Train In Data

Nested Cross-Validation

Uses of Cross-Validation

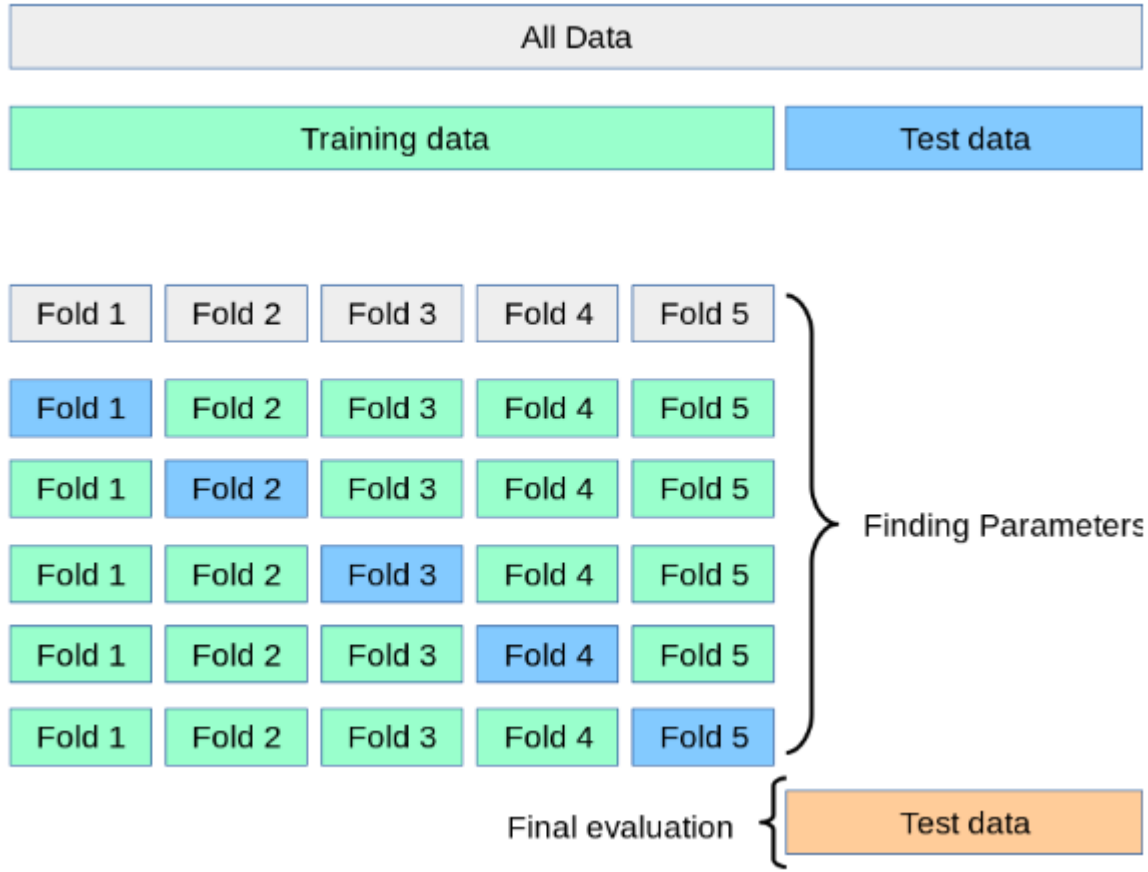
- Estimate the generalization error of a given model
- Select best performing model from a group of models
  - Different algorithms
  - Different feature subsets
- Select hyperparameters

Uses of Cross-Validation

- Estimate the generalization error of a given model
- Select best performing model from a group of models
  - Different algorithms
  - Different feature subsets
- Select hyperparameters

Cross-Validation & Generalization Error

When we have 1 model:



Generalization Error:

