GridSearchCV, RandomizedSearchCV, and Cross-Validation Strategies Cheat Sheet

GridSearchCV Parameters

Parameter	Meaning	Typical Values / Notes
estimator	Model to train	Any sklearn estimator
param_grid	Dictionary of parameters to try	{'param': [values]}
scoring	Metric for evaluation	'accuracy', 'neg_mean_squared_error', etc.
cv	Cross-validation strategy	int, CV splitter object
n_jobs	Parallel jobs	-1 for all cores
verbose	Output verbosity	0, 1, 2
return_train_score	Include training scores	True/False
error_score	Score if fit fails	np.nan or 'raise'

RandomizedSearchCV Parameters

Parameter	Meaning	Typical Values / Notes
estimator	Model to train	Any sklearn estimator
param_distributions	Dictionary or distributions to sample	{'param': scipy.stats distribution}
n_iter	Number of parameter settings to sample	10–100
scoring	Metric for evaluation	Same as GridSearchCV
cv	Cross-validation strategy	int, CV splitter object
n_jobs	Parallel jobs	-1 for all cores
verbose	Output verbosity	0, 1, 2
random_state	Seed for reproducibility	int
error_score	Score if fit fails	np.nan or 'raise'

Cross-Validation Strategies

Strategy	Meaning	Key Parameters / Notes
KFold	Splits data into K folds	n_splits=5–10, shuffle=True
StratifiedKFold	KFold preserving class distribution	n_splits=5–10, shuffle=True
GroupKFold	KFold ensuring same group not in train/test	n_splits=5–10
TimeSeriesSplit	Splits for time-ordered data	n_splits=5–10, no shuffle
RepeatedKFold	KFold repeated multiple times	n_splits=5–10, n_repeats=2–10
RepeatedStratifiedKFold	Stratified KFold repeated	n_splits=5–10, n_repeats=2–10

LeaveOneOut	Each sample is its own test set	No params
LeavePOut	P samples as test set	p=1–5
ShuffleSplit	Random train/test splits	n_splits=10, test_size=0.1–0.5
StratifiedShuffleSplit	ShuffleSplit preserving class balance	Same params as ShuffleSplit

LeaveOneOut	Each sample is its own test set	No params
LeavePOut	P samples as test set	p=1–5
ShuffleSplit	Random train/test splits	n_splits=10, test_size=0.1–0.5