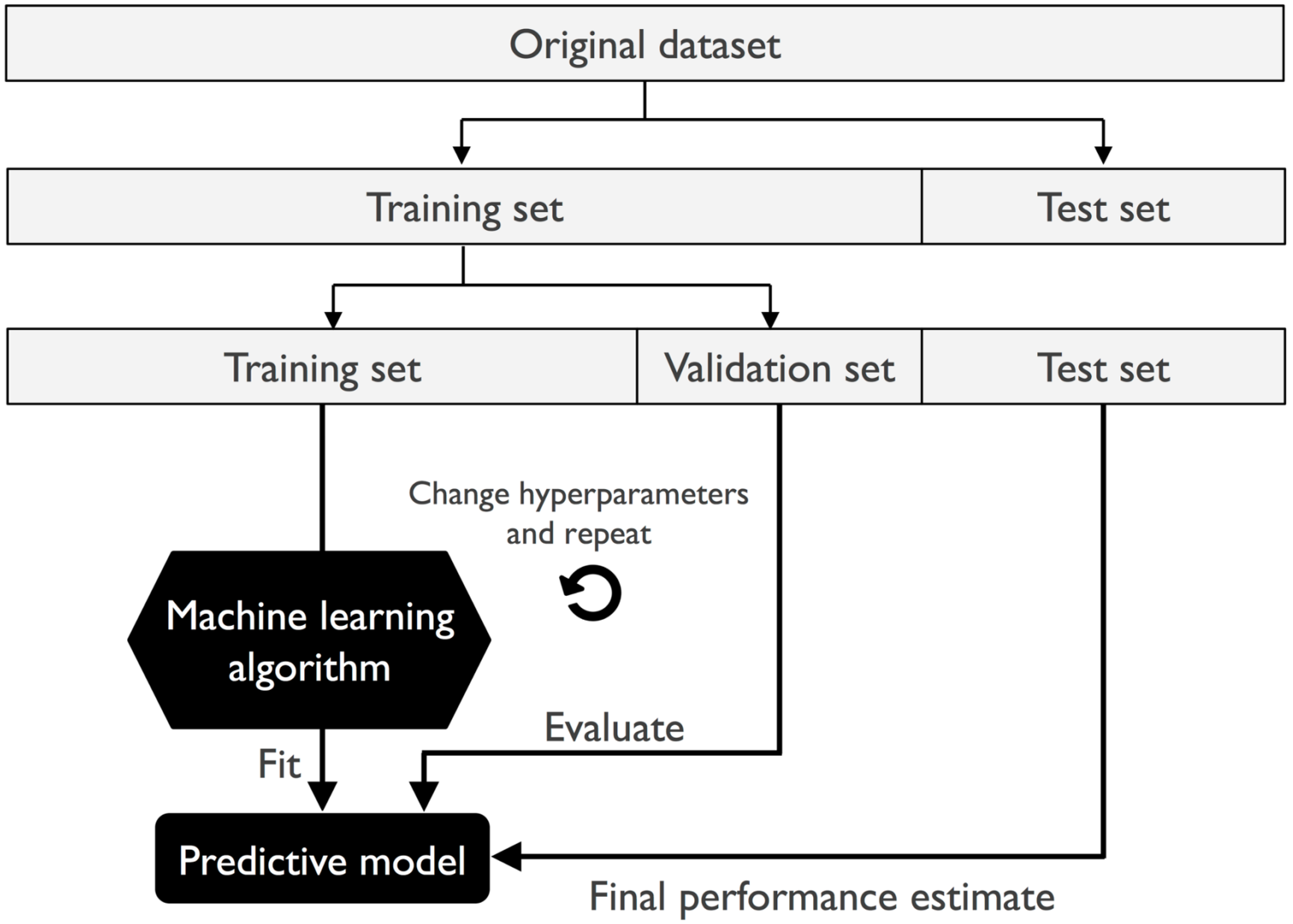
