $-\nabla_F \mathcal{L}(F)$
- equivalent to AdaBoost