Considering model complexity

- Modeling in another word

  Reduce the complexity of the data to something that is more comprehensible

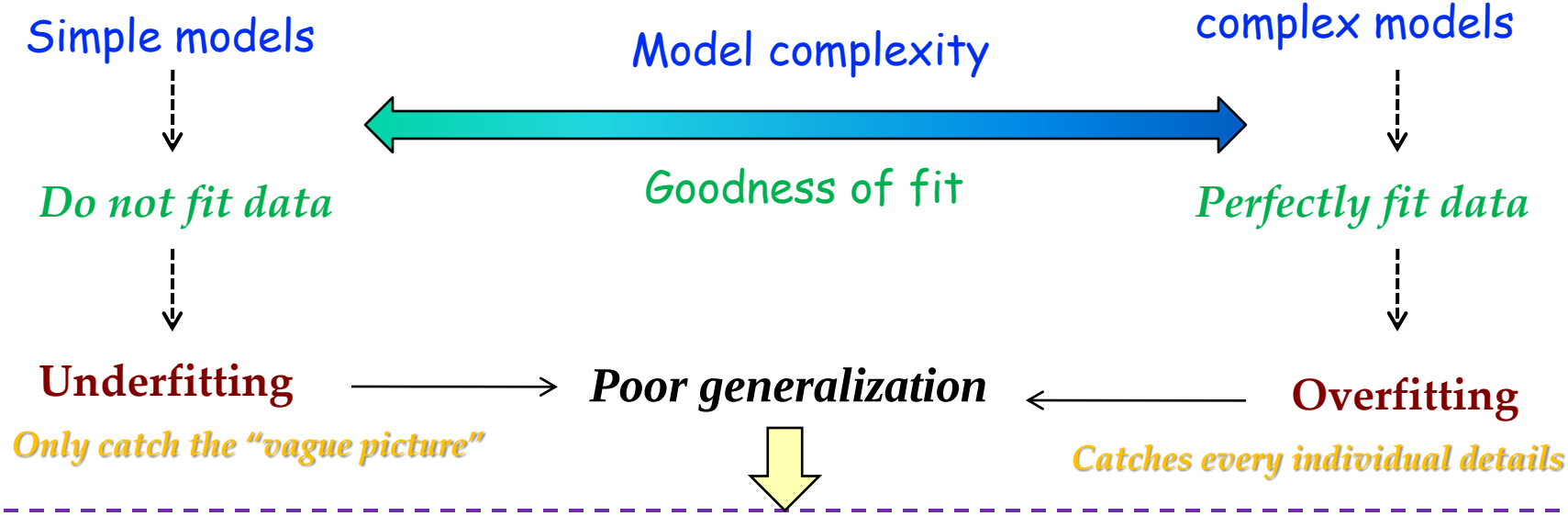
  - **High goodness of fit:** Can fit the training data reasonably well
  - **Low model complexity**: model should not be too difficult to describe

- Score function should reflects the difficulties of describing the current system

  $S_I(\theta, M)$ = # bits to describe the data given the model

  + # bits to describe the model (and parameter)

Why consider model complexity?



Improving generalization ability

$$MSE = Variance + Bias^2$$

$$= E[\hat{y} - \mu_y]^2 = \underline{E[\hat{y} - E(\hat{y})]^2} + \underline{E[E(\hat{y}) - \mu_y]^2}$$

Low complexity

Well-fit of data

Penalize complexity

- To achieve better compromise, the score function should penalizes:
  - Error made be by underfitting of the model
  - Complexity of the model

Score (model) = error(model) + penalty(model)