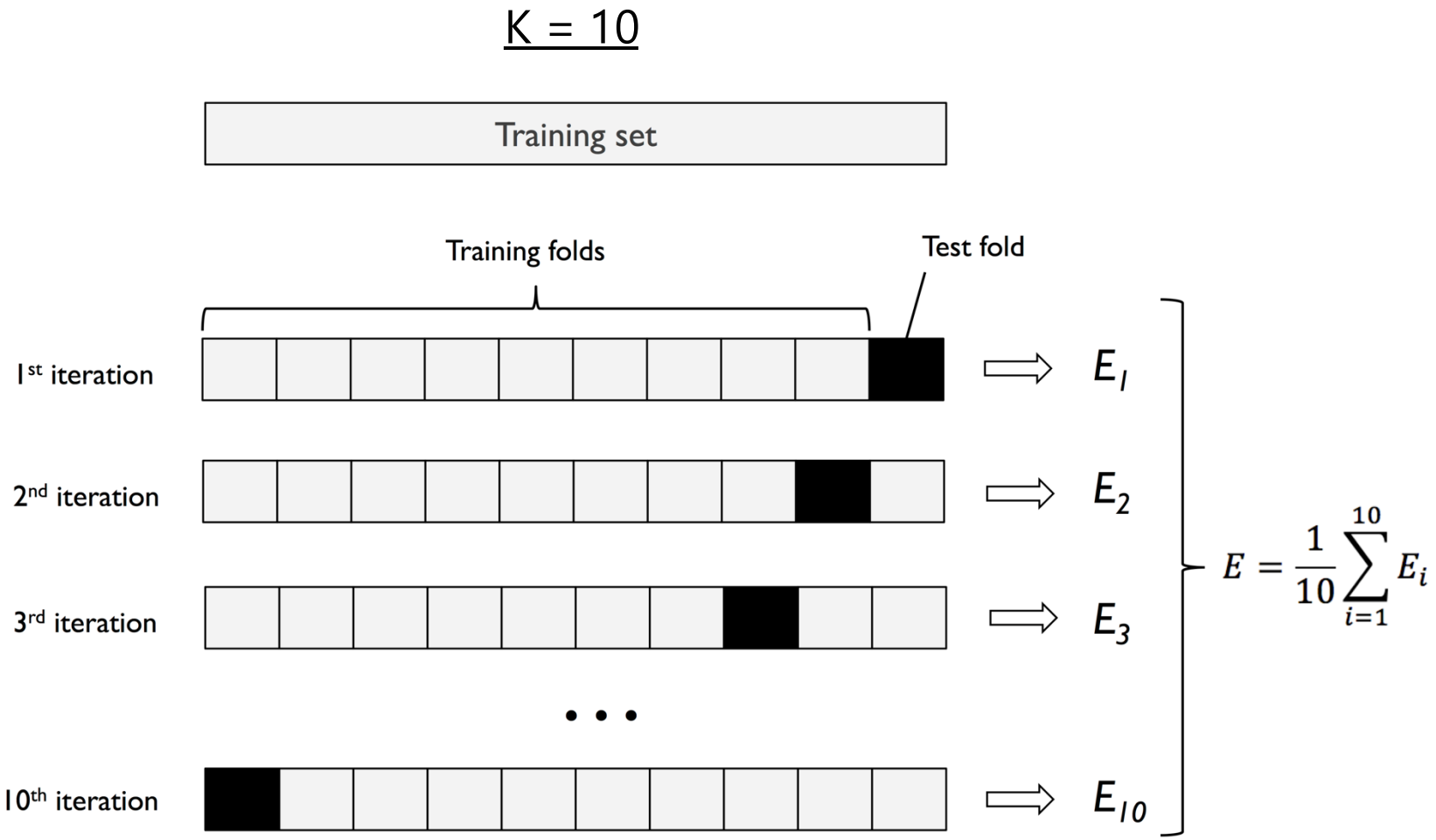
Parameters vs hyperparameters

