**Parameter spaces for different tabular deep models**

For different tabular deep learning models, we use BO and CV to find the best parameters. Here is a reference for the parameter space of different models:

**LightGBM:**

learning rate: 0,003463 – 0.00515

max\_depth: 3 - 5

num\_leaves: 5 - 8

num\_iterations: 800 - 1000

n\_estimators: 400 - 800

**LassoNet**

M: 7.0 - 30

Lambda: 0.1 – 20

node\_first\_layer: 200 – 800

node\_second\_layer: 200 – 400

batch size: 100 – 300

learning rate: 1,0e-5 – 1.0e-4

**TabR**

width: 100 – 400

attention dropout rate: 0,5

dropout rate in FFN: 0,5

learning rate: 1,0e-4 – 1.0E-6

**TabNet**

n\_d: 200-400

n\_a: 100 – 300

batch\_size: 100 - 400

step size: 100 – 300

L\_sparse: 0.1 – 0.005

gamma: 0 - 0.50,205

alpha: 0,5

**NODE**

total\_tree\_count: 200 – 600

tree\_output\_dim: 1 – 5

num\_layers: 1-5

tree\_depth: 3-6

**TabTransformer**

depth: 3 - 6

heads = 2 - 8

attn\_dropout: 0,1

ff\_dropout: 0.1 - 0.5

learning rate: 1e-6 – 1E-4

**FT Transformer**

layer: 1-6

feature embedding size: 100-500

head count: 4-8

attention dropout rate; 0,5

FN dropout rate: 0,45

learning rate: 1e-5

**AutoInt**

layer\_num: 2 - 6

att\_embedding\_size: 12 - 18

learning rate: 1.0e-5 – 2e-4

l2\_reg\_linear = 1e-05

l2\_reg\_embedding = 1e-05

**GANDALF**

learning rate: 0.0001 – 0.005

batch size: 400 – 2000

optimizer: adam

num\_trees: 4-18

depth trees: 3-8