FM3 - FREE SAMPLE

ML Algorithms Reference Card

Free sample from the 12-sheet AI/ML Interview Cheat Sheet Bundle

SAMPLE SHEET

ML Algorithms Selection Guide

Problem ??' Algorithm Map

- Tabular classification (structured): Logistic Regression baseline ??' XGBoost/LightGBM ??' Neural net if huge data
- Tabular regression: Linear/Ridge ??' GBDT ??' avoid deep learning unless sequence/spatial structure
- Imbalanced binary: class weights, focal loss, threshold tuning, PR-AUC over accuracy
- Ranking/LTR: pointwise (regression on labels) ??' pairwise (RankNet) ??' listwise (LambdaMART)
- Clustering: K-Means (spherical) ??' DBSCAN (arbitrary shape) ??' Hierarchical (dendrogram cuts)

Key Equations

- Logistic: P(y=1|x) = ??(w?-x + b); cross-entropy loss
- Softmax: P(y=k|x) = exp(w_k?-x) / ??_j exp(w_j?-x)
- Gini impurity (trees): 1 - ??_k p_k??; Information gain = parent - weighted children
- Bias-variance: E[(y-?-)??] = Bias?? + Variance + Irreducible noise

Hyperparameter Quick Reference

- XGBoost: max_depth 4-8, learning_rate 0.01-0.1, n_estimators tune with early stopping
- Random Forest: n_estimators 100-500, max_features sqrt(p) classification
- SVM: C (regularization), kernel RBF gamma inversely related to feature scale
- Neural nets: Adam lr 1e-3 to 1e-4, batch norm, dropout 0.1-0.5