CEMS
Chemical Engineering
& Materials Science

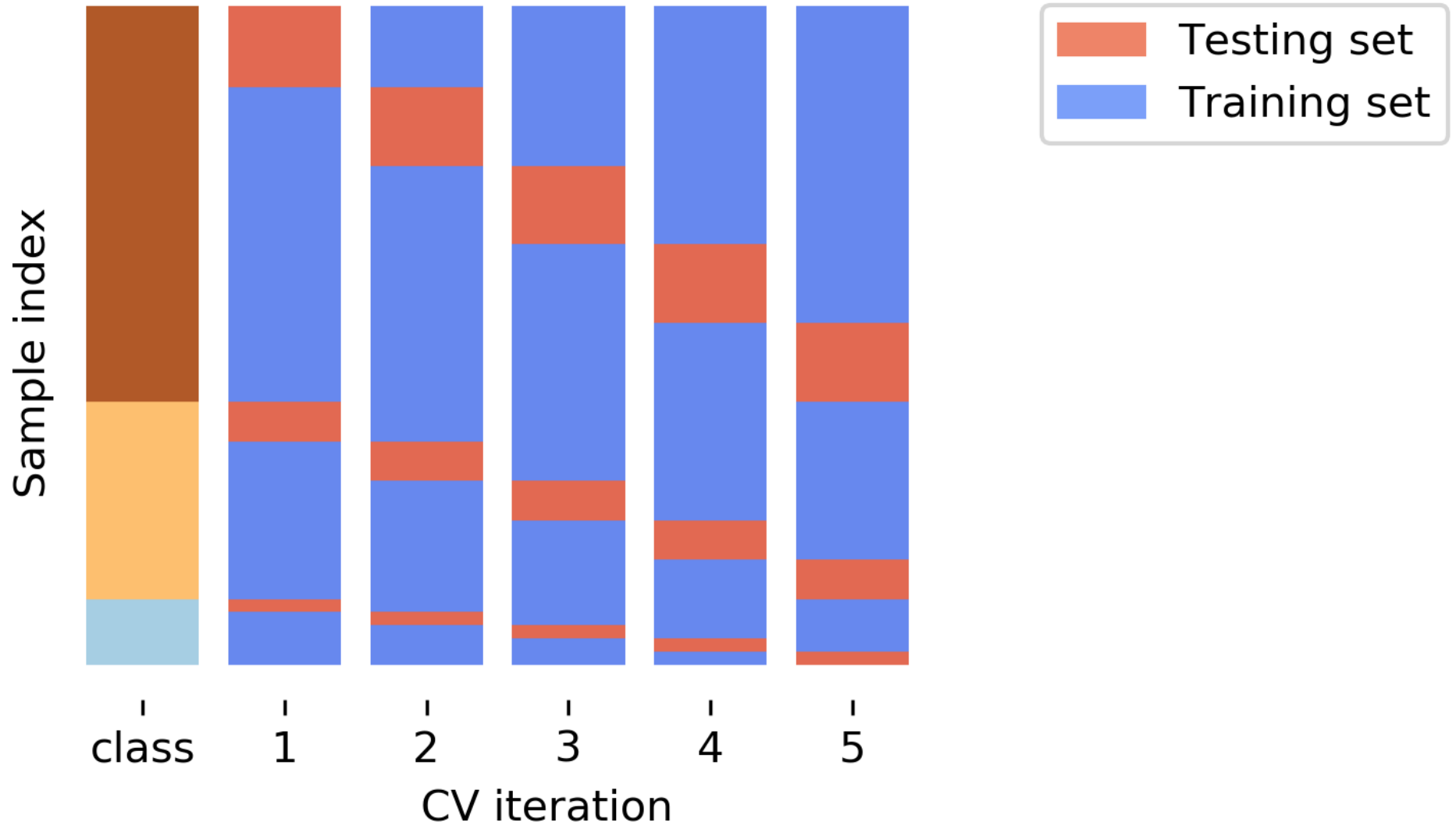
Validation and testing workflow



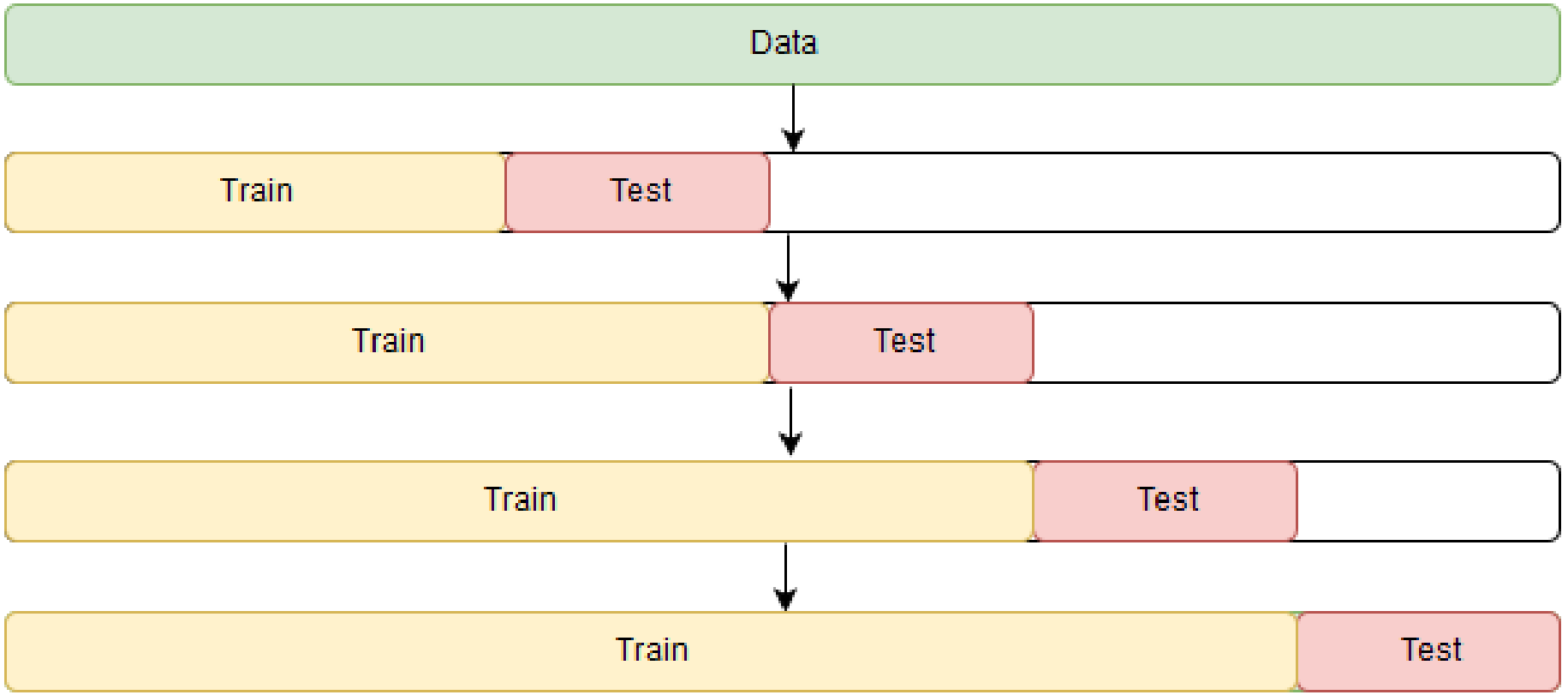
K-fold cross validation



Class imbalance: stratified cross validation



