

Figure 1: Comparison of RoBERTa model variants on the MRPC dataset across training epochs. full represents the original RoBERTa model. $r[x]$ represents RoBERTa model with injected LoRA layer with rank= α = x .

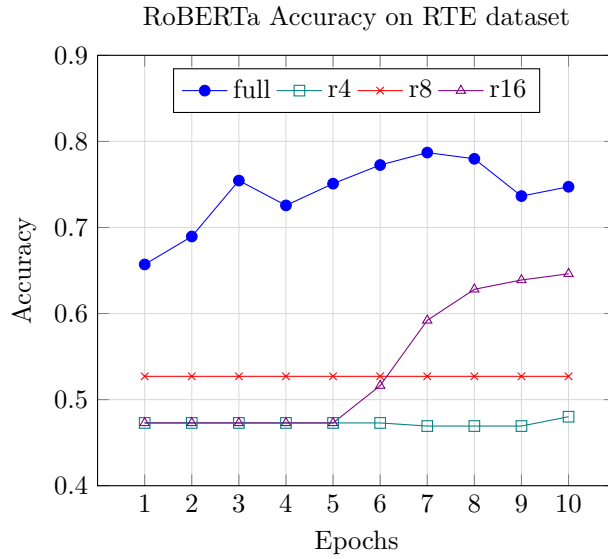


Figure 2: Comparison of RoBERTa model variants on the RTE dataset across training epochs.

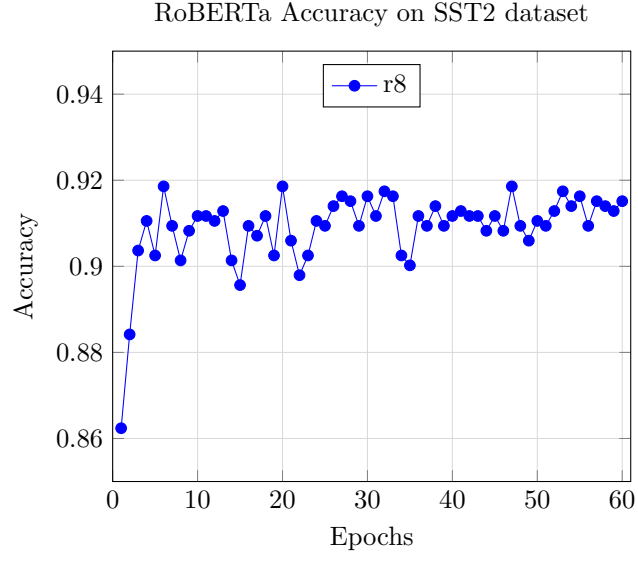


Figure 3: RoBERTa model on the SST2 dataset across 60 training epochs with injected LoRA layer of rank= α =8.

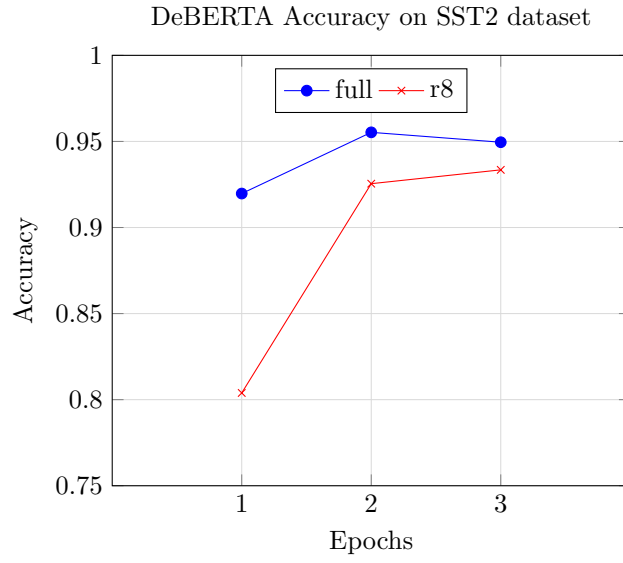


Figure 4: DeBERTa model on the SST2 dataset across 3 training epochs with injected LoRA layer of rank= α =8 compared to the original model.