

**Algorithm** FORWARDFITTING

**Input**     $M$ , labeled training set  $\mathcal{D}$

**Initialize**    $f = 0$

**repeat**

**for**    $k = 1, 2, \dots M$

            fit  $k$ -th predictor  $\beta^k, b^k = \operatorname{argmin} \hat{L}(f + \beta b)$

            update  $f = f + b^k \beta^k$

**until**   change in  $\hat{L}$  small enough (or, change in  $b^k$  small enough)

**Output**    $f(x) = \sum_{k=1}^M \beta^k b^k(x)$