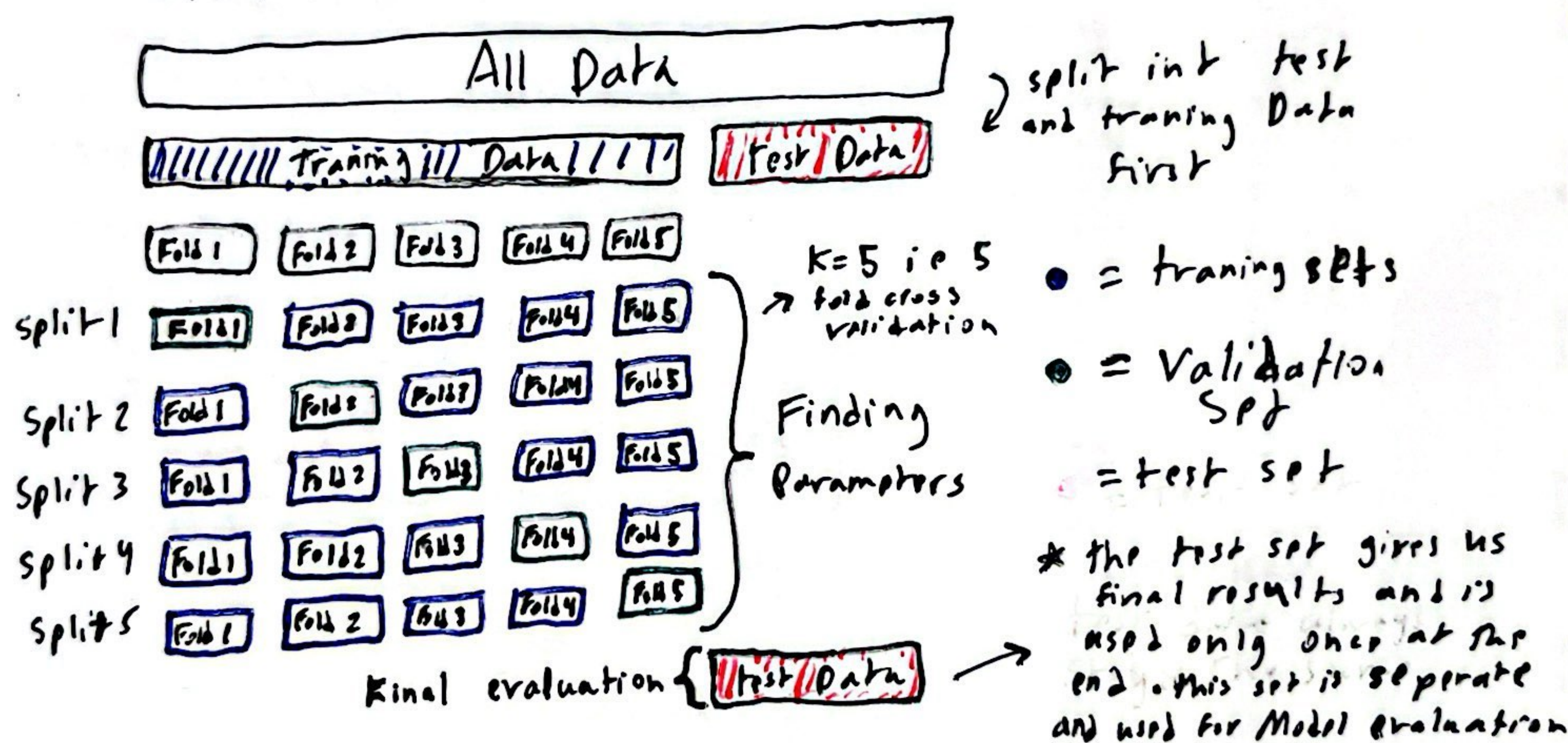# Cross Validation

- extends the concept of <u>validation</u> by repeatedly traning and validating the model on different splits of the data • Ex in 5 Fold cross validation troning data is divided into 5 parts and the model is traned 5 times each time using a diffrent part as the validation set <u>rest for traning. Another way of preventing over/under fitting</u>

* this is like rotating bettwen parts of troning data and at each validation set stoping, validating Model and then Figuringout what hyperparameters to change / other Model choices~ all in hope that it will improve the Model after each split.


```
┌─────────────────────────────────┐
│          All Data               │    ⟩ split int test
└─────────────────────────────────┘    ↙ and traning Data
                                            first
[////// Traning // Data /// ]  [ Test Data ]

[Fold 1] [Fold 2] [Fold 3] [Fold 4] [Fold 5]      K=5 ie 5      ● = traning sets
                                               ↗ fold cross
split 1 [Fold1] [Fold2] [Fold3] [Fold4] [Fold5]    validation
                                                              ● = Validation
Split 2 [Fold1] [Fold2] [Fold3] [Fold4] [Fold5]                    set
                                                 Finding
Split 3 [Fold1] [Fold2] [Fold3] [Fold4] [Fold5]  Parameters       = test set
Split 4 [Fold1] [Fold2] [Fold3] [Fold4] [Fold5]
Split 5 [Fold1] [Fold2] [Fold3] [Fold4] [Fold5]           * the test set gives us
                                                            final results and is
          Final evaluation { [Test Data] } →                used only once at the
                                                            end . this set is seperate
                                                            and used for Model evaluation
```

