

S1. The setting of parameters for four traditional machine learning methods.

K-nearest neighbor (`n_neighbors=5`, `weights='uniform'`, `leaf_size=30`, `p=2`, `metric='minkowski'`)

Random forest (`n_estimators=100`, `criterion='gini'`, `min_samples_split=2`, `min_samples_leaf=1`, `min_weight_fraction_leaf=0.0`, `max_features='auto'`, `bootstrap=True`)

Logistic regression (`penalty='l2'`, `tol=0.0001`, `C=1.0`, `intercept_scaling=1`, `class_weight=None`, `max_iter=100`)

Neural network (`hidden_layer_sizes=(100, )`, `activation='relu'`, `*`, `solver='adam'`, `alpha=0.0001`, `batch_size='auto'`, `learning_rate_init=0.001`, `momentum=0.9`)

S2. The computational cost for all machine learning approaches based on CGR representation.

