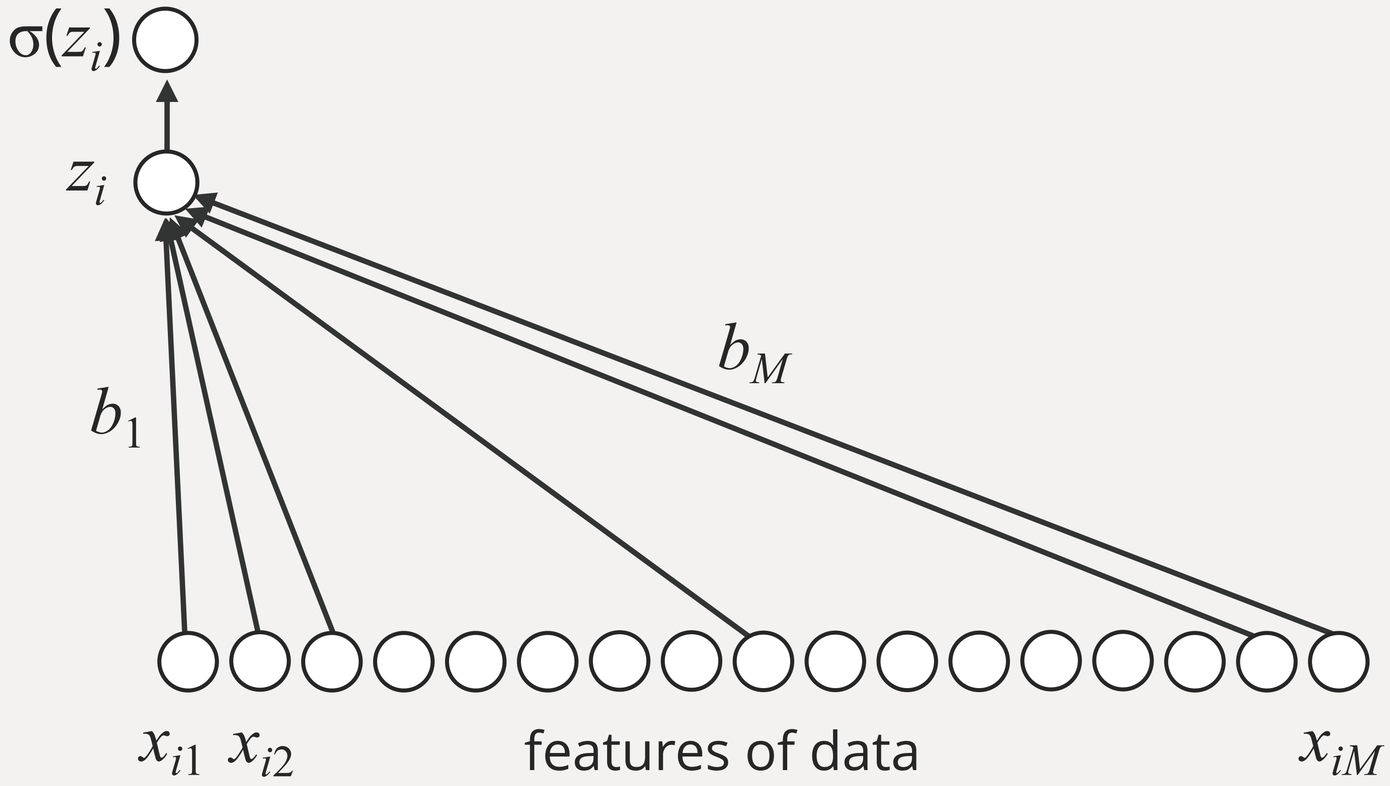
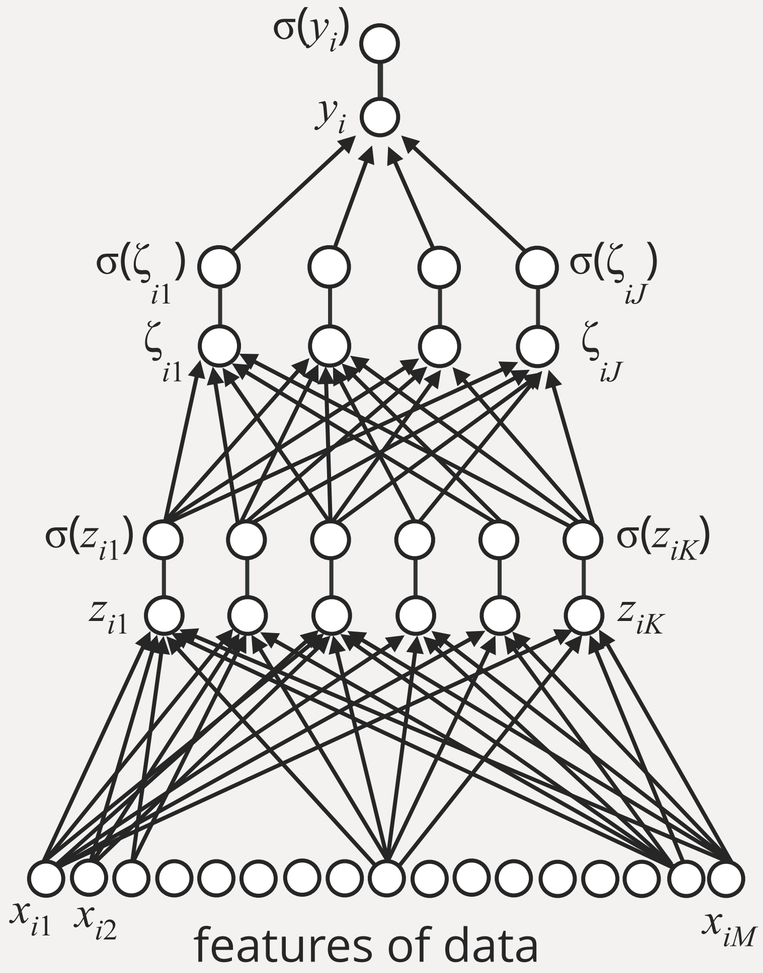
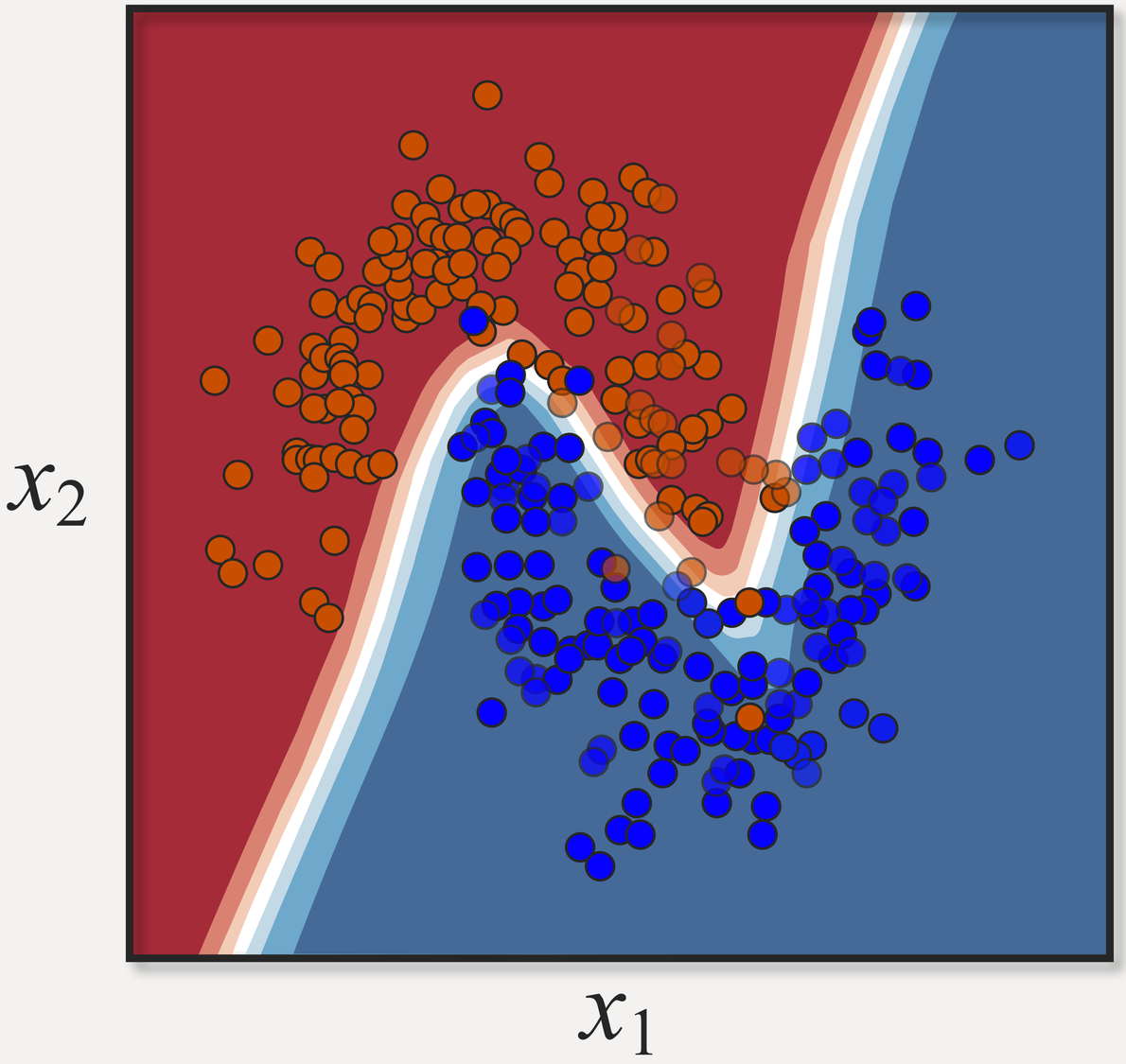
Logistic Regression



