

KL Divergence During Reasoning Training on MATH-500