R2 = 0.8 ± 0.1

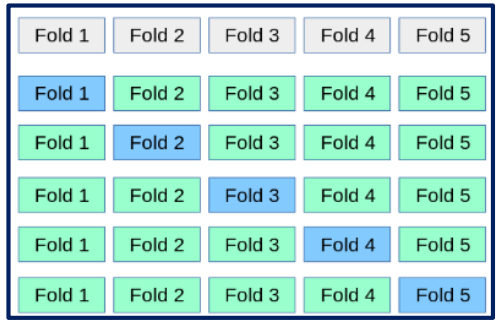
https://scikit-learn.org/stable/modules/cross_validation.html

Model selection, CV & Generalization Error

Generalization Error:

R2 = 0.8 ± 0.1

R2 = 0.8
R2 = 0.7
R2 = 0.9
R2 = 0.7
R2 = 0.9

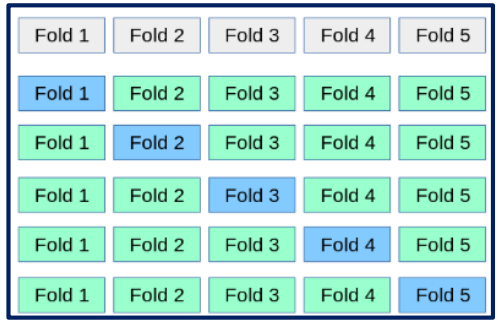


Hyperparams 1

Generalization Error:

R2 = 0.84 ± 0.0 6

R2 = 0.8
R2 = 0.8
R2 = 0.9
R2 = 0.8
R2 = 0.9

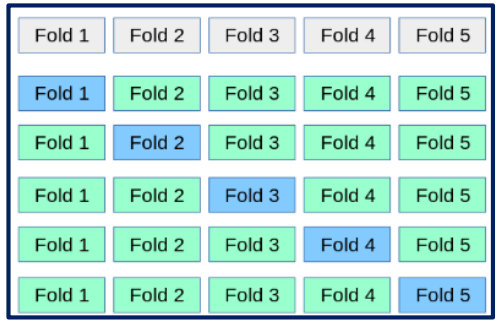


Hyperparams 2

Generalization Error:

R2 = 0.78 ± 0.11

R2 = 0.7
R2 = 0.7
R2 = 0.9
R2 = 0.7
R2 = 0.9



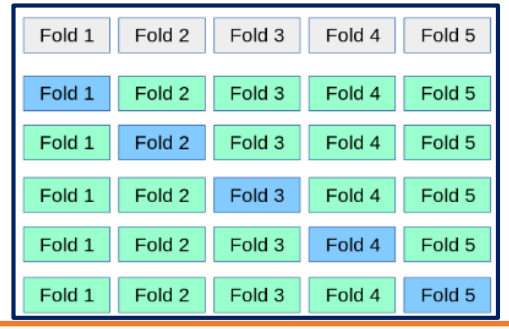
Hyperparams 3

Model selection, CV & Generalization Error

Generalization Error:

R2 = 0.8 ± 0.1

R2 = 0.8
R2 = 0.7
R2 = 0.9
R2 = 0.7
R2 = 0.9

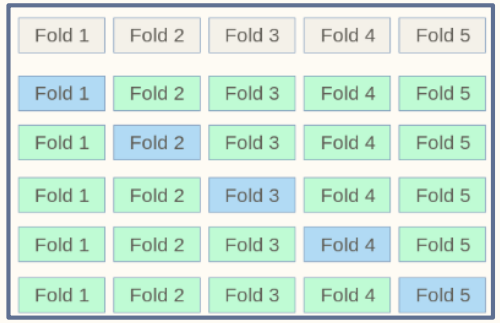


Hyperparams 1

Generalization Error:

R2 = 0.84 ± 0. 06

R2 = 0.8
R2 = 0.8
R2 = 0.9
R2 = 0.8
R2 = 0.9

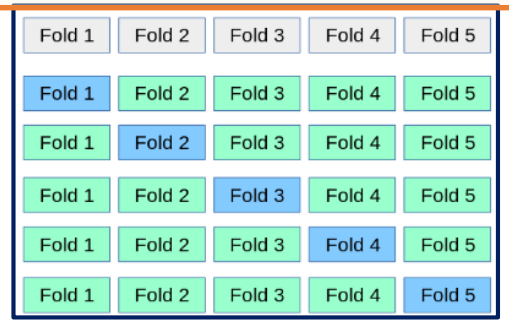


Hyperparams 2

Generalization Error:

