kokchun giang

searching for the right model with the right parameters require **pipelines and gridsearch**

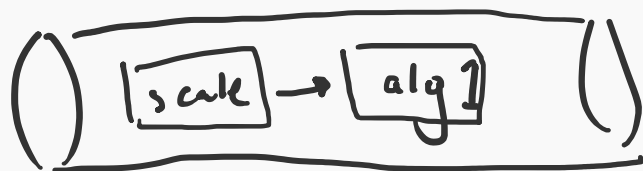# which **model** to choose and what **hyper parameters**?

for classification we have many many algorithms
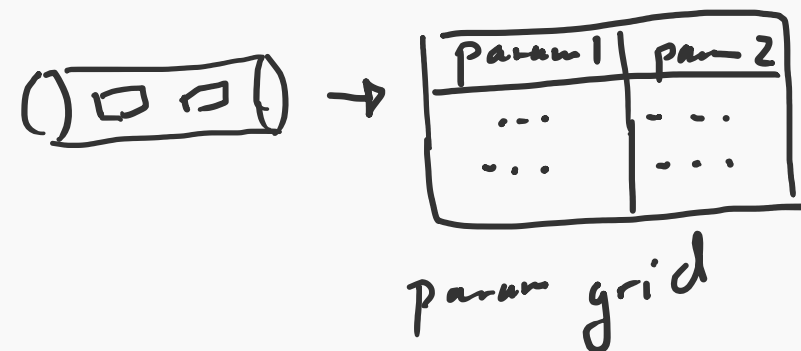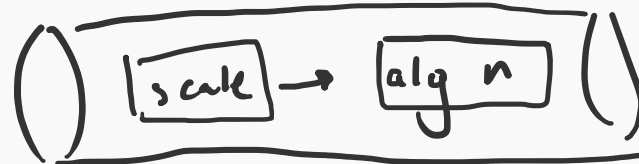
e.g. log.reg.
KNN
Random forest
XG Boost
...

each of them have a set of

hyperparameters that should be set differently depending on the dataset

Pipeline

param grid

→ for each pipeline each parameter combinations are tested using k-fold cross validation

↳ compare metrics