

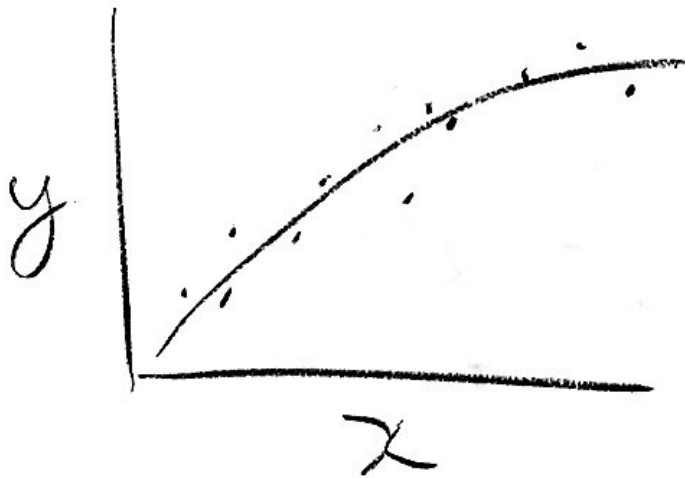
Model selection

- Two basic approaches
 - Hypothesis tests
 - Forward selection, backward selection
 - Predictive performance
 - emphasized in machine learning
 - basis for xIC (AIC, BIC, DIC, WAIC, etc)
- Code for the ant dataset
15_8_ants_model_comparison.md

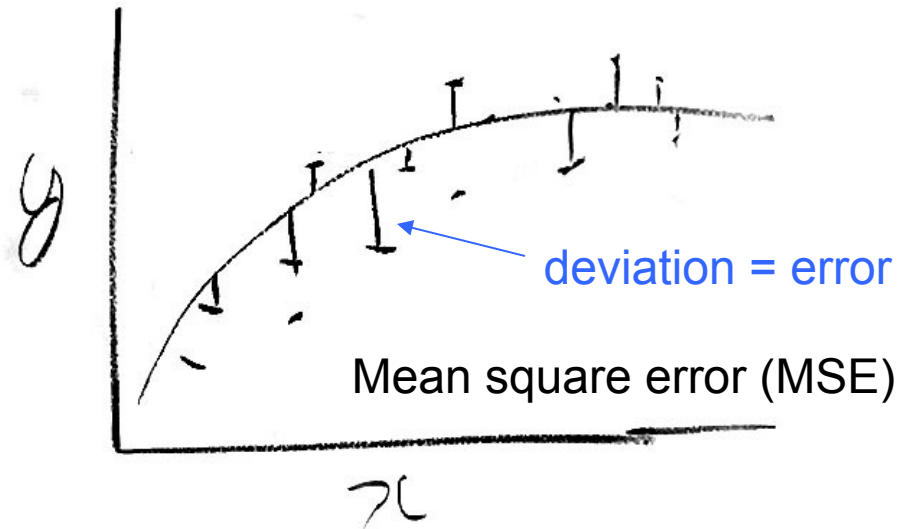
Predictive performance

Basic idea: out-of-sample validation

Fit model to training dataset

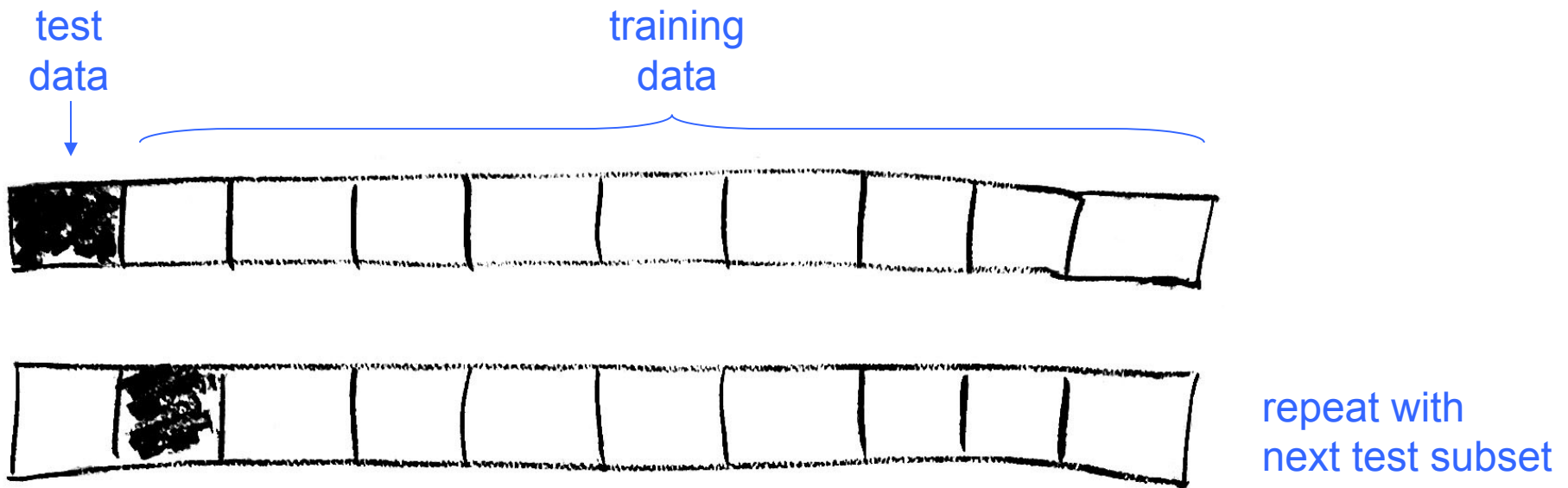


Test model on validation dataset



k-fold cross validation (CV)

Divide dataset into k parts (preferably randomly)



... repeat with each test subset

Leave-one-out cross validation

- LOOCV
- = k-fold CV for $k = n$

Algorithm

for each data point

 fit model without point

 predict for that point

 measure prediction error (compare to observed)

CV_error = mean error across points

Information criteria

- AIC
 - frequentist LOOCV asymptotically

$$\underbrace{-2\ln(L)}_{2*\text{nll} = \text{deviance}} + 2k$$

complexity penalty

k : number of estimated parameters in model, including stochastic parameters (e.g. variance)

Information criteria

- AICc
 - finite sample correction

$$\text{AIC}_c = \text{AIC} + \underbrace{\frac{2k(k+1)}{n - k - 1}}_{\text{finite sample correction}}$$

n: number of data points

k: number of estimated parameters
(including stochastic, e.g.
variance)

Information criteria

- WAIC
 - widely applicable information criterion
 - Bayesian LOOCV asymptotically
- LOOIC
 - leave-one-out information criterion
 - Bayesian LOOCV for finite samples
 - `loo()` function (knows what to do with `rstanarm`, `rstan`, `brms` etc objects)