Regularization

- refers to techniques used to prevent <u>overfitting</u> by adding constraints or penelties that discourage a model from becoming too complex or fitting to closely to traning dah . it <u>keeps model parameters small</u>. its like squeezing the regretion line .amount of regularization is a hyperparameter

Ex: impat of reg.

without Regularization (over fit)

with Regularization 13 (Good fit)
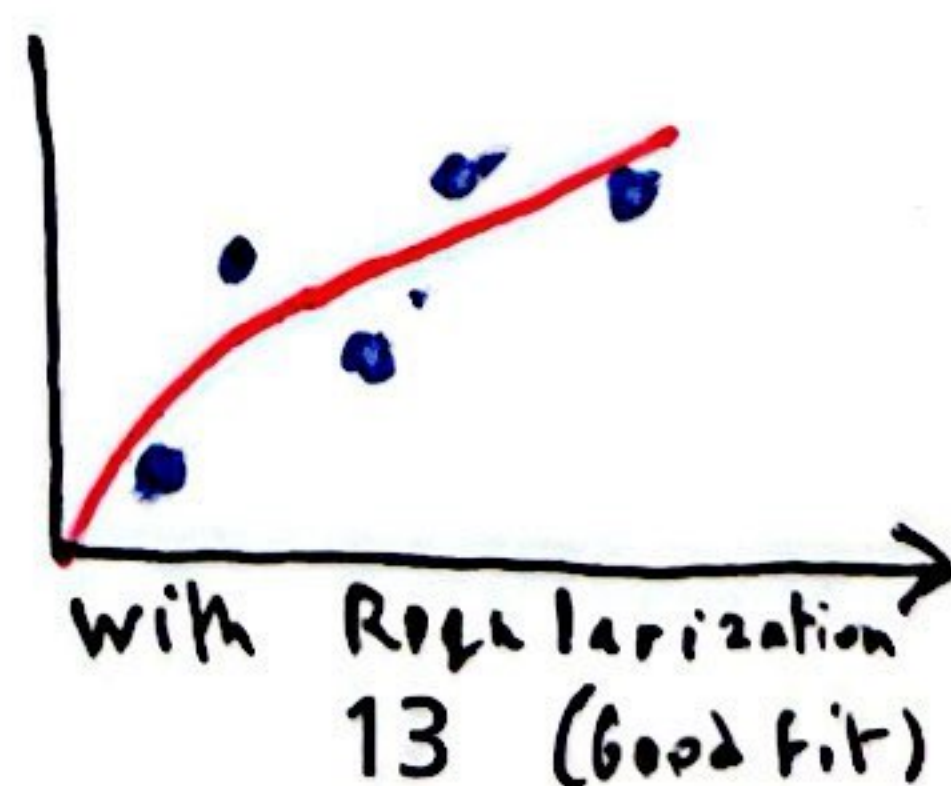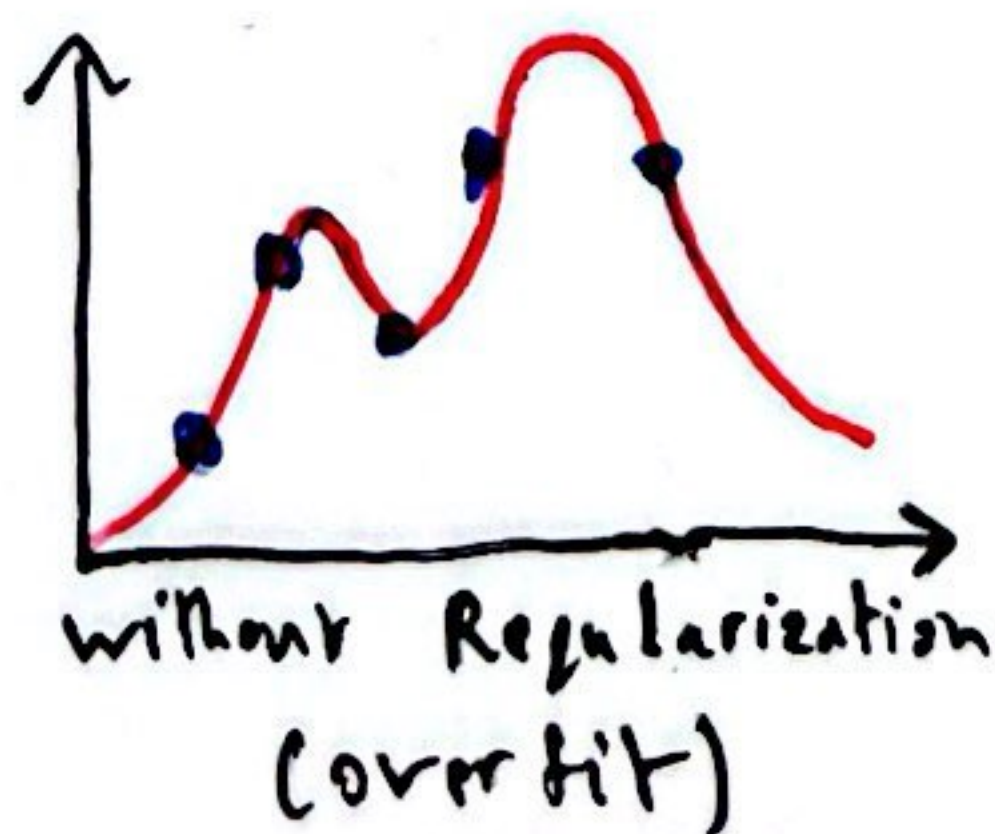
☆
too much regularization
=
Underfitting