Multilayer Perceptron





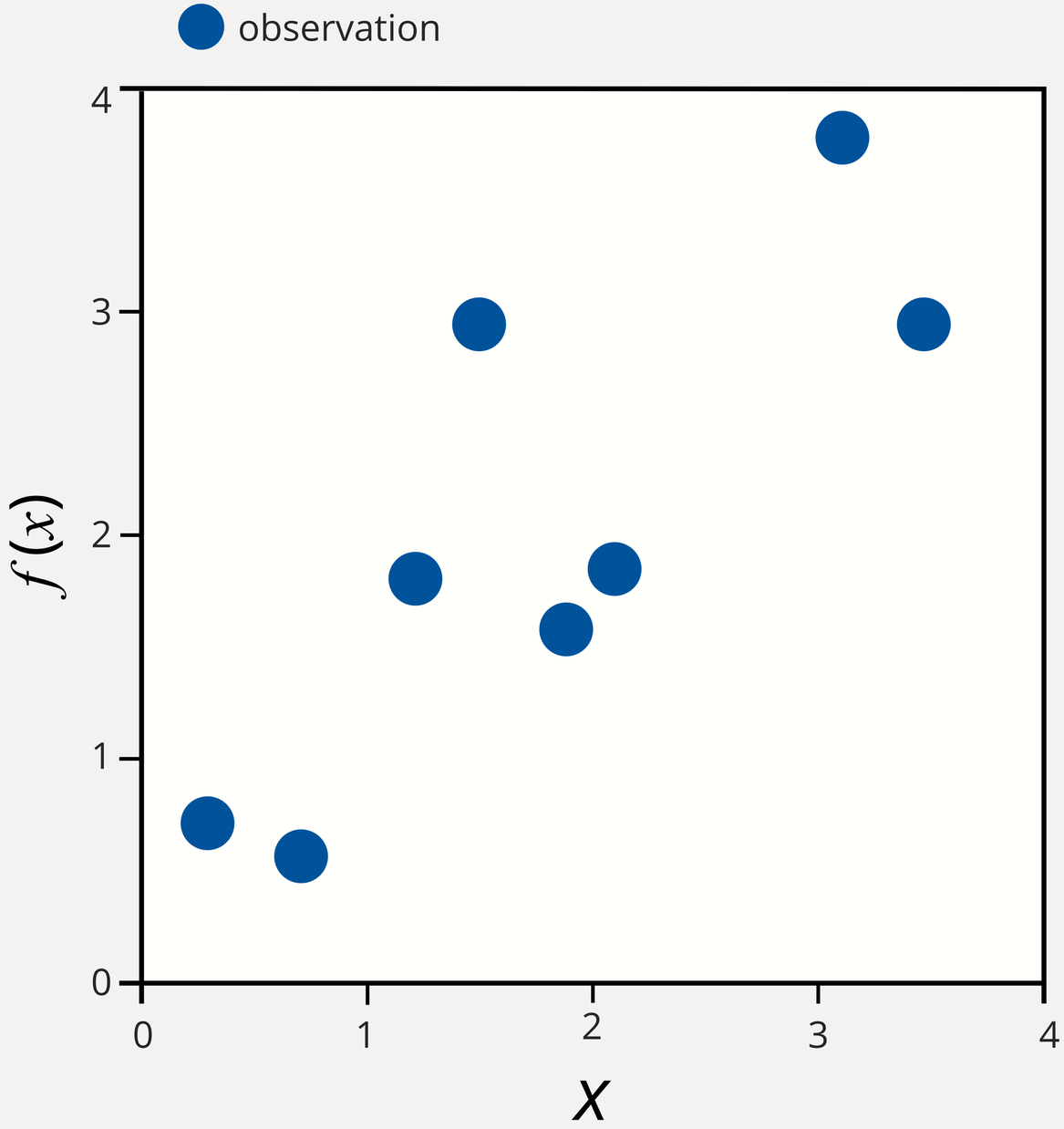
Complex Relationships Using Deep Learning

- Can be captured by using deep neural networks
- Can be represented accurately and predicted well
- Can give perfect performance in the training set
- Can perform poorly in the real world
- Needs to be validated

