

GRID SEARCH = exhaustive search over all possible combinations of hyperparameters.

RANDOM SEARCH = sample hyperparam from prior distribution
(explore space randomly as per chosen distribⁿ)
faster than grid.

STOCHASTIC HILL CLIMBING = Initial set of hyperparam → random adjustment → improved^N
(till local optima) ↓ accept

BAYESIAN = optimises next choice instead of random — balances exploration & exploitation.
grid > bayesian > random

- = assume performance fⁿ from gaussian, maintains posterior
- = optimizes expected improvement on curr. best result. (upper confidence bound)
- = quickly finds optimum of non-convex function
- = evaluations are computationally heavy, but few.