

Table 1: Performance Results for SVM\200samplesn10featuresn_C1Problem

Optimizer	Mean Final Value	Std Dev	Best Value	Worst Value	Mean Func Evals	Success Rate (%)	Mean Time (s)
GD-WeightDecay	6.86×10^{-1}	7.74×10^{-7}	6.86×10^{-1}	6.86×10^{-1}	172.5	0.0	0.168
QQN-Backtracking	6.86×10^{-1}	2.04×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	704.7	0.0	0.318
QQN-StrongWolfe	6.86×10^{-1}	4.49×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	544.7	0.0	0.334
QQN-MoreThuente	6.86×10^{-1}	4.88×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	441.9	0.0	0.339
QQN-Bisection-1	6.86×10^{-1}	5.28×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	474.2	0.0	0.338
L-BFGS-Conservative	6.86×10^{-1}	4.84×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	639.5	0.0	0.268
QQN-Bisection-2	6.86×10^{-1}	1.08×10^{-5}	6.86×10^{-1}	6.86×10^{-1}	367.1	0.0	0.200
QQN-CubicQuadraticInterpolation	6.86×10^{-1}	5.18×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	394.2	0.0	0.355
QQN-GoldenSection	6.86×10^{-1}	8.54×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	712.9	0.0	0.242
GD	6.86×10^{-1}	1.17×10^{-5}	6.86×10^{-1}	6.86×10^{-1}	333.8	0.0	0.327
L-BFGS-Aggressive	6.86×10^{-1}	1.43×10^{-5}	6.86×10^{-1}	6.86×10^{-1}	133.8	0.0	0.055
Adam-WeightDecay	6.86×10^{-1}	4.23×10^{-4}	6.86×10^{-1}	6.87×10^{-1}	493.4	0.0	0.311
L-BFGS	7.00×10^{-1}	4.33×10^{-2}	6.86×10^{-1}	8.73×10^{-1}	276.4	0.0	0.124
GD-Nesterov	7.01×10^{-1}	3.71×10^{-3}	6.93×10^{-1}	7.08×10^{-1}	25.6	0.0	0.024
GD-Momentum	7.08×10^{-1}	5.43×10^{-3}	7.00×10^{-1}	7.20×10^{-1}	25.7	0.0	0.024
Adam-Fast	7.18×10^{-1}	2.35×10^{-2}	6.86×10^{-1}	7.55×10^{-1}	44.4	0.0	0.028
Adam-AMSGrad	7.41×10^{-1}	2.79×10^{-2}	7.00×10^{-1}	8.12×10^{-1}	502.0	0.0	0.318
Adam	7.46×10^{-1}	2.91×10^{-2}	7.00×10^{-1}	8.10×10^{-1}	502.0	0.0	0.316
Trust Region-Standard	1.05	5.60×10^{-1}	6.93×10^{-1}	3.44	14.9	0.0	0.008
Trust Region-Conservative	2.85	1.27	6.88×10^{-1}	3.96	140.3	0.0	0.071
Trust Region-Adaptive	8.00×10^1	2.71×10^2	7.03×10^{-1}	1.25×10^3	24.4	0.0	0.012