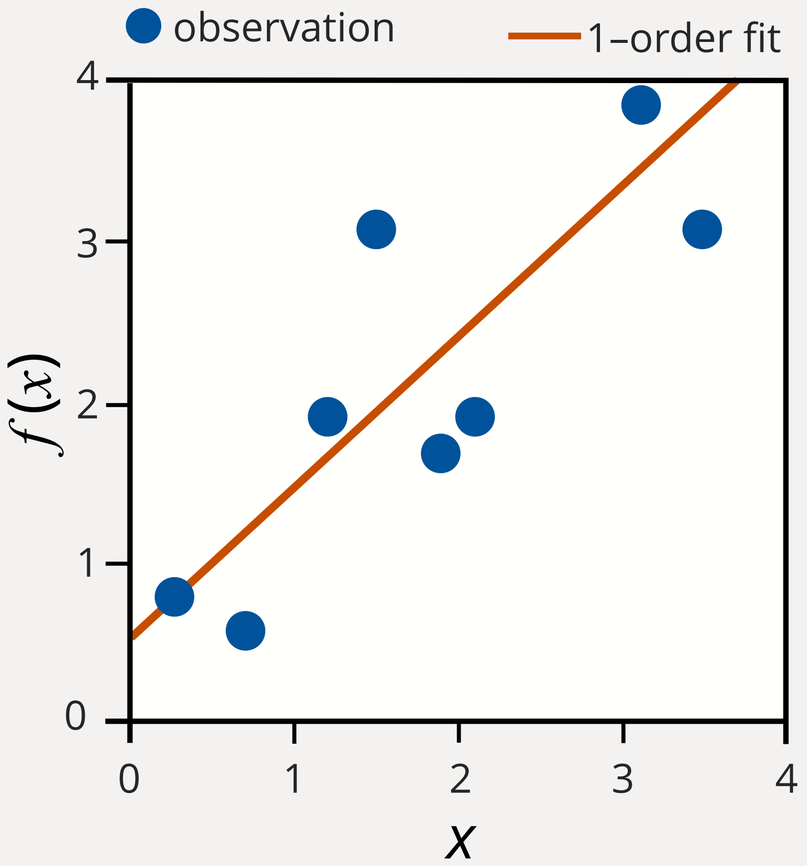
Overfitting is when the learned model increases complexity to fit the observed training data too well

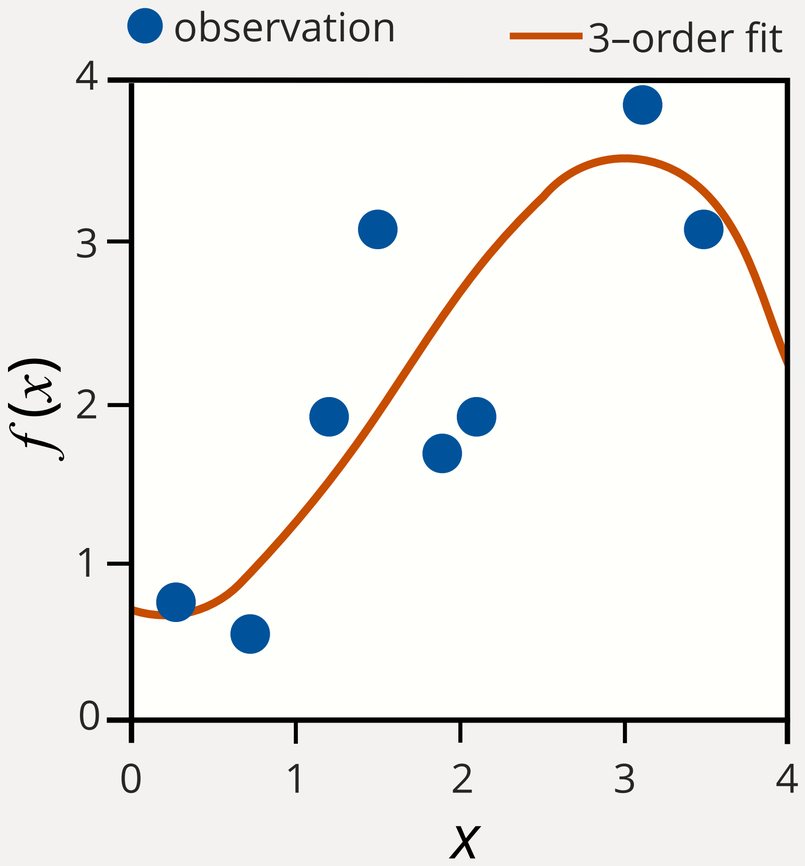
Will not work on future data in the real world

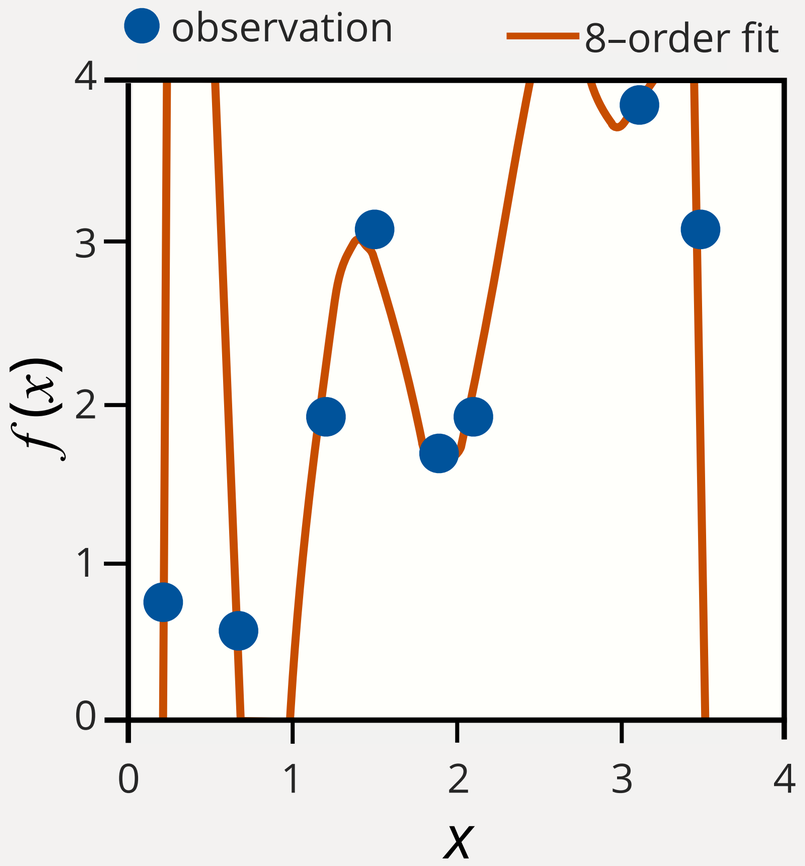
Want to come up with function to predict observation given *x*