R2 = 0.78 ± 0.11

R2 = 0.7
R2 = 0.7
R2 = 0.9
R2 = 0.7
R2 = 0.9



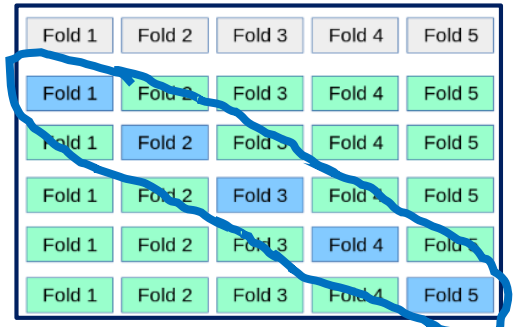
Hyperparams 3

Model selection, CV & Generalization Error

Generalization Error:

R2 = 0.8 ± 0.1

R2 = 0.8
R2 = 0.7
R2 = 0.9
R2 = 0.7
R2 = 0.9

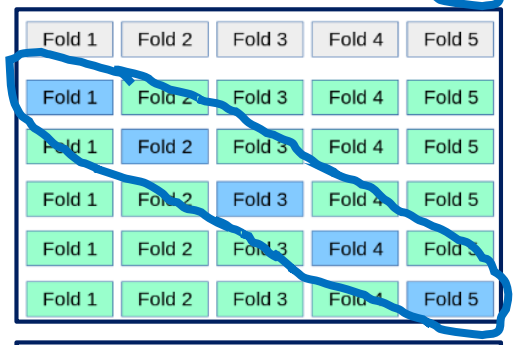


Hyperparams 1

Generalization Error:

R2 = 0.84 ± 0. 06

R2 = 0.8
R2 = 0.8
R2 = 0.9
R2 = 0.8
R2 = 0.9

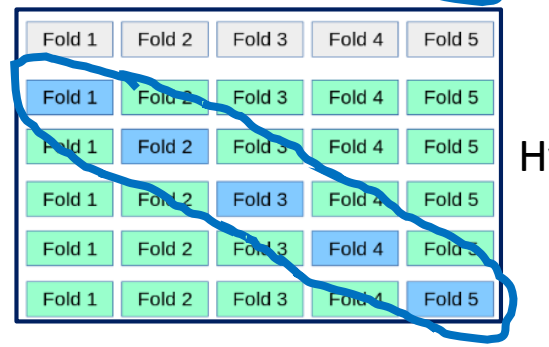


Hyperparams 2

Generalization Error:

R2 = 0.78 ± 0.11

R2 = 0.7
R2 = 0.7
R2 = 0.9
R2 = 0.7
R2 = 0.9



Hyperparams 3

- Validation tests are the **same** for all models
- Validation set "leaks" information to model "selection" procedure
- Generalization error is optimistically biased