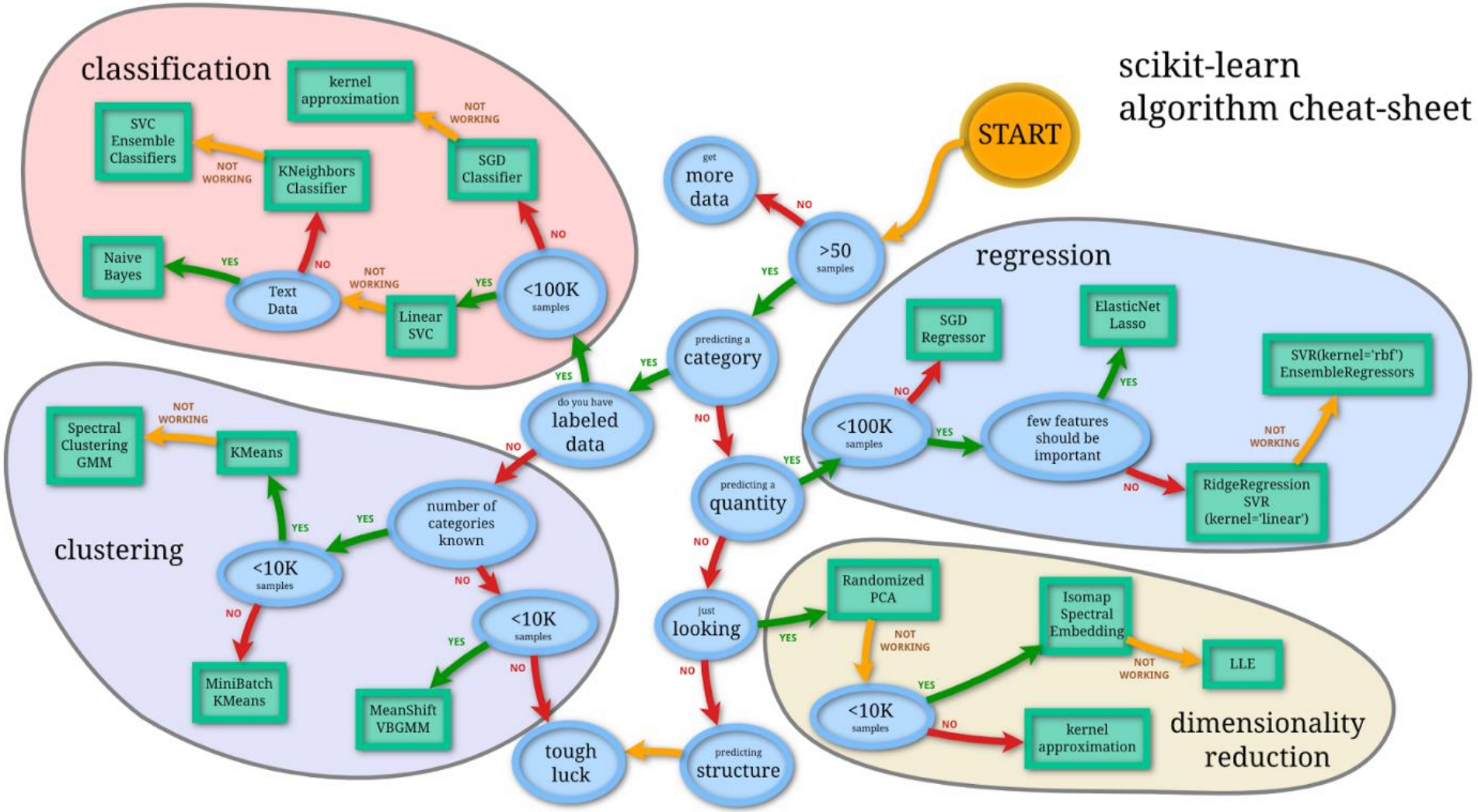
Time series: “Rolling validation”



Understanding data validation: IID vs OOD

CEMS
Chemical Engineering
& Materials Science

Which model to use??



