Table 1: Optimizer Family vs Problem Family Performance Matrix  $\,$ 

Problem Family	Adam	GD	L-BFGS	QQN	Trust Region
Ackley	16.3 / 12.0 Adam-AMSGrad Adam-Fast	16.5 / 9.7 GD GD-Momentum	6.5 / 2.3 L-BFGS Conservative	5.1 / 1.0  Bisection-2  GoldenSection	20.7 / 13.7 Conservative Aggressive
Barrier	$8.5 \ / \ 3.0$ WeightDecay Adam-Fast	$6.2 \ / \ 1.0$ GD AdaptiveMom	3.7 / 2.3 Conservative L-BFGS-Limited	inf / inf N/A N/A	13.1 / 8.0 Conservative Aggressive
Beale	19.0 / 8.0 WeightDecay Adam-Fast	8.8 / 3.0 GD-Nesterov GD-Momentum	10.0 / 2.0  MoreThuente  Aggressive	8.8 / 1.0 GoldenSection Bisection-2	18.4 / 15.0  Precise  Standard
Booth	19.2 / 11.0 WeightDecay Adam-Robust 13.2 / 10.0	14.6 / 10.0 GD GD-Momentum 15.2 / 9.0	11.0 / 6.0  MoreThuente  Aggressive  10.4 / 5.0	3.0 / 1.0 CubicQuadIn GoldenSection 3.2 / 1.0	17.2 / 12.0 Adaptive Conservative 23.0 / 21.0
$\operatorname{GoldsteinPrice}$	Adam-AMSGrad Adam-Fast 17.7 / 12.0	GD-Momentum $GD$ $12.0 / 7.7$	MoreThuente Aggressive 7.9 / 3.7	GoldenSection Bisection-2 6.3 / 1.0	Aggressive Precise 21.1 / 13.7
Griewank	Adam-Fast Adam-Robust 18.8 / 11.0	GD-Momentum GD 14.6 / 9.0	Aggressive L-BFGS 11.2 / 5.0	StrongWolfe CubicQuadIn 3.4 / 1.0	Conservative Aggressive 17.0 / 8.0
Himmelblau	WeightDecay Adam-Robust 14.2 / 9.0	GD AdaptiveMom $12.5 / 7.0$	L-BFGS-Limited Aggressive 12.5 / 4.7	GoldenSection Bisection-1 4.1 / 1.7	Adaptive Conservative 21.8 / 16.7
Ill Conditioned Rosenbrock	WeightDecay Adam-Robust 14.4 / 11.0	GD-Nesterov GD-Momentum 14.6 / 9.0	MoreThuente Aggressive 11.6 / 3.0	CubicQuadIn GoldenSection 3.8 / 1.0	Aggressive Conservative 20.6 / 13.0
Levi	Adam-Robust Adam-Fast 15.9 / 10.0	GD-Momentum GD-WeightDecay 16.2 / 7.7	L-BFGS-Limited Aggressive 9.1 / 6.7	GoldenSection Bisection-1 3.0 / 1.0	Conservative Aggressive 20.8 / 16.0
Levy	WeightDecay Adam-AMSGrad 13.2 / 10.0	GD-WeightDecay AdaptiveMom 16.0 / 12.0	Conservative Aggressive 8.8 / 3.0	Bisection-2 GoldenSection 4.0 / 1.0	Conservative Aggressive 23.0 / 21.0
Matyas	Adam-Fast Adam-AMSGrad 6.2 / 1.0	GD-Momentum AdaptiveMom 12.1 / 6.7	L-BFGS Aggressive 14.3 / 7.0	StrongWolfe Bisection-1 11.9 / 6.7	Conservative Precise 20.5 / 16.3
Michalewicz	Adam Adam-Robust 9.1 / 3.5	AdaptiveMom GD-WeightDecay 19.4 / 16.0	MoreThuente Aggressive 11.3 / 8.0	Bisection-2 CubicQuadIn 3.6 / 1.0	Conservative Aggressive 21.6 / 18.5
Neural Networks	WeightDecay Adam-Robust 16.9 / 8.7	GD-WeightDecay AdaptiveMom 8.8 / 5.3	Conservative L-BFGS 7.4 / 1.0	Bisection-2 StrongWolfe 9.9 / 2.7	Adaptive Aggressive 19.9 / 16.3
NoisySphere	Adam-Fast Adam 8.1 / 4.3	AdaptiveMom GD-WeightDecay 12.3 / 9.7	Conservative Aggressive 14.3 / 5.7	StrongWolfe CubicQuadIn 7.5 / 1.0	Conservative Precise 22.9 / 20.7
PenaltyI	Adam-AMSGrad Adam-Fast 11.4 / 4.7	GD GD-Nesterov $14.2 \ / \ 7.7$	Conservative Aggressive 14.1 / 3.7	CubicQuadIn Bisection-2 9.9 / 3.0	Adaptive Precise 15.4 / 7.0
Rastrigin	Adam-AMSGrad Adam-Fast	GD-WeightDecay GD-Momentum	MoreThuente Aggressive	CubicQuadIn Bisection-2	Adaptive Conservative