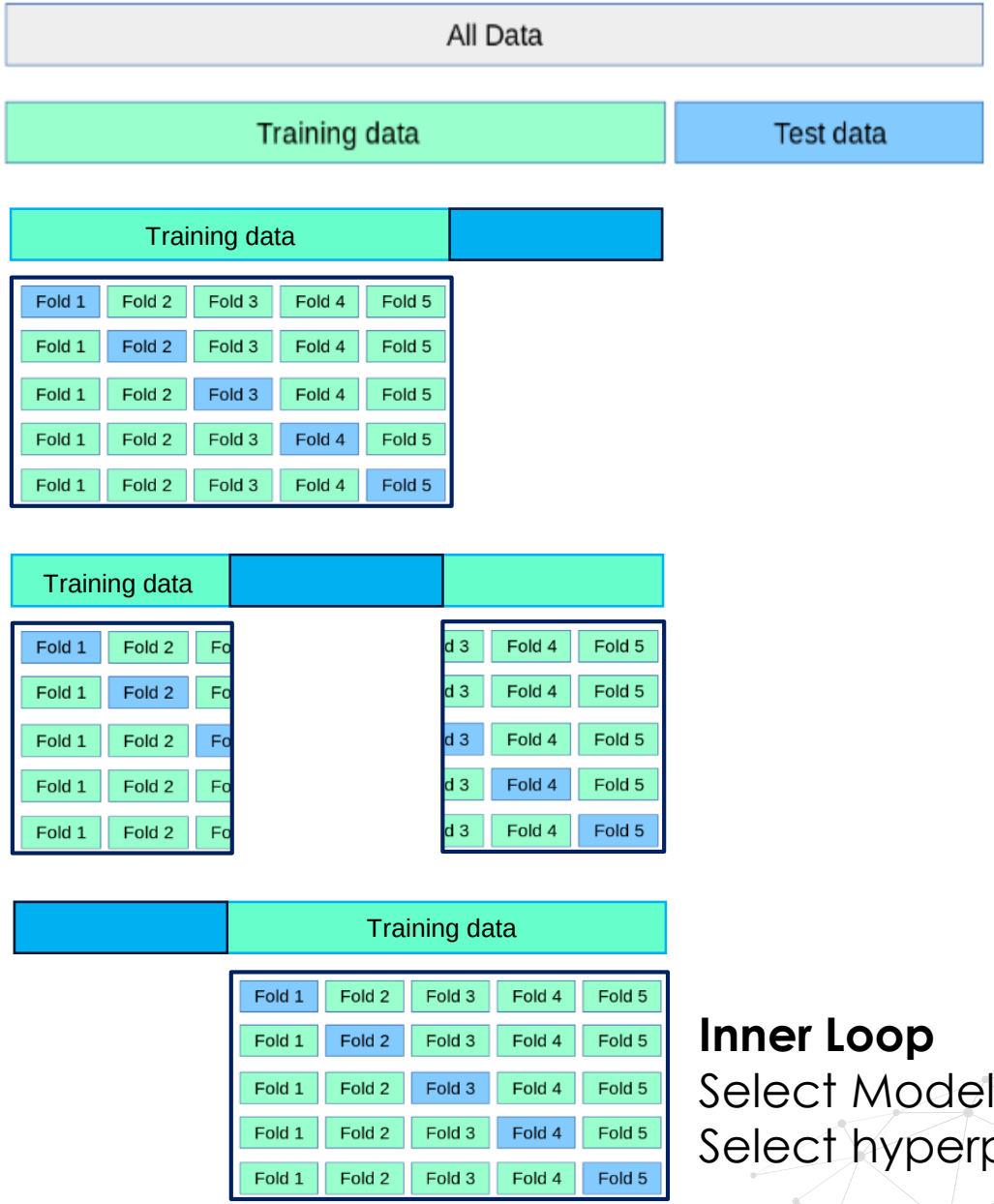
Model selection, CV & Generalization Error

We need a *different* test set to get an **unbiased** evaluation of the generalization error of selected model (or selected hyperparameters for the purpose of this course)

Particularly important for data science competitions.

