

Model performance across benchmarks with different freezing strategies

