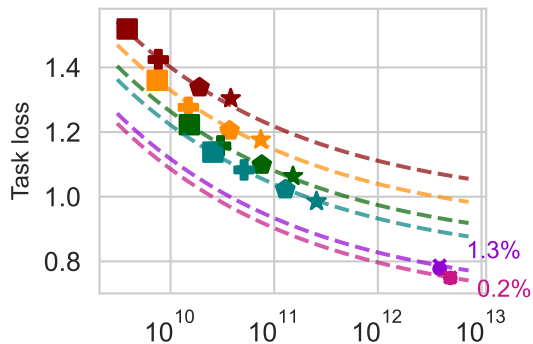
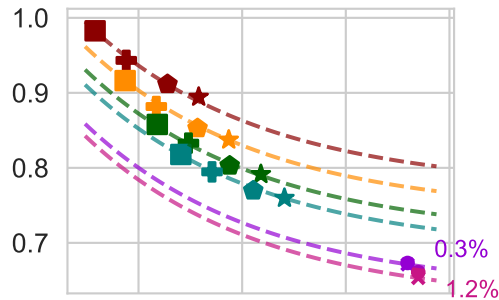


--- 190M (fitted)
 --- 370M (fitted)
 --- 760M (fitted)
 --- 1.3B (fitted)
 ● 7B-4T (target)
 × 7B-4T (predicted)
 ● 13B-5T (target)
 × 13B-5T (predicted)

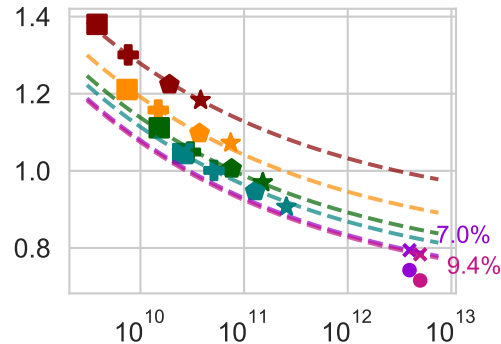
MMLU (Fitting error: 0.46%)



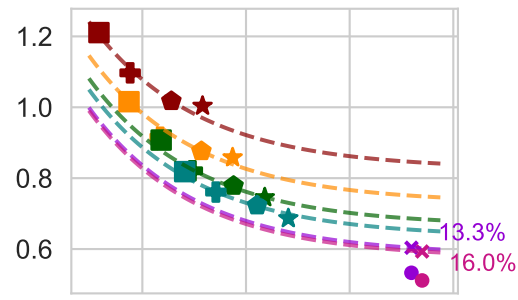
HellaSwag (Fitting error: 0.26%)



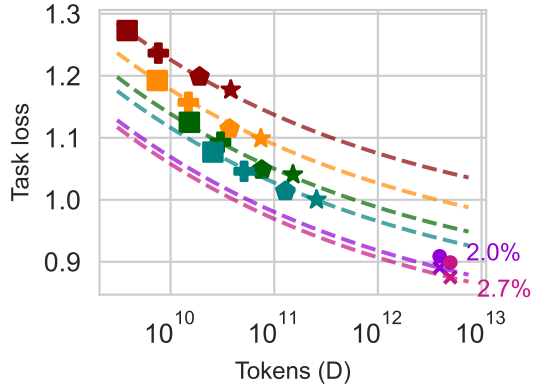
ARC-Challenge (Fitting error: 0.48%)



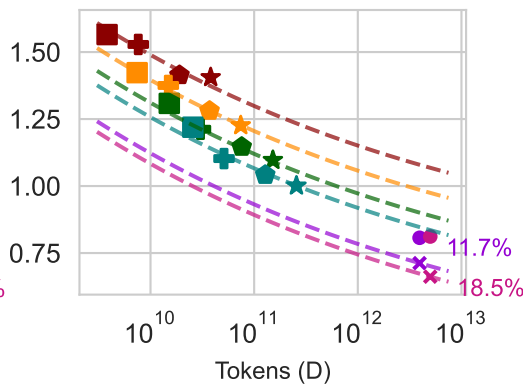
ARC-Easy (Fitting error: 1.18%)