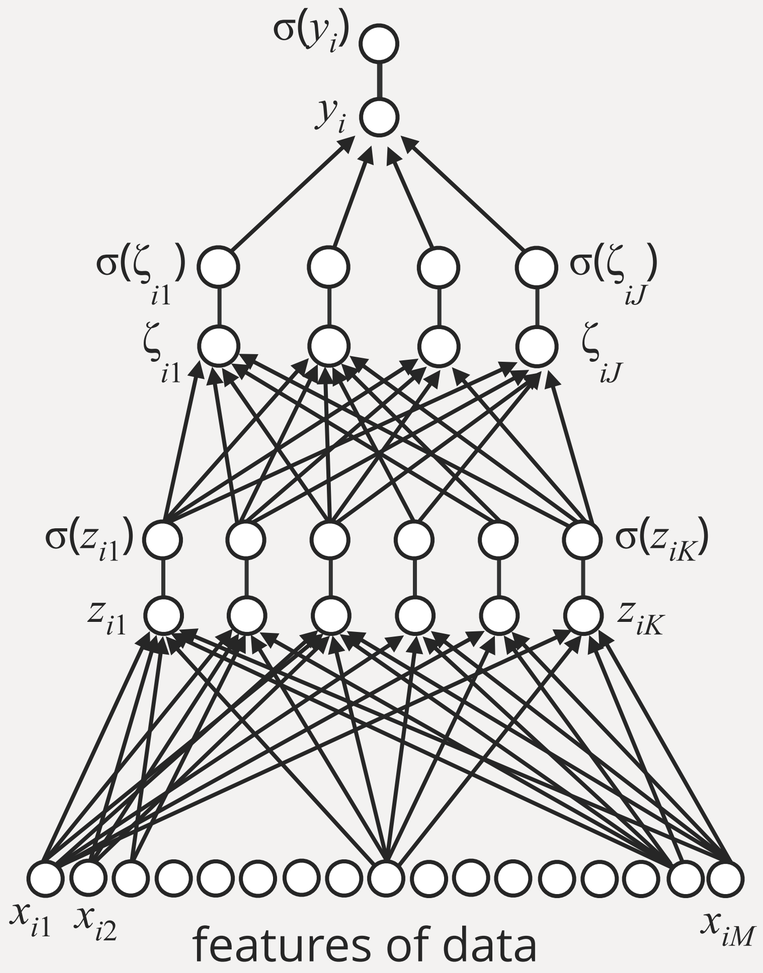
Increasing Polynomial Order





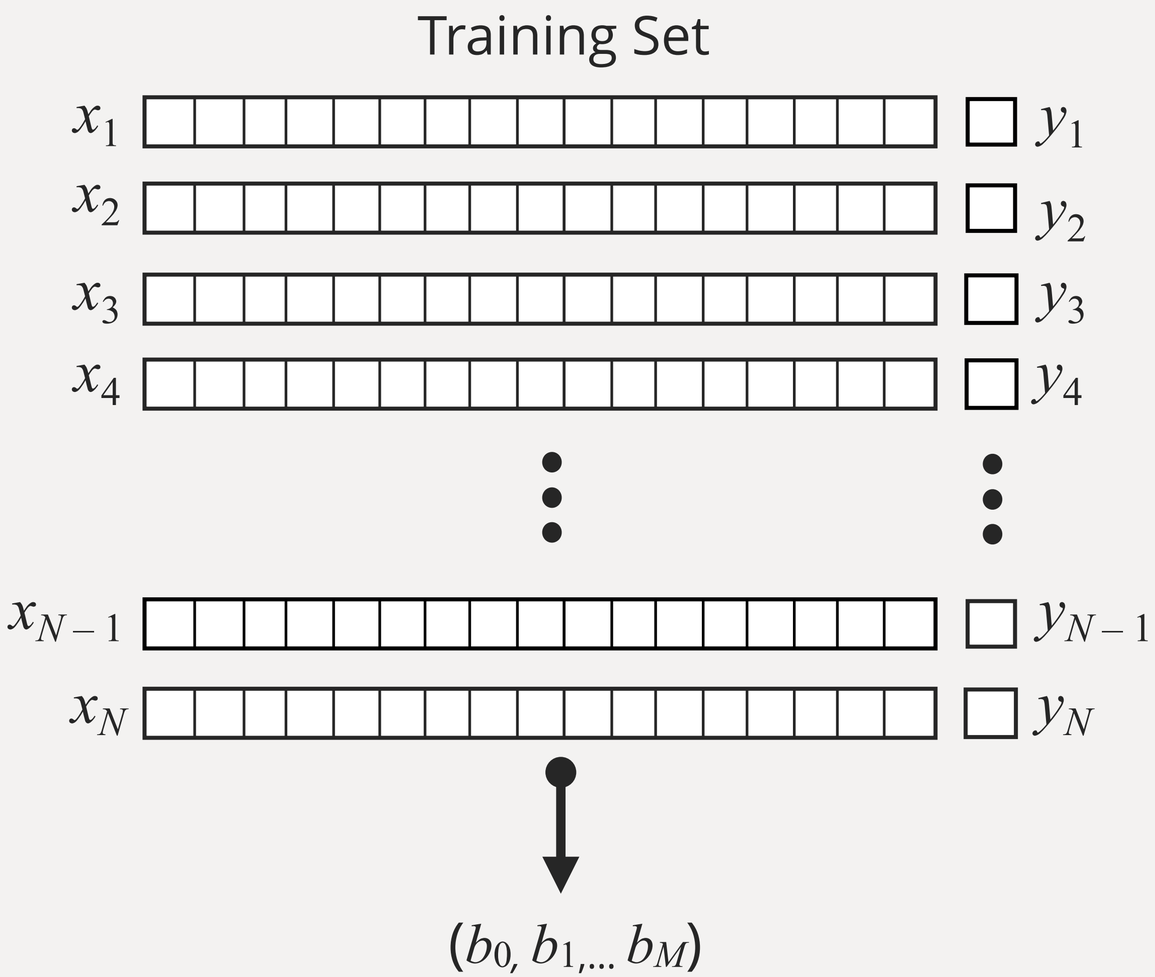


Multilayer Perceptron



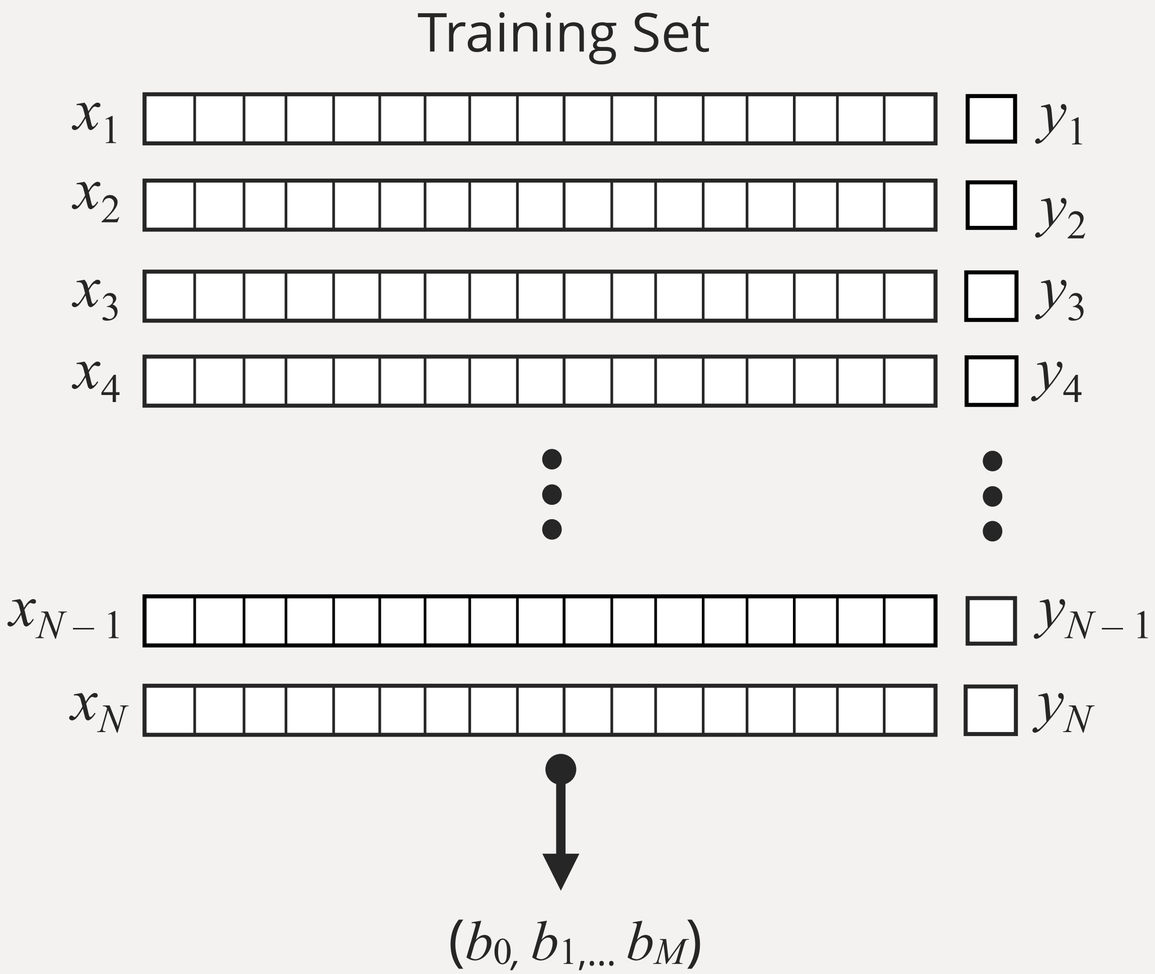
Problems with Overfitting

- Increasing parameters increases error rate
- Complex relationship may be too complex for reality



