Comparison Training Loss Over Iterations
300 samples, 2D points, 5 clusters, 0 sparse dimensions