

Table 1: Performance Results for NeuralNetwork\100samplesn_layersn_10n_20n_5Problem

Optimizer	Mean Final Value	Std Dev	Best Value	Worst Value	Mean Func Evals	Success Rate (%)	Mean Time (s)
Adam-Fast	4.13×10^{-2}	6.27×10^{-3}	2.82×10^{-2}	5.46×10^{-2}	198.4	30.0	0.308
QQN-StrongWolfe	5.00×10^{-2}	1.08×10^{-2}	3.31×10^{-2}	7.12×10^{-2}	474.6	0.0	0.824
QQN-Bisection-2	5.21×10^{-2}	4.83×10^{-3}	4.17×10^{-2}	5.98×10^{-2}	535.4	0.0	0.769
QQN-MoreThuente	5.58×10^{-2}	2.13×10^{-2}	3.42×10^{-2}	9.60×10^{-2}	501.9	0.0	0.805
QQN-Bisection-1	5.61×10^{-2}	4.70×10^{-3}	4.46×10^{-2}	6.44×10^{-2}	503.3	0.0	0.791
QQN-Backtracking	6.08×10^{-2}	5.66×10^{-3}	5.00×10^{-2}	7.12×10^{-2}	544.4	0.0	0.793
Adam-WeightDecay	6.69×10^{-2}	6.33×10^{-3}	5.57×10^{-2}	8.14×10^{-2}	502.0	0.0	0.788
L-BFGS-Conservative	6.70×10^{-2}	1.12×10^{-2}	4.58×10^{-2}	9.09×10^{-2}	568.1	0.0	0.752
L-BFGS-Aggressive	8.56×10^{-2}	4.11×10^{-3}	7.95×10^{-2}	9.59×10^{-2}	705.8	0.0	0.656
QQN-GoldenSection	8.57×10^{-2}	3.80×10^{-3}	7.47×10^{-2}	9.50×10^{-2}	895.5	0.0	0.550
Adam-AMSGrad	1.16×10^{-1}	5.80×10^{-3}	1.02×10^{-1}	1.25×10^{-1}	502.0	0.0	0.788
Adam	1.17×10^{-1}	4.31×10^{-3}	1.08×10^{-1}	1.24×10^{-1}	502.0	0.0	0.786
QQN-CubicQuadraticInterpolation	1.18×10^{-1}	3.98×10^{-3}	1.06×10^{-1}	1.26×10^{-1}	394.2	0.0	0.915
L-BFGS	1.18×10^{-1}	8.87×10^{-2}	7.22×10^{-2}	4.83×10^{-1}	493.2	0.0	0.580
GD-WeightDecay	1.43×10^{-1}	2.34×10^{-3}	1.39×10^{-1}	1.47×10^{-1}	335.0	0.0	0.899
GD-Nesterov	1.50×10^{-1}	2.57×10^{-2}	1.01×10^{-1}	1.68×10^{-1}	98.9	0.0	0.263
Trust Region-Conservative	1.50×10^{-1}	2.63×10^{-3}	1.46×10^{-1}	1.56×10^{-1}	132.8	0.0	0.153
GD	1.52×10^{-1}	1.96×10^{-3}	1.49×10^{-1}	1.56×10^{-1}	335.0	0.0	0.884
GD-Momentum	1.60×10^{-1}	1.56×10^{-2}	1.12×10^{-1}	1.71×10^{-1}	51.9	0.0	0.136
Trust Region-Adaptive	1.77×10^{-1}	3.44×10^{-2}	1.30×10^{-1}	3.17×10^{-1}	18.6	0.0	0.022
Trust Region-Standard	3.98×10^{-1}	3.07×10^{-1}	2.06×10^{-1}	1.14	6.8	0.0	0.008