

# Ensemble Methods

## Merge predictions

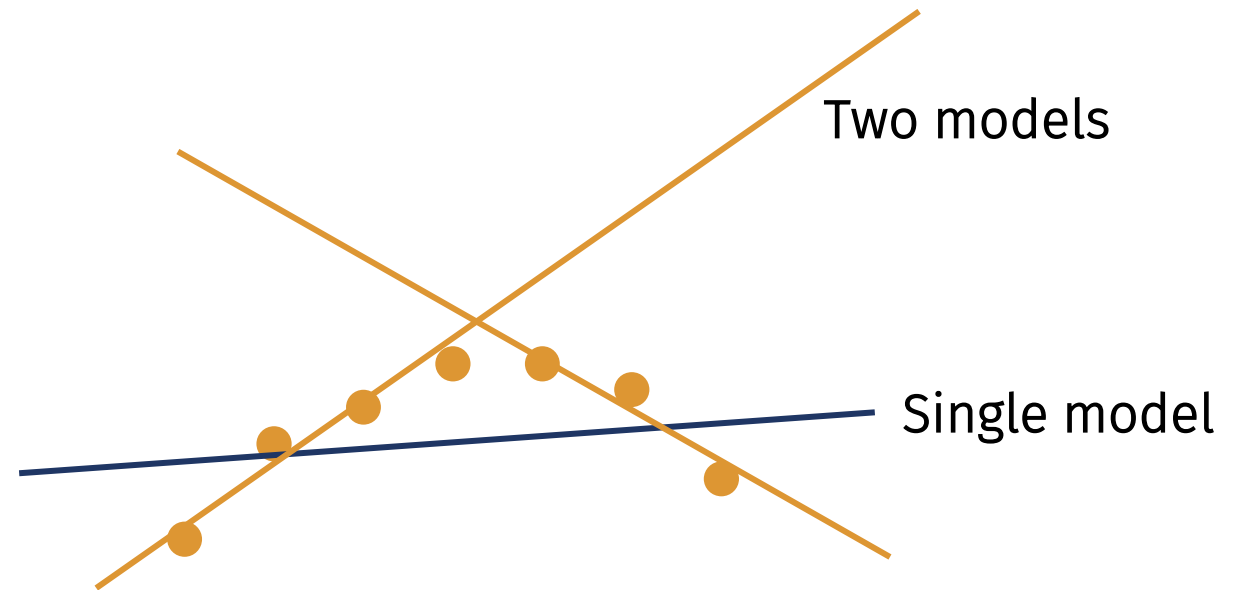
- Individual predictions might be off
- Perhaps bias too high

## Core idea

- Build (at least 2, typically more) models
- Models may be unstable
- Either
  - Learn when to trust which one
  - Reduce variance by aggregation

## Meta Algorithm

- Concrete base model is arbitrary
- Model mixing possible



## Merge predictions

- Take random subset of training data **with replacement**
- Random subset of features
- Fit decision trees

## Classification

- Majority vote

## Regression

- Average prediction

## Random forests

- Base learner is decision tree



## Refine predictions

- Build base model
- Focus on mispredictions, build new or stronger weighted model for those
- Many flavours

## Example: AdaBoost

Build a new weak learner, assign residual errors

Until end:

Train new weak learner  $f$  with training instances weighted by current error  
Find global weight  $w$  for prediction s.t. total error is minimized:  $P_+ = P + wf$

## Many learners

- Build diverse base models  $m_j$ 
  - Could even be different methods

## Combiner model

- Take all predictions  $m_i(x)$  as input
  - Optionally augmented by model features
- Combine prediction
  - Either learn reliability depending on features
  - Weights depending on regions

