Table 1: Performance Results for LogisticRegression$\backslash_2 00 samples n_1 0 features n_r eg 0.01 Problem$

Optimizer	Mean Final Value	Std Dev	Best Value	Worst Value	Mean Func Evals	Success Rate (%)	Mean Time (s)
QQN-StrongWolfe	3.23×10^{-1}	1.34×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	80.3	0.0	0.059
QQN-Backtracking	3.23×10^{-1}	2.05×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	80.0	0.0	0.051
QQN-MoreThuente	3.23×10^{-1}	1.36×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	139.4	0.0	0.092
QQN-Bisection-2	3.23×10^{-1}	1.54×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	112.3	0.0	0.067
QQN-GoldenSection	3.23×10^{-1}	2.35×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	365.6	0.0	0.140
QQN-CubicQuadraticInterpolation	3.23×10^{-1}	2.49×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	177.6	0.0	0.143
QQN-Bisection-1	3.23×10^{-1}	2.82×10^{-12}	3.23×10^{-1}	3.23×10^{-1}	115.3	0.0	0.083
L-BFGS-Conservative	3.23×10^{-1}	8.34×10^{-11}	3.23×10^{-1}	3.23×10^{-1}	546.0	0.0	0.324
L-BFGS	3.24×10^{-1}	2.41×10^{-3}	3.23×10^{-1}	3.34×10^{-1}	399.0	0.0	0.220
L-BFGS-Aggressive	3.24×10^{-1}	9.27×10^{-4}	3.23×10^{-1}	3.26×10^{-1}	622.6	0.0	0.259
Adam-Fast	3.24×10^{-1}	2.61×10^{-4}	3.24×10^{-1}	3.25×10^{-1}	76.5	0.0	0.050
GD-Momentum	3.47×10^{-1}	6.95×10^{-4}	3.46×10^{-1}	3.48×10^{-1}	335.0	0.0	0.333
GD-Nesterov	3.47×10^{-1}	6.91×10^{-4}	3.46×10^{-1}	3.49×10^{-1}	335.0	0.0	0.334
Trust Region-Conservative	4.44×10^{-1}	2.94×10^{-3}	4.39×10^{-1}	4.50×10^{-1}	85.8	0.0	0.046
Adam-WeightDecay	4.61×10^{-1}	1.35×10^{-2}	4.28×10^{-1}	4.86×10^{-1}	502.0	0.0	0.328
GD-WeightDecay	4.72×10^{-1}	7.20×10^{-3}	4.60×10^{-1}	4.91×10^{-1}	335.0	0.0	0.334
GD	5.75×10^{-1}	1.80×10^{-2}	5.28×10^{-1}	6.13×10^{-1}	335.0	0.0	0.332
Adam	5.75×10^{-1}	1.77×10^{-2}	5.41×10^{-1}	6.07×10^{-1}	502.0	0.0	0.327
Adam-AMSGrad	5.80×10^{-1}	2.46×10^{-2}	5.46×10^{-1}	6.38×10^{-1}	502.0	0.0	0.328
Trust Region-Adaptive	6.99×10^{-1}	2.55×10^{-2}	6.51×10^{-1}	7.55×10^{-1}	5.0	0.0	0.003
Trust Region-Standard	7.05×10^{-1}	2.30×10^{-2}	6.47×10^{-1}	7.39×10^{-1}	5.0	0.0	0.003