Continued on next page

Table 1 – continued from previous page

Problem Family	Adam	GD	L-BFGS	QQN	Trust Region
	$18.5 \ / \ 13.2$	$13.6 \ / \ 8.2$	8.9 / 4.8	3.4 / 1.0	20.6 / 17.2
Regression	Adam-Fast	${\bf Adaptive Mom}$	Conservative	Bisection-1	Adaptive
	Adam-Robust	GD	L-BFGS	GoldenSection	Conservative
	13.4 / 6.0	12.1 / 5.0	12.6 / 4.0	4.9 / 2.0	22.0 / 17.7
Rosenbrock	Adam-Fast	GD-Nesterov	MoreThuente	StrongWolfe	Aggressive
	Adam-Robust	GD-Momentum	Aggressive	GoldenSection	Conservative
CITIZA	13.5 / 8.5	13.9 / 5.0	$9.6 \ / \ 3.0$	6.3 / 2.5	21.7 / 17.0
$\mathbf{SVM}$	WeightDecay	GD-WeightDecay	Conservative	StrongWolfe	Conservative
	Adam-Fast	AdaptiveMom	L-BFGS-Limited	GoldenSection	Aggressive
C -1 f-1	20.1 / 10.7	10.5 / 6.7	10.5 / 4.3	4.3 / 1.0	19.7 / 16.7
Schwefel	Adam-Fast	GD-WeightDecay	Conservative	StrongWolfe	Standard
	Adam-Robust	$^{ m GD}$ $14.5 \ / \ 10.5$	Aggressive $6.4 / 1.5$	GoldenSection 4.9 / 1.5	Conservative
SparseQuadratic	18.9 / 12.5 WeightDecay	GD-WeightDecay	MoreThuente	4.9 / 1.9 GoldenSection	20.3 / 14.5 Precise
Sparse Quadratic	Adam-AMSGrad	AdaptiveMom	L-BFGS	Bisection-1	Aggressive
	12.9 / 8.0	11.8 / 6.0	15.2 / 4.5	3.7 / 1.0	21.4 / 19.0
SparseRosenbrock	Adam-Fast	GD-Nesterov	L-BFGS-Limited	CubicQuadIn	Standard
Sparsertosenbrock	Adam-Robust	GD-Nesterov GD-Momentum	Aggressive	Bisection-2	Conservative
	20.1 / 14.5	13.9 / 10.0	6.1 / 1.0	5.3 / 3.0	19.6 / 14.0
Sphere	WeightDecay	GD-Momentum	Aggressive	StrongWolfe	Conservative
P	Adam-AMSGrad	AdaptiveMom	Conservative	GoldenSection	Aggressive
	16.6 / 4.3	14.2 / 7.3	10.3 / 2.3	8.7 / 1.7	15.3 / 4.3
StyblinskiTang	WeightDecay	GD-WeightDecay	Conservative	GoldenSection	Standard
Ţ G	Adam-Robust	AdaptiveMom	Aggressive	StrongWolfe	Conservative
	12.7 / 7.3	14.3 / 5.0	12.7 / 5.3	4.0 / 1.0	21.3 / 17.7
Trigonometric	Adam	GD	MoreThuente	CubicQuadIn	Precise
	Adam-Fast	GD-Momentum	Aggressive	Bisection-2	Aggressive
	13.4 / 9.3	14.7 / 7.3	11.7 / 6.0	3.0 / 1.0	22.2 / 19.0
Zakharov	WeightDecay	$\operatorname{GD}$	MoreThuente	GoldenSection	Adaptive
	Adam-Robust	AdaptiveMom	L-BFGS	StrongWolfe	Conservative

Legend: Each cell contains:

- Top line: Average Ranking / Best Rank Average (lower is better)
- Middle line: Best performing variant in this optimizer family
- Bottom line: Worst performing variant in this optimizer family

Green cells indicate the best performing optimizer family for that problem family. Red cells indicate the worst performing optimizer family.