Nested CV

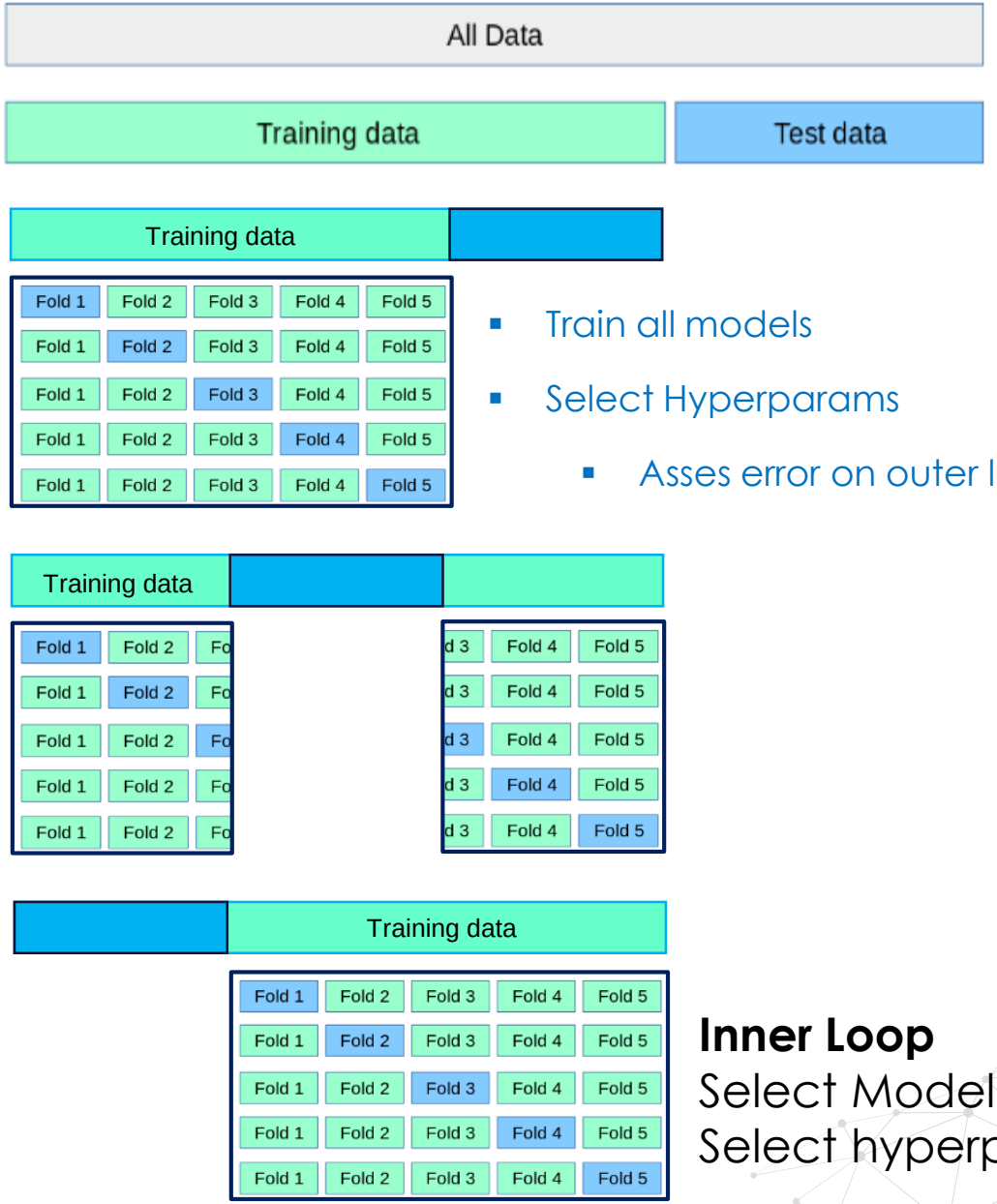
Outer Loop
Estimate
Generalization Error