Basics of linear regression

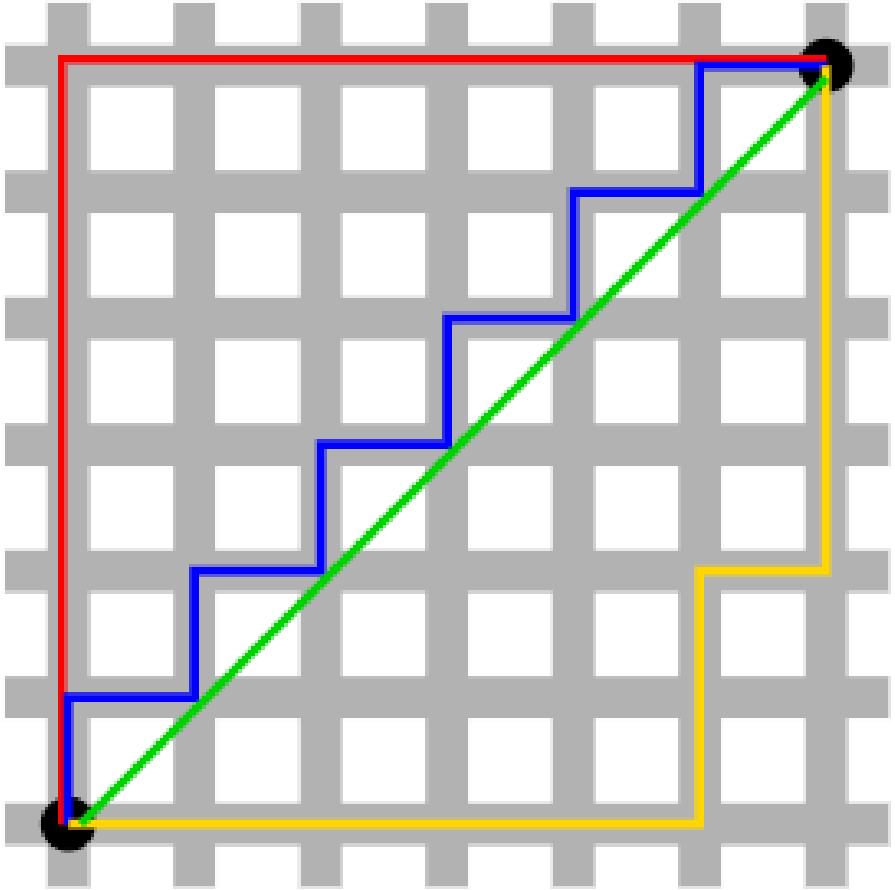
Basics of linear regression

L1 vs L2 norm

$$\hat{y}(w, x) = w_0 + w_1 x_1 + \ldots + w_p x_p$$

LASSO (L1 norm): $\min_w ||Xw - y||_2^2 + \alpha ||w||_1$

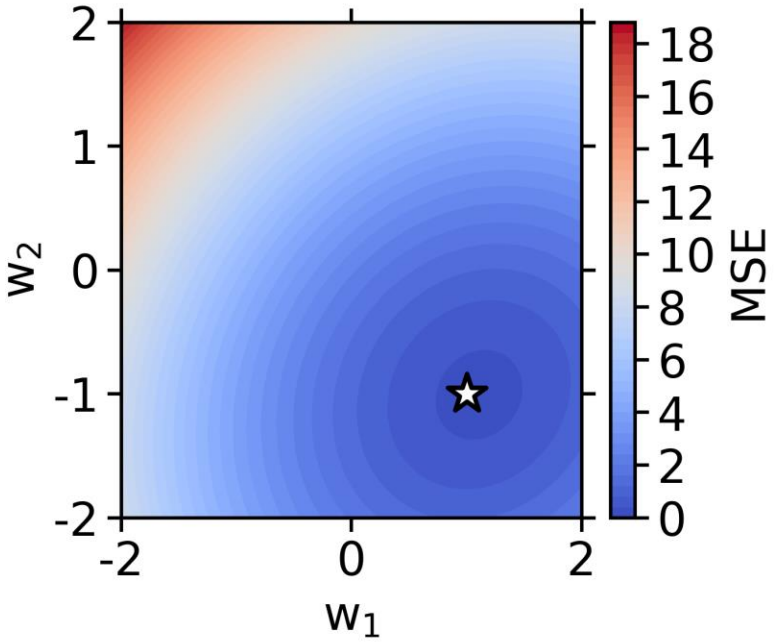
Ridge (L2 norm): $\min_w ||Xw - y||_2^2 + \alpha ||w||_2^2$

