
Prediction records
JSON / CSV
Protocol identity
rows, split, seed
Evaluation definition
metric + population
Reported estimate
with uncertainty
Invalid predictions and failed runs retain their original evaluation populations.