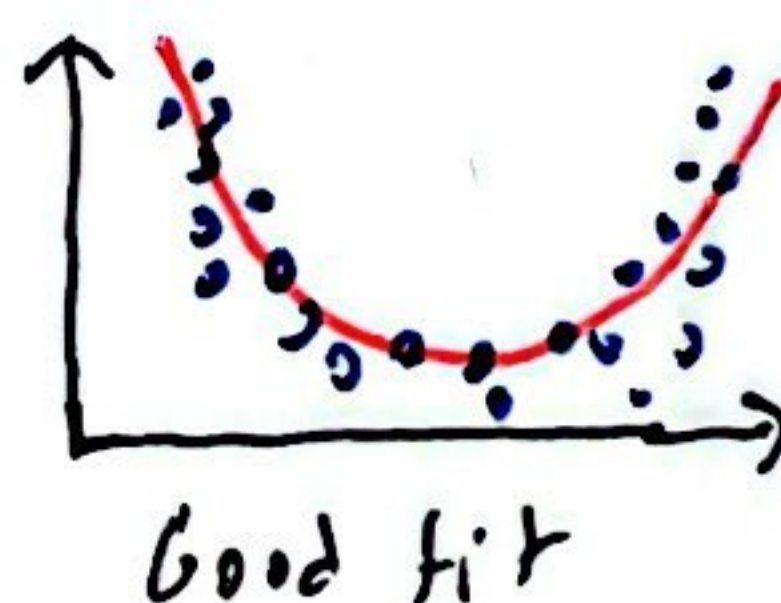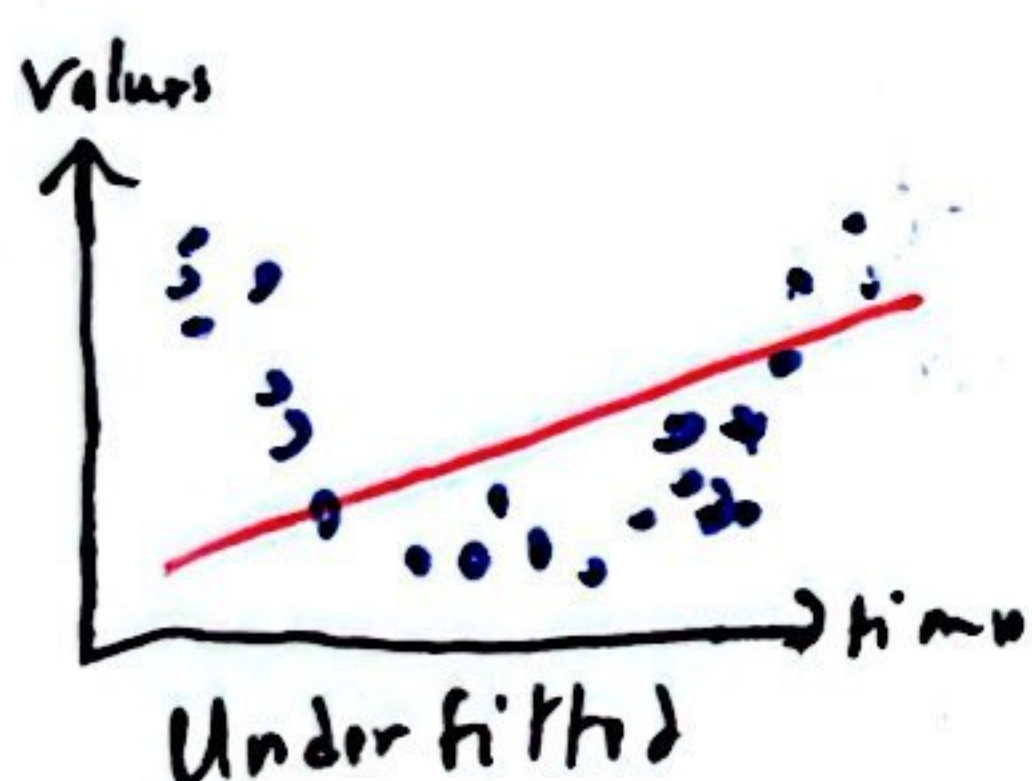# Over fitting & Under fitting

- Overfitting happens when the ML model learns the noise and random fluctuations in training data rather than learning the true underlying patterns its like memorizing the answers it does well on training data but bad on new unseen data.
  - happens when Model is too complex for task or training too long on too little data

- Under fitting occurs when a ML model is too simple to capture important patterns in data resulting in poor performance in training and test data. it makes over simplified assumptions. this happens when a model has a high bias. for ex using linear model for non-linear tasks

Exs:

Under fitted   Good fit   Overfitted

Validation

- Validation is the practice of evaluating a models performance on data it hasent been trained on (havent seen) by setting aside a portion of training data called "validation set" to simulate how well a model will perform on new unseen data
  - # Validation set is used to fine tune the model in training process before testing test set is used to evaluate the final completed model.
  - * this is one way of estimating bias and variance in training

A training test
Single dataset

B training Validation test

12