■ Machine Learning Algorithms – One Page Cheat Sheet

Supervised Learning - Regression

- Linear Regression – Predicts continuous values. Parameters: fit_intercept, normalize, n_jobs. Cost Function: MSE.
- Logistic Regression – Classification. Parameters: penalty, C, solver, max_iter, multi_class. Cost Function: Cross-Entropy (Log Loss).

Supervised Learning - Classification

- Decision Tree – Parameters: criterion, max_depth, min_samples_split, max_features. Cost Function: Gini/Entropy.
- Random Forest – Parameters: n_estimators, max_depth, bootstrap, oob_score.
- Support Vector Machine (SVM) – Parameters: C, kernel, gamma, degree. Cost Function: Hinge Loss.
- K-Nearest Neighbors (KNN) – Parameters: n_neighbors, weights, p, metric.
- Naïve Bayes – Parameters: alpha (Laplace smoothing), var_smoothing.

Unsupervised Learning

- K-Means Clustering – Parameters: n_clusters, init, max_iter, tol. Cost Function: Inertia.
- Hierarchical Clustering – Parameters: linkage (ward, complete, average).
- PCA (Dimensionality Reduction) – Parameters: n_components, svd_solver, whiten.

Ensemble & Boosting

- Gradient Boosting / XGBoost / LightGBM – Parameters: n_estimators, learning_rate, max_depth, subsample.

Regularization

- L1 (Lasso): Shrinks some weights to zero (feature selection).
- L2 (Ridge): Shrinks weights smoothly.
- Elastic Net: Combination of both.

Cost Functions (Quick Recap)

- Regression → MSE, MAE, Huber Loss
- Classification → Cross-Entropy, Hinge Loss
- Clustering → Inertia, Silhouette Score

Golden Rule

- Use simpler models first.
- Prevent overfitting with regularization.

- Tune hyperparameters for best performance.