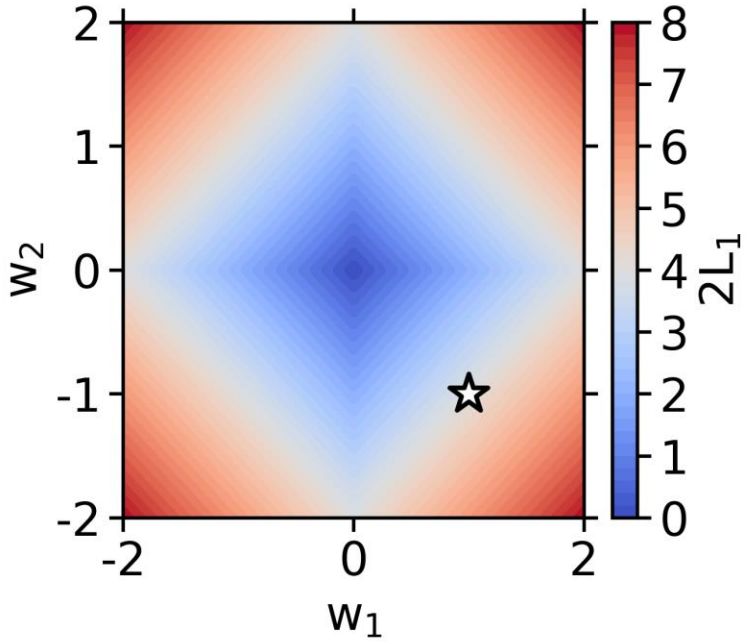


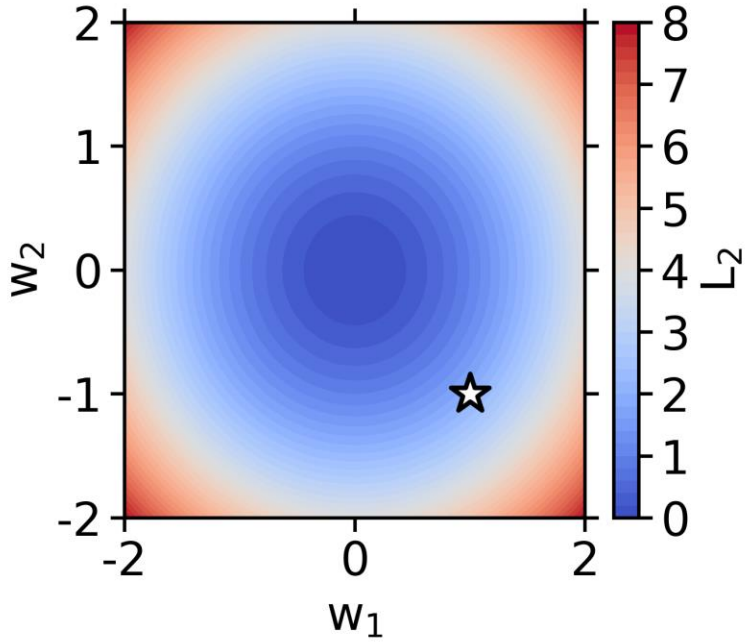
L1 vs L2 norm

$$y = 1.0x_1 - 1.0x_2 + N(0, 0.25)$$

$$\hat{y} = w_1 x_1 + w_2 x_2$$



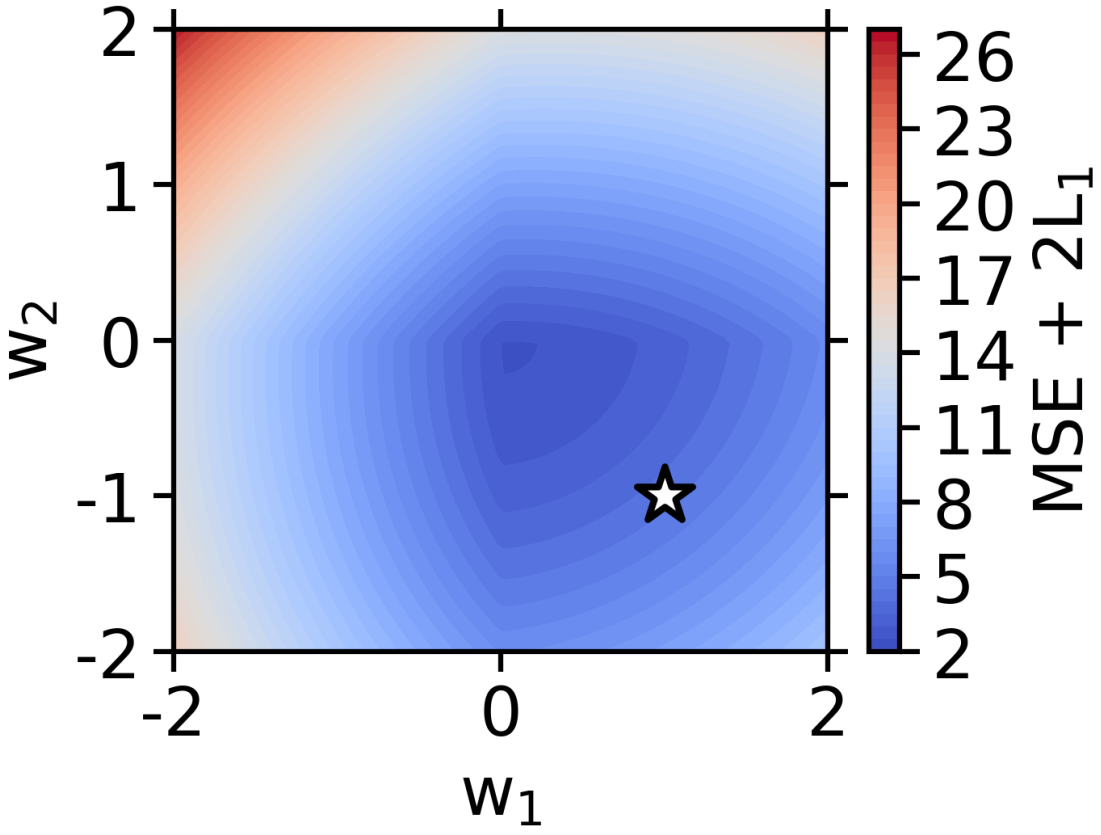


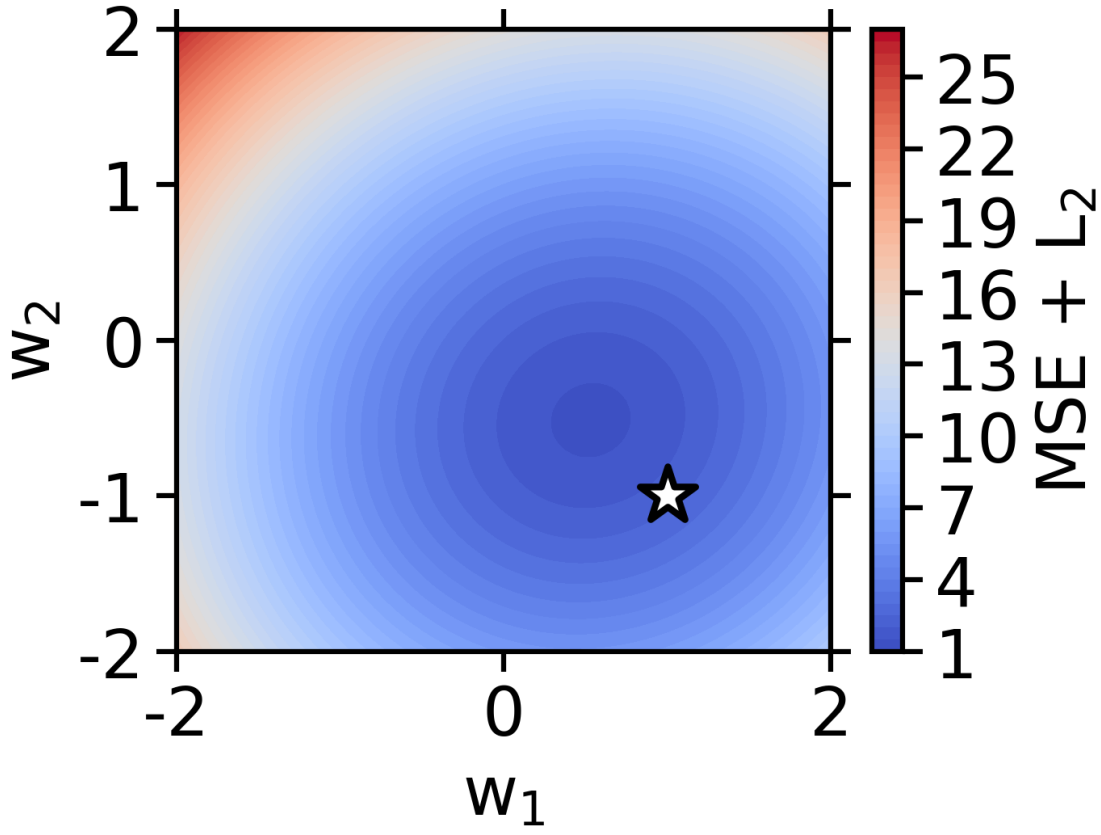


L1 vs L2 norm

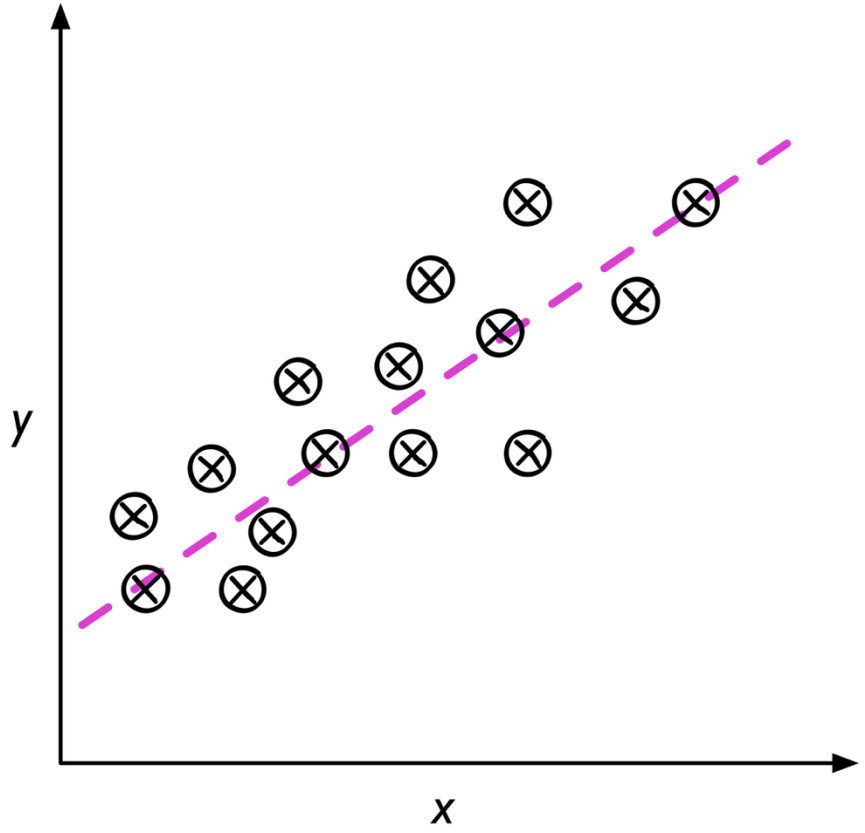
$$y = 1.0x_1 - 1.0x_2 + N(0, 0.25)$$

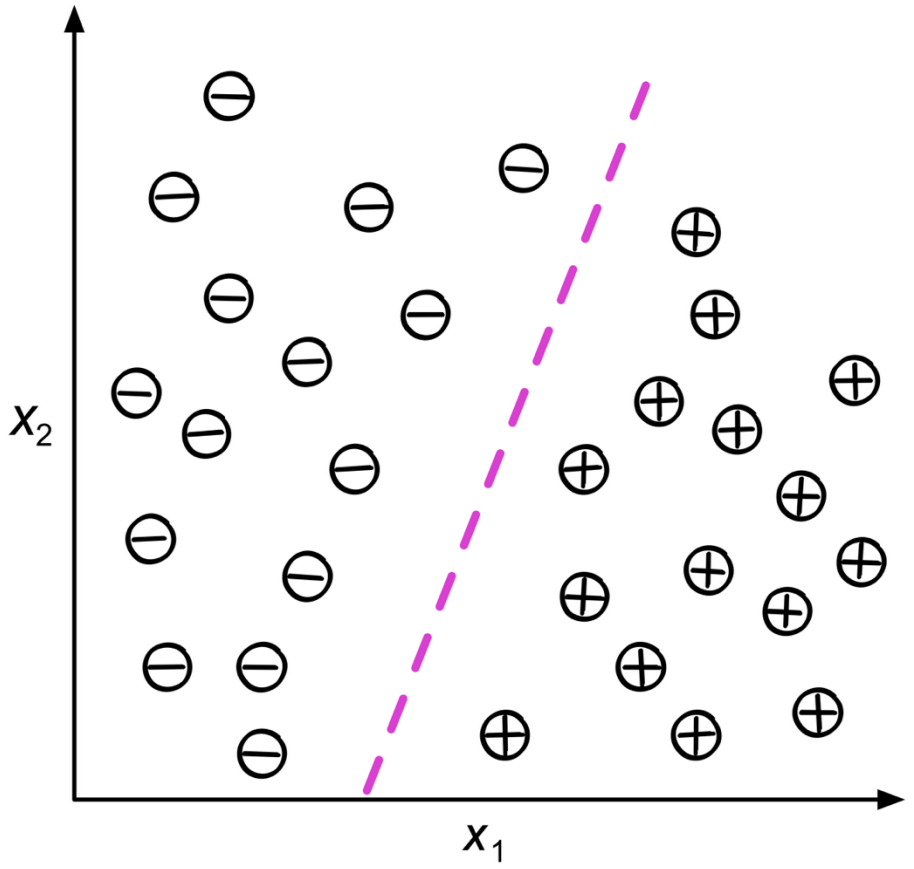
$$\hat{y} = w_1 x_1 + w_2 x_2$$





How does this look for classification?





How does this look for classification?

For binary classification, y is no longer continuous, but binomial:

y = [1, 1, 1, -1, -1, 1, -1, -1, …]