

Table 1: Performance Results for NeuralNetwork\100samplesn_layersn_5n_10n_3Problem

Optimizer	Mean Final Value	Std Dev	Best Value	Worst Value	Mean Func Evals	Success Rate (%)	Mean Time (s)
Adam-Fast	1.34×10^{-1}	1.35×10^{-2}	1.09×10^{-1}	1.67×10^{-1}	231.8	60.0	0.194
QQN-Bisection-2	1.43×10^{-1}	1.12×10^{-2}	1.17×10^{-1}	1.70×10^{-1}	551.8	0.0	0.431
QQN-Backtracking	1.45×10^{-1}	9.11×10^{-3}	1.23×10^{-1}	1.62×10^{-1}	545.4	0.0	0.448
QQN-Bisection-1	1.46×10^{-1}	1.02×10^{-2}	1.30×10^{-1}	1.78×10^{-1}	498.6	0.0	0.448
QQN-StrongWolfe	1.47×10^{-1}	1.32×10^{-2}	1.23×10^{-1}	1.73×10^{-1}	476.8	0.0	0.466
QQN-MoreThuente	1.53×10^{-1}	2.23×10^{-2}	1.16×10^{-1}	1.89×10^{-1}	496.8	0.0	0.452
L-BFGS-Conservative	1.61×10^{-1}	5.23×10^{-3}	1.53×10^{-1}	1.70×10^{-1}	574.9	0.0	0.403
L-BFGS	1.68×10^{-1}	1.26×10^{-2}	1.54×10^{-1}	2.11×10^{-1}	559.0	0.0	0.366
L-BFGS-Aggressive	1.72×10^{-1}	6.30×10^{-3}	1.59×10^{-1}	1.87×10^{-1}	696.4	0.0	0.359
QQN-GoldenSection	1.72×10^{-1}	7.27×10^{-3}	1.60×10^{-1}	1.95×10^{-1}	902.1	0.0	0.311
Adam-WeightDecay	1.78×10^{-1}	8.36×10^{-3}	1.60×10^{-1}	1.95×10^{-1}	502.0	0.0	0.422
QQN-CubicQuadraticInterpolation	1.94×10^{-1}	4.04×10^{-3}	1.87×10^{-1}	2.03×10^{-1}	394.7	0.0	0.505
Adam-AMSGrad	2.01×10^{-1}	5.44×10^{-3}	1.90×10^{-1}	2.13×10^{-1}	502.0	0.0	0.422
Adam	2.03×10^{-1}	4.28×10^{-3}	1.96×10^{-1}	2.15×10^{-1}	502.0	0.0	0.420
GD-WeightDecay	2.08×10^{-1}	2.76×10^{-3}	2.04×10^{-1}	2.14×10^{-1}	335.0	0.0	0.478
GD	2.16×10^{-1}	2.35×10^{-3}	2.10×10^{-1}	2.20×10^{-1}	335.0	0.0	0.474
Trust Region-Conservative	2.19×10^{-1}	2.81×10^{-3}	2.14×10^{-1}	2.24×10^{-1}	119.2	0.0	0.075
GD-Nesterov	2.28×10^{-1}	3.96×10^{-3}	2.22×10^{-1}	2.39×10^{-1}	21.2	0.0	0.029
GD-Momentum	2.32×10^{-1}	5.60×10^{-3}	2.18×10^{-1}	2.40×10^{-1}	21.4	0.0	0.029
Trust Region-Standard	5.20×10^{-1}	6.47×10^{-1}	1.90×10^{-1}	2.59	36.5	0.0	0.023
Trust Region-Adaptive	9.39×10^{-1}	2.20	1.77×10^{-1}	1.01×10^1	59.1	0.0	0.037