

Table 1: Performance Results for LogisticRegression\100samplesn5featuresnreg0.01Problem

Optimizer	Mean Final Value	Std Dev	Best Value	Worst Value	Mean Func Evals	Success Rate (%)	Mean Time (s)
QQN-Backtracking	3.15×10^{-1}	2.65×10^{-12}	3.15×10^{-1}	3.15×10^{-1}	74.0	0.0	0.027
QQN-StrongWolfe	3.15×10^{-1}	2.50×10^{-12}	3.15×10^{-1}	3.15×10^{-1}	73.7	0.0	0.031
QQN-Bisection-1	3.15×10^{-1}	3.08×10^{-12}	3.15×10^{-1}	3.15×10^{-1}	102.5	0.0	0.042
QQN-CubicQuadraticInterpolation	3.15×10^{-1}	2.98×10^{-12}	3.15×10^{-1}	3.15×10^{-1}	155.8	0.0	0.072
QQN-GoldenSection	3.15×10^{-1}	3.25×10^{-12}	3.15×10^{-1}	3.15×10^{-1}	320.0	0.0	0.071
QQN-MoreThuente	3.15×10^{-1}	2.52×10^{-12}	3.15×10^{-1}	3.15×10^{-1}	124.3	0.0	0.047
QQN-Bisection-2	3.15×10^{-1}	1.96×10^{-12}	3.15×10^{-1}	3.15×10^{-1}	98.0	0.0	0.033
L-BFGS-Conservative	3.15×10^{-1}	5.45×10^{-11}	3.15×10^{-1}	3.15×10^{-1}	556.0	0.0	0.187
L-BFGS	3.15×10^{-1}	6.90×10^{-5}	3.15×10^{-1}	3.16×10^{-1}	591.9	0.0	0.183
L-BFGS-Aggressive	3.15×10^{-1}	1.76×10^{-4}	3.15×10^{-1}	3.16×10^{-1}	609.4	0.0	0.165
Adam-Fast	3.16×10^{-1}	1.06×10^{-4}	3.16×10^{-1}	3.16×10^{-1}	90.4	0.0	0.033
GD-Momentum	3.33×10^{-1}	6.13×10^{-4}	3.32×10^{-1}	3.35×10^{-1}	335.0	0.0	0.184
GD-Nesterov	3.34×10^{-1}	5.75×10^{-4}	3.33×10^{-1}	3.35×10^{-1}	335.0	0.0	0.185
Trust Region-Conservative	4.25×10^{-1}	2.55×10^{-3}	4.20×10^{-1}	4.29×10^{-1}	87.2	0.0	0.026
GD-WeightDecay	4.51×10^{-1}	5.13×10^{-3}	4.42×10^{-1}	4.59×10^{-1}	335.0	0.0	0.185
Adam-WeightDecay	4.71×10^{-1}	1.44×10^{-2}	4.44×10^{-1}	4.92×10^{-1}	502.0	0.0	0.182
GD	5.57×10^{-1}	1.52×10^{-2}	5.28×10^{-1}	5.88×10^{-1}	335.0	0.0	0.183
Adam-AMSGrad	5.98×10^{-1}	2.17×10^{-2}	5.63×10^{-1}	6.32×10^{-1}	502.0	0.0	0.184
Adam	6.01×10^{-1}	1.92×10^{-2}	5.63×10^{-1}	6.31×10^{-1}	502.0	0.0	0.182
Trust Region-Standard	6.92×10^{-1}	2.02×10^{-2}	6.46×10^{-1}	7.36×10^{-1}	5.0	0.0	0.002
Trust Region-Adaptive	6.94×10^{-1}	2.60×10^{-2}	6.43×10^{-1}	7.44×10^{-1}	5.0	0.0	0.002