Inner Loop
Select Model –
Select hyperparameters

Nested CV

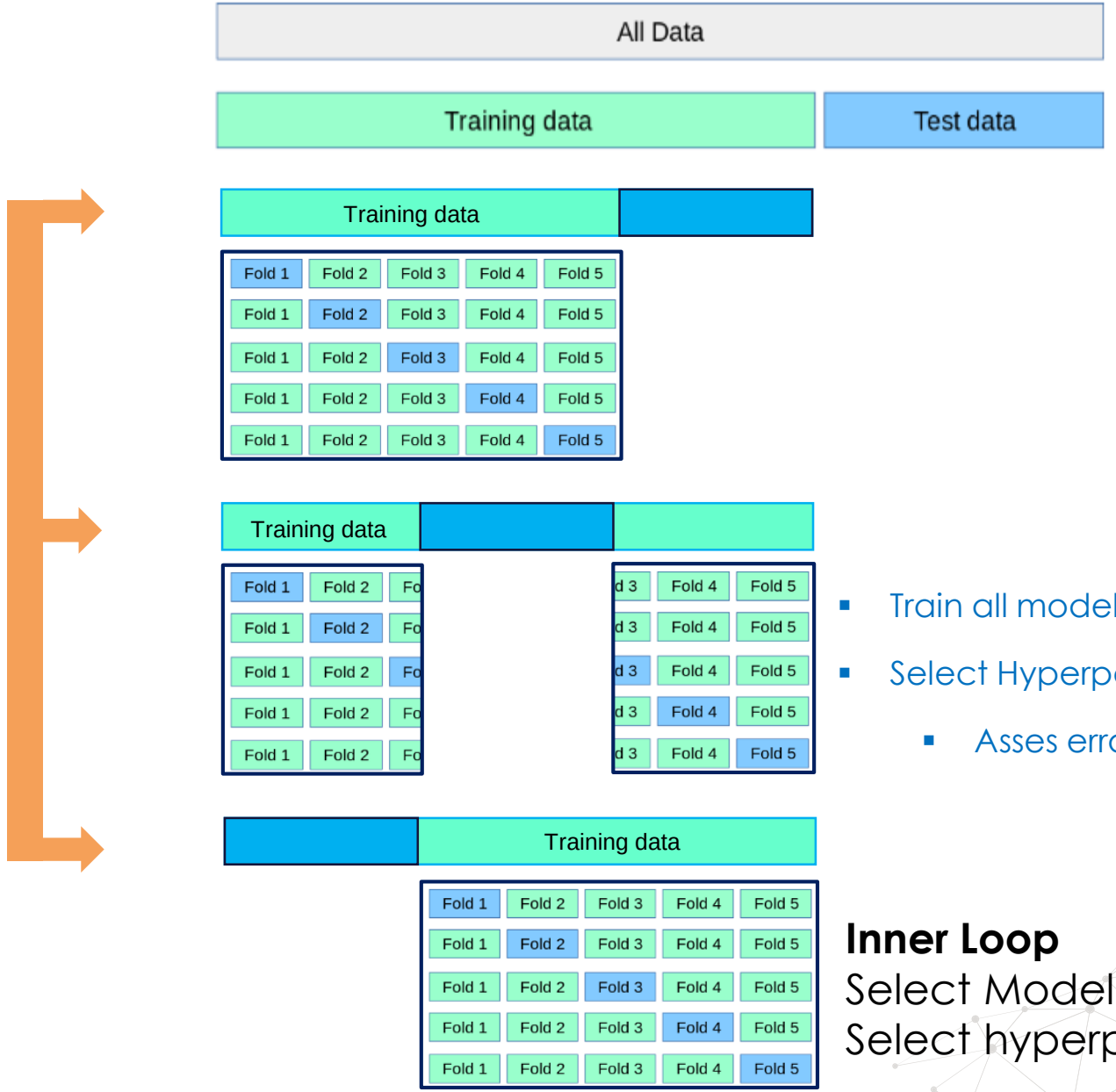
Outer Loop
Estimate
Generalization Error



Inner Loop
Select Model –
Select hyperparameters

Nested CV

Outer Loop
Estimate
Generalization Error



- Train all models
- Select Hyperparams
  - Asses error on outer loop

Inner Loop
Select Model –
Select hyperparameters

Final evaluation { Test data

Nested CV

Outer Loop
Estimate
Generalization Error

- Train all models
- Select Hyperparams
- Asses error on outer loop

Inner Loop
Select Model –
Select hyperparameters

Final evaluation { Test data

Nested CV, considerations

- Computationally expensive
- Useful when we need a good estimation of the generalization error
- Different inner models may have different hyperparameters, although I would expect to be among the top performing hyperparameters.

Train In Data

THANK YOU

www.trainindata.com