Problems with Overfitting

- Increasing parameters increases error rate
- Complex relationship may be too complex for reality
- Models and analysis are not generalized

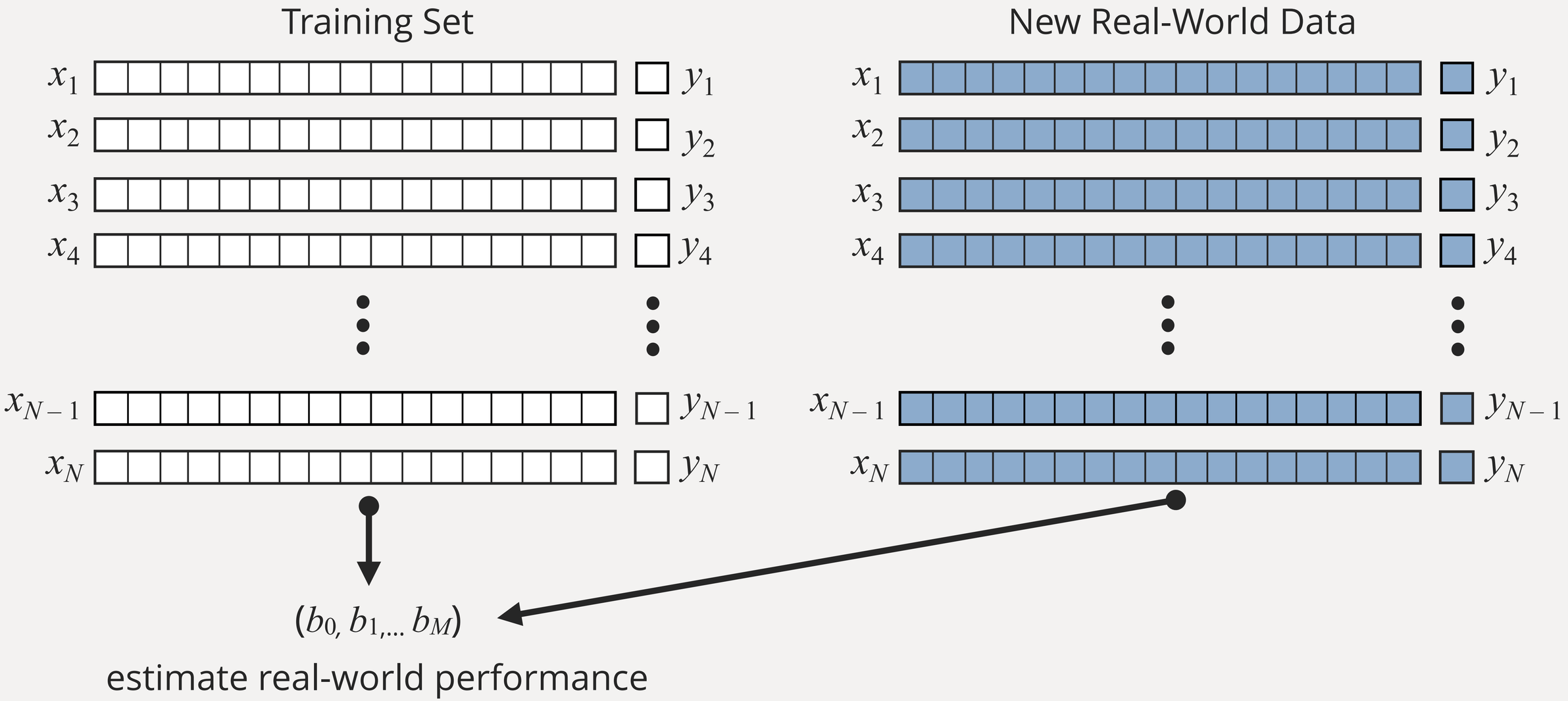


how well will this work in the real world?

Problems with Overfitting

- Increasing parameters increases error rate
- Complex relationship may be too complex for reality
- Models and analysis are not generalized

