

**TABLE 1** Hyperparameter space for DNNs.

Hyperparameter	Values
# nodes for first hidden layer	30, 40, 50, 75, 100, 120, 150
# hidden layers	4, 5, 6, 8, 10, 15
batch size	8, 16, 32, 64, 128
activation function	ReLu, sigmoid, tanh, SELU, ELU
optimizer	SGD, RMSprop, Adam, Adadelata, Adagrad, Adamax, Nadam, Ftrl
learning rate	$10^{-5}$ , $10^{-4}$ , $10^{-3}$ , $10^{-2}$
# epochs	100, 300, 400, 600, 1000, 1500
preprocessing method	StandartScaler, MinMaxScaler
regularization function	None, L2, L1, Dropout
regularization rate	$10^{-5}$ , $10^{-4}$ , $10^{-3}$ , $10^{-2}$
dropout rate	0.2, 0.3, 0.4, 0.5
weight initializer	Xavier Normal or Uniform, He Normal or Uniform, Random Normal or Uniform, Ones