

Table 1: Performance Results for LogisticRegression\200samplesn10featuresnreg0.01Problem

Optimizer	Mean Final Value	Std Dev	Best Value	Worst Value	Mean Func Evals	Success Rate (%)	Mean Time (s)
QQN-StrongWolfe	3.23×10^{-1}	1.85×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	80.3	0.0	0.056
QQN-Backtracking	3.23×10^{-1}	1.71×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	79.6	0.0	0.048
QQN-CubicQuadraticInterpolation	3.23×10^{-1}	2.06×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	173.4	0.0	0.133
QQN-GoldenSection	3.23×10^{-1}	2.02×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	359.4	0.0	0.136
QQN-MoreThuente	3.23×10^{-1}	1.81×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	135.1	0.0	0.085
QQN-Bisection-2	3.23×10^{-1}	1.74×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	109.5	0.0	0.062
QQN-Bisection-1	3.23×10^{-1}	2.44×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	117.6	0.0	0.082
L-BFGS-Conservative	3.23×10^{-1}	7.50×10^{-11}	3.23×10^{-1}	3.23×10^{-1}	545.8	0.0	0.320
L-BFGS	3.24×10^{-1}	2.04×10^{-3}	3.23×10^{-1}	3.33×10^{-1}	454.8	0.0	0.252
L-BFGS-Aggressive	3.24×10^{-1}	8.77×10^{-4}	3.23×10^{-1}	3.26×10^{-1}	638.6	0.0	0.266
Adam-Fast	3.25×10^{-1}	2.41×10^{-4}	3.24×10^{-1}	3.25×10^{-1}	76.0	0.0	0.049
GD-Momentum	3.47×10^{-1}	6.05×10^{-4}	3.46×10^{-1}	3.48×10^{-1}	335.0	0.0	0.332
GD-Nesterov	3.47×10^{-1}	1.02×10^{-3}	3.45×10^{-1}	3.49×10^{-1}	335.0	0.0	0.330
Trust Region-Conservative	4.46×10^{-1}	2.79×10^{-3}	4.41×10^{-1}	4.50×10^{-1}	83.0	0.0	0.043
Adam-WeightDecay	4.59×10^{-1}	1.26×10^{-2}	4.34×10^{-1}	4.80×10^{-1}	502.0	0.0	0.324
GD-WeightDecay	4.69×10^{-1}	7.59×10^{-3}	4.53×10^{-1}	4.83×10^{-1}	335.0	0.0	0.332
GD	5.69×10^{-1}	1.17×10^{-2}	5.51×10^{-1}	5.95×10^{-1}	335.0	0.0	0.330
Adam-AMSGrad	5.79×10^{-1}	1.85×10^{-2}	5.44×10^{-1}	6.17×10^{-1}	502.0	0.0	0.325
Adam	5.81×10^{-1}	2.14×10^{-2}	5.30×10^{-1}	6.15×10^{-1}	502.0	0.0	0.326
Trust Region-Adaptive	6.96×10^{-1}	2.75×10^{-2}	6.46×10^{-1}	7.58×10^{-1}	5.0	0.0	0.003
Trust Region-Standard	6.98×10^{-1}	2.71×10^{-2}	6.41×10^{-1}	7.56×10^{-1}	5.0	0.0	0.003