

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}	8.95×10^{-7}	6.86×10^{-1}	6.86×10^{-1}	178.7	0.0	0.181
QQN-StrongWolfe	6.86×10^{-1}	2.94×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	551.6	0.0	0.348
QQN-Backtracking	6.86×10^{-1}	2.72×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	679.4	0.0	0.307
QQN-Bisection-1	6.86×10^{-1}	6.62×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	478.8	0.0	0.356
L-BFGS-Conservative	6.86×10^{-1}	7.92×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	555.9	0.0	0.246
QQN-MoreThuente	6.86×10^{-1}	6.37×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	443.4	0.0	0.344
QQN-Bisection-2	6.86×10^{-1}	1.30×10^{-5}	6.86×10^{-1}	6.86×10^{-1}	324.1	0.0	0.187
QQN-GoldenSection	6.86×10^{-1}	6.09×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	672.1	0.0	0.234
QQN-CubicQuadraticInterpolation	6.86×10^{-1}	5.45×10^{-6}	6.86×10^{-1}	6.86×10^{-1}	395.0	0.0	0.362
GD	6.86×10^{-1}	1.26×10^{-5}	6.86×10^{-1}	6.86×10^{-1}	334.8	0.0	0.342
L-BFGS-Aggressive	6.86×10^{-1}	1.47×10^{-5}	6.86×10^{-1}	6.86×10^{-1}	135.1	0.0	0.057
Adam-WeightDecay	6.86×10^{-1}	4.19×10^{-4}	6.86×10^{-1}	6.88×10^{-1}	481.4	0.0	0.317
L-BFGS	6.87×10^{-1}	1.75×10^{-3}	6.86×10^{-1}	6.94×10^{-1}	269.8	0.0	0.125
GD-Nesterov	7.02×10^{-1}	3.42×10^{-3}	6.96×10^{-1}	7.09×10^{-1}	25.7	0.0	0.025
Adam-Fast	7.10×10^{-1}	2.29×10^{-2}	6.86×10^{-1}	7.63×10^{-1}	55.8	0.0	0.036
GD-Momentum	7.11×10^{-1}	6.48×10^{-3}	7.01×10^{-1}	7.29×10^{-1}	25.8	0.0	0.025
Adam-AMSGrad	7.37×10^{-1}	3.08×10^{-2}	6.92×10^{-1}	7.99×10^{-1}	502.0	0.0	0.328
Adam	7.48×10^{-1}	3.72×10^{-2}	6.88×10^{-1}	8.32×10^{-1}	502.0	0.0	0.329
Trust Region-Standard	9.97×10^{-1}	6.11×10^{-1}	6.96×10^{-1}	3.60	21.9	0.0	0.012
Trust Region-Conservative	2.67	1.46	6.89×10^{-1}	4.04	128.8	0.0	0.067
Trust Region-Adaptive	8.45	1.29×10^1	7.44×10^{-1}	3.47×10^1	21.8	0.0	0.011