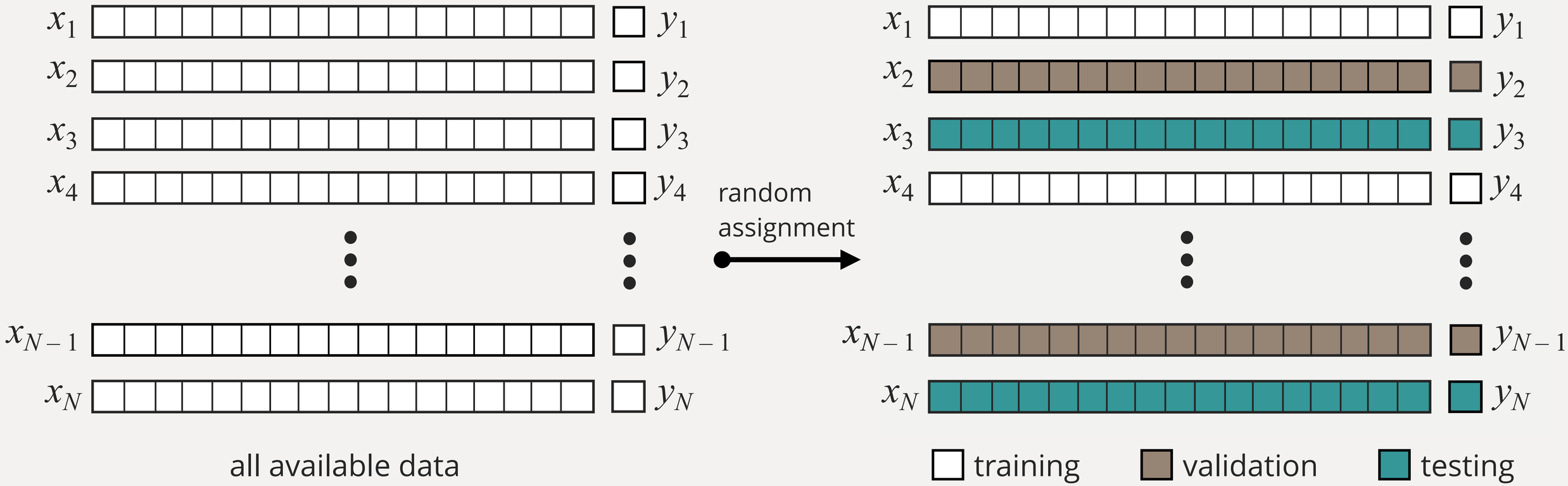
Standard Validation Strategy



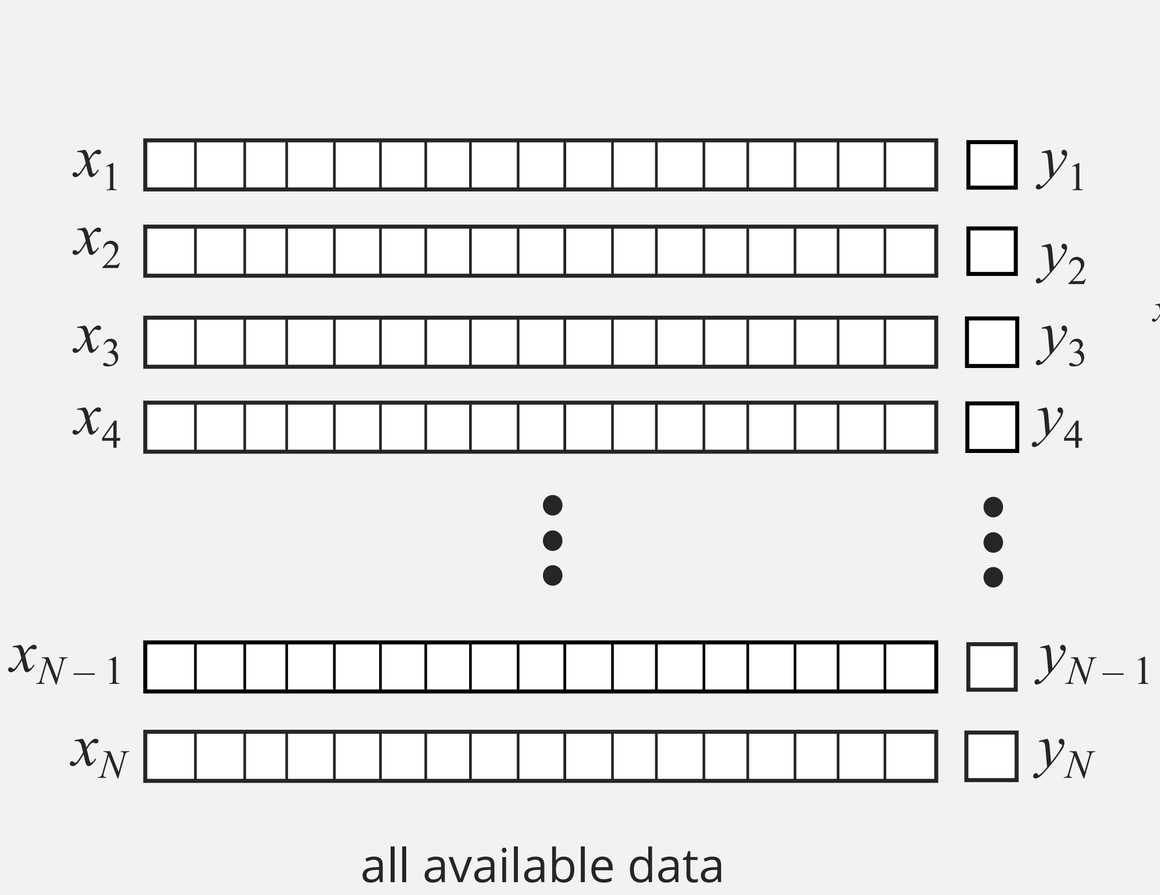
Standard Validation Strategy

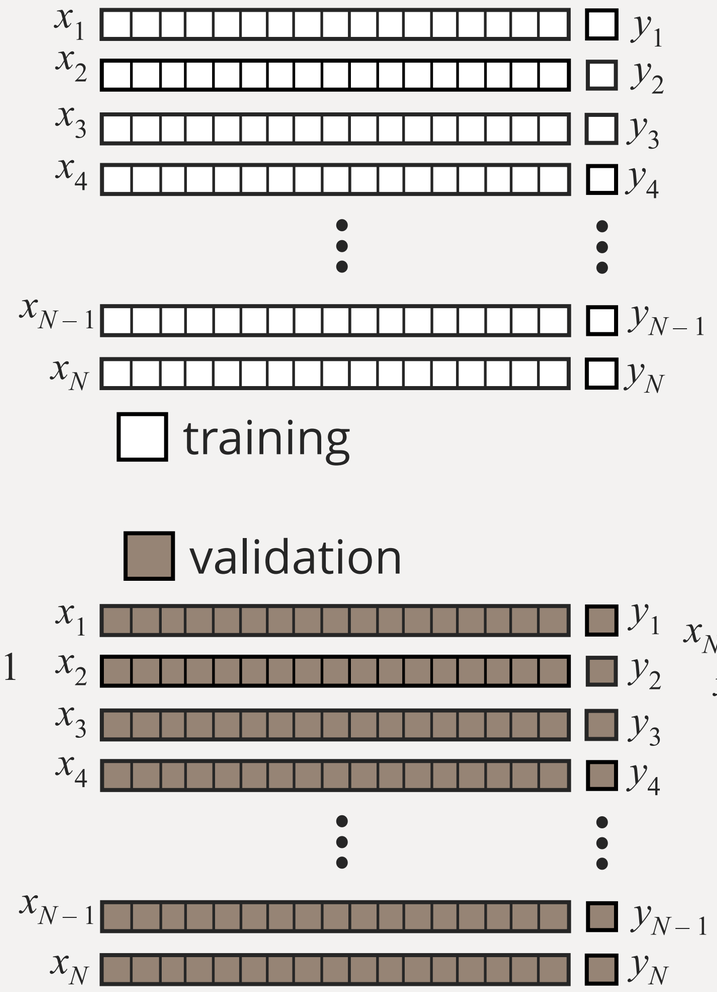
Is costly, can we use existing data to estimate performance?

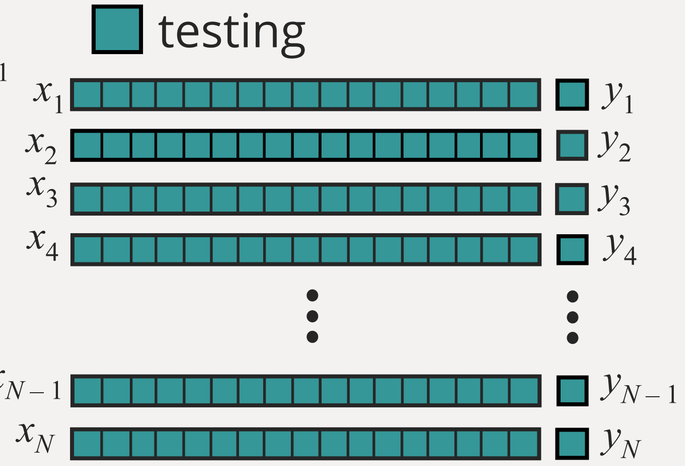
Split Data in Separate Groups

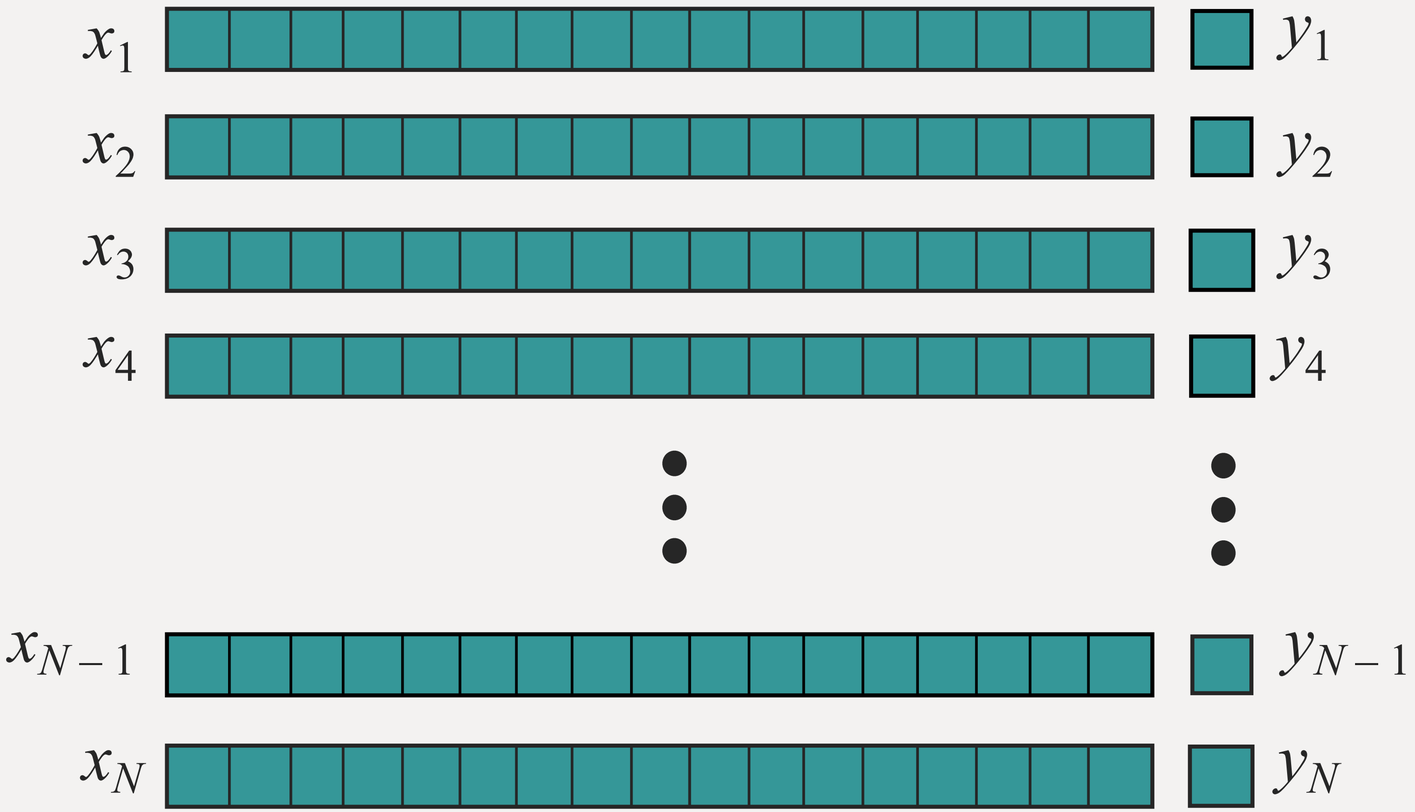


Split Data in Separate Groups



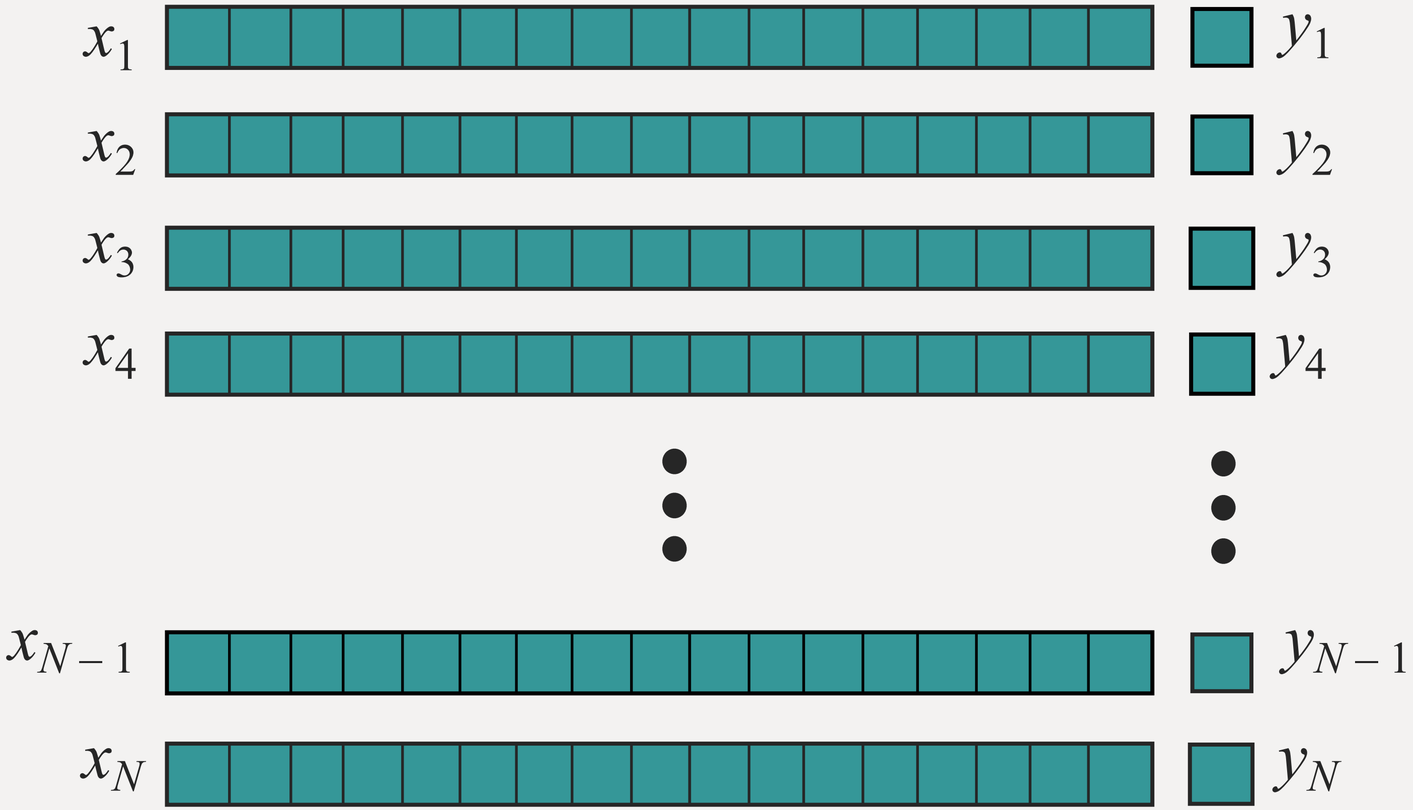






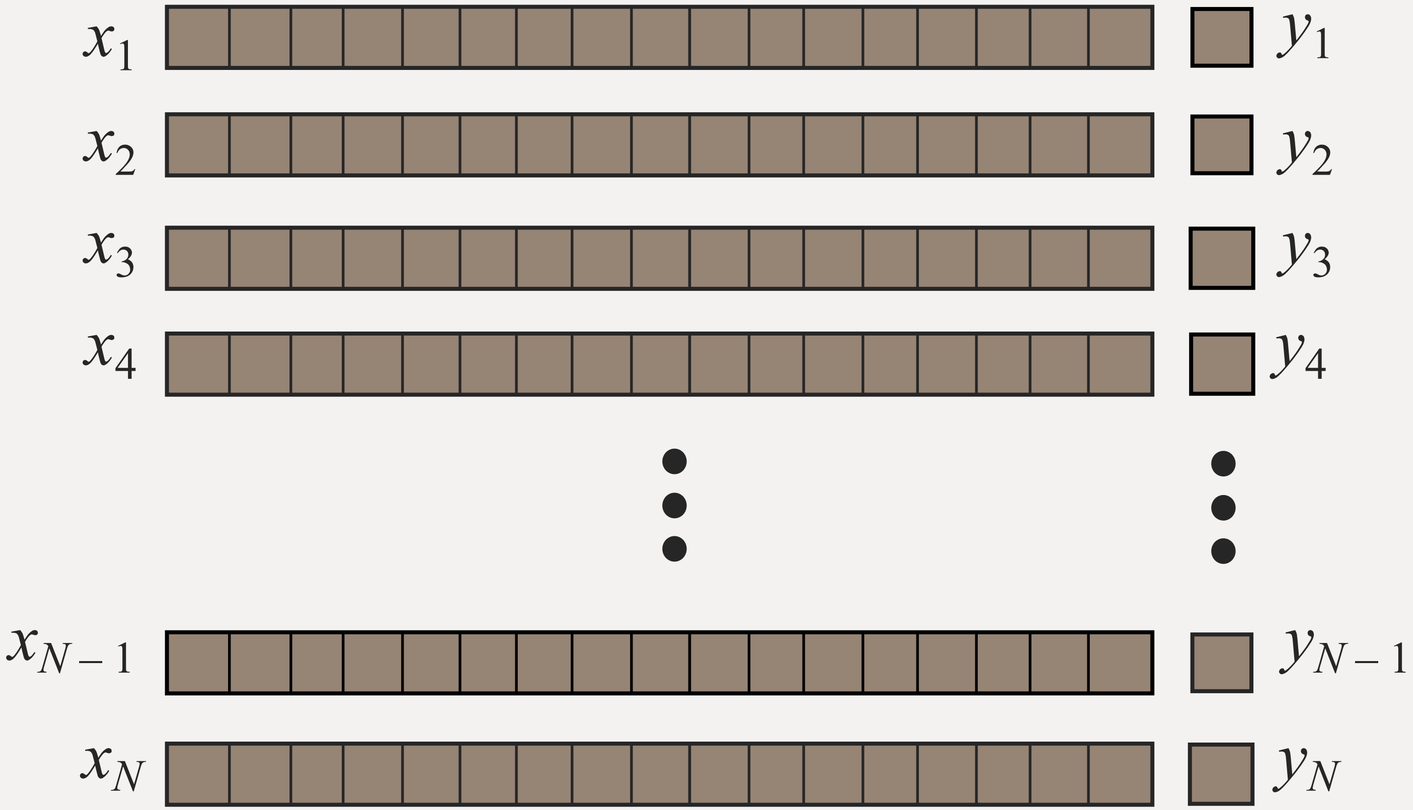
Test Set

- Standard practice in machine learning
- Created prior to any analysis
- Will never be used to learn or fit any parameters
- Can evaluate performance of network on test set
- Analogous to running a new experiment



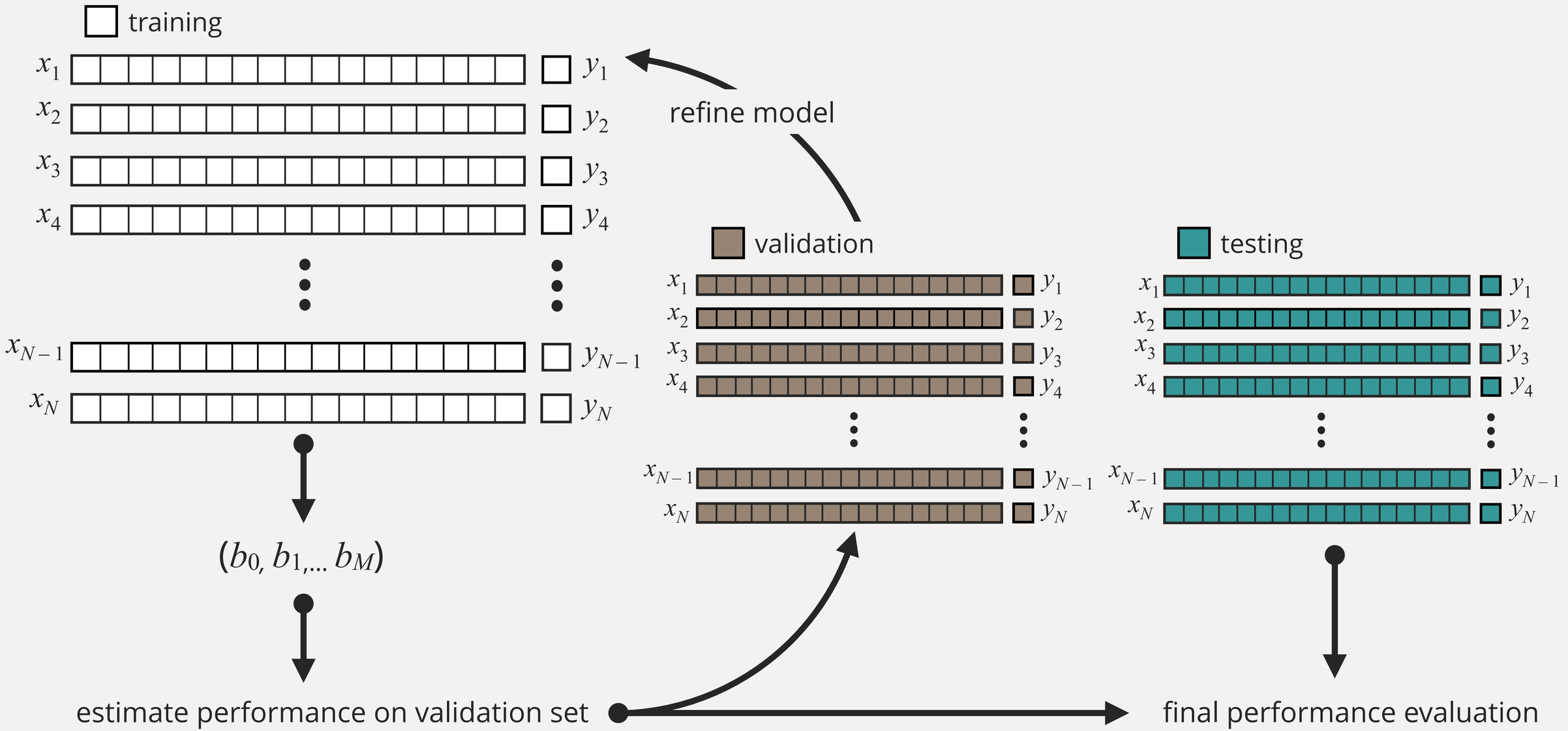
Test Set

- Should ideally only be used once
- Reusing a test set will lead to bias
- Bias results will lead to optimistic performance estimates



Validation Set

- Can be used to compare which approach is best
- Not used to learn parameters
- Used repeatedly to estimate the performance of a model
- Can be used to pick out the best performance model


training
x_1 y_1
x_2 y_2
x_3 y_3
x_4 y_4
x_{N-1} y_{N-1}
x_N y_N
refine model
validation
testing
$(b_0, b_1, \ldots b_M)$
estimate performance on validation set
final performance evaluation