Table 1: Performance Results for SVM$\backslash_1 00 samples n_5 features n_C 1 Problem$

Optimizer	Mean Final Value	Std Dev	Best Value	Worst Value	Mean Func Evals	Success Rate (%)	Mean Time (s)
GD-WeightDecay	6.43×10^{-1}	7.06×10^{-7}	6.43×10^{-1}	6.43×10^{-1}	176.7	100.0	0.095
L-BFGS	6.43×10^{-1}	4.38×10^{-6}	6.43×10^{-1}	6.43×10^{-1}	346.1	100.0	0.083
QQN-Backtracking	6.43×10^{-1}	5.36×10^{-7}	6.43×10^{-1}	6.43×10^{-1}	354.4	100.0	0.087
L-BFGS-Conservative	6.43×10^{-1}	9.99×10^{-6}	6.43×10^{-1}	6.43×10^{-1}	420.9	100.0	0.102
L-BFGS-Aggressive	6.43×10^{-1}	1.55×10^{-4}	6.43×10^{-1}	6.44×10^{-1}	138.3	97.5	0.031
QQN-MoreThuente	6.43×10^{-1}	9.92×10^{-6}	6.43×10^{-1}	6.43×10^{-1}	409.4	50.0	0.168
QQN-GoldenSection	6.43×10^{-1}	2.58×10^{-5}	6.43×10^{-1}	6.43×10^{-1}	410.1	10.0	0.079
QQN-StrongWolfe	6.43×10^{-1}	5.20×10^{-6}	6.43×10^{-1}	6.43×10^{-1}	396.6	5.0	0.129
GD	6.43×10^{-1}	1.28×10^{-5}	6.43×10^{-1}	6.43×10^{-1}	334.6	2.5	0.180
Adam-Fast	6.82×10^{-1}	2.53×10^{-2}	6.43×10^{-1}	7.45×10^{-1}	34.7	2.5	0.012
QQN-Bisection-1	6.43×10^{-1}	7.55×10^{-6}	6.43×10^{-1}	6.43×10^{-1}	467.2	0.0	0.187
QQN-Bisection-2	6.43×10^{-1}	1.56×10^{-5}	6.43×10^{-1}	6.43×10^{-1}	229.9	0.0	0.074
QQN-CubicQuadraticInterpolation	6.43×10^{-1}	5.12×10^{-6}	6.43×10^{-1}	6.43×10^{-1}	396.2	0.0	0.197
Adam-WeightDecay	6.44×10^{-1}	9.10×10^{-4}	6.43×10^{-1}	6.47×10^{-1}	502.0	0.0	0.175
GD-Nesterov	6.62×10^{-1}	5.38×10^{-3}	6.52×10^{-1}	6.70×10^{-1}	24.8	0.0	0.013
GD-Momentum	6.74×10^{-1}	6.54×10^{-3}	6.66×10^{-1}	6.92×10^{-1}	25.1	0.0	0.013
Adam-AMSGrad	7.24×10^{-1}	4.19×10^{-2}	6.61×10^{-1}	8.10×10^{-1}	502.0	0.0	0.175
Adam	7.39×10^{-1}	5.10×10^{-2}	6.51×10^{-1}	8.40×10^{-1}	502.0	0.0	0.174
Trust Region-Conservative	3.09	1.13	6.46×10^{-1}	3.97	151.9	0.0	0.041
Trust Region-Adaptive	6.69	9.44	1.14	3.52×10^{1}	20.7	0.0	0.006
Trust Region-Standard	6.82×10^{2}	1.60×10^{3}	6.89×10^{-1}	5.00×10^{3}	20.1	0.0	0.006