

# Machine Learning Models & Their Key Hyperparameters

## Tree-Based Models

Model	Hyperparameter	Meaning	Typical Values / Notes
CatBoost	iterations	Number of boosting rounds (trees)	500–2000
CatBoost	depth	Max depth of trees	4–10 (6–8 common)
CatBoost	learning_rate	Step size shrinkage	0.01–0.3
CatBoost	l2_leaf_reg	L2 regularization	1–10
CatBoost	border_count	Number of bins for numeric features	32–255
CatBoost	bagging_temperature	Randomness in sampling	0–1
CatBoost	random_strength	Randomness in split scores	1–10
XGBoost	n_estimators	Number of trees	100–1000
XGBoost	max_depth	Max depth of trees	3–10
XGBoost	learning_rate	Step size shrinkage	0.01–0.3
XGBoost	subsample	Row sampling fraction	0.5–1.0
XGBoost	colsample_bytree	Feature sampling fraction	0.5–1.0
XGBoost	gamma	Min loss reduction for split	0–5
XGBoost	lambda	L2 regularization	0–10
XGBoost	alpha	L1 regularization	0–10
LightGBM	n_estimators	Number of boosting rounds	100–1000
LightGBM	max_depth	Max depth of trees	-1 (no limit) or 3–10
LightGBM	learning_rate	Step size shrinkage	0.01–0.3
LightGBM	num_leaves	Max leaves per tree	31–255
LightGBM	min_child_samples	Min data per leaf	10–50
LightGBM	subsample	Row sampling fraction	0.5–1.0
LightGBM	colsample_bytree	Feature sampling fraction	0.5–1.0
LightGBM	reg_lambda	L2 regularization	0–10
LightGBM	reg_alpha	L1 regularization	0–10
Random Forest	n_estimators	Number of trees	100–1000
Random Forest	max_depth	Max depth of trees	None or 5–50
Random Forest	max_features	Features per split	'auto', 'sqrt', float
Random Forest	min_samples_split	Min samples to split	2–20
Random Forest	min_samples_leaf	Min samples per leaf	1–10
Random Forest	bootstrap	Sampling with replacement	True/False

## Linear Models

Model	Hyperparameter	Meaning	Typical Values / Notes
Linear Regression	fit_intercept	Include intercept term	True/False
Ridge/Lasso	alpha	Regularization strength	0.001–10
Ridge/Lasso	max_iter	Max iterations	1000–10000
Logistic Regression	penalty	Type of regularization	'l1', 'l2', 'elasticnet', 'none'
Logistic Regression	C	Inverse regularization strength	0.001–100
Logistic Regression	solver	Optimization method	'liblinear', 'lbfgs', 'saga'
Logistic Regression	max_iter	Max iterations	100–1000

## Support Vector Machines

Model	Hyperparameter	Meaning	Typical Values / Notes
SVC/SVR	C	Regularization	0.1–100
SVC/SVR	kernel	Kernel type	'linear', 'rbf', 'poly', 'sigmoid'
SVC/SVR	degree	Polynomial degree	2–5 (for poly)
SVC/SVR	gamma	Kernel coefficient	'scale', 'auto', float
SVR	epsilon	Insensitive loss margin	0.01–1.0

## K-Nearest Neighbors

Hyperparameter	Meaning	Typical Values
n_neighbors	Number of neighbors	3–50
weights	Voting scheme	'uniform', 'distance'
metric	Distance metric	'minkowski', 'euclidean', etc.
p	Power for Minkowski	1=Manhattan, 2=Euclidean

## Naive Bayes

Model	Hyperparameter	Meaning	Typical Values
GaussianNB	var_smoothing	Stability term	1e-9–1e-6
MultinomialNB	alpha	Additive smoothing	0–1
MultinomialNB	fit_prior	Learn class priors	True/False

## Neural Networks (TensorFlow/Keras)

Hyperparameter	Meaning	Typical Values
epochs	Number of training passes	10–200

Hyperparameter	Meaning	Typical Values
batch_size	Samples per update	16, 32, 64, 128
learning_rate	Step size	0.0001–0.1
optimizer	Optimization algorithm	Adam, SGD, RMSprop
activation	Activation function	ReLU, sigmoid, tanh
units	Neurons per layer	32–1024
dropout_rate	Fraction dropped	0–0.5
layers	Number of hidden layers	1–10

## Clustering

Model	Hyperparameter	Meaning	Typical Values
KMeans	n_clusters	Number of clusters	2–20
KMeans	init	Initialization method	'k-means++', 'random'
KMeans	max_iter	Max iterations	100–1000
KMeans	n_init	Number of initializations	10–50
DBSCAN	eps	Neighborhood radius	0.1–10
DBSCAN	min_samples	Min points per cluster	3–10
DBSCAN	metric	Distance metric	'euclidean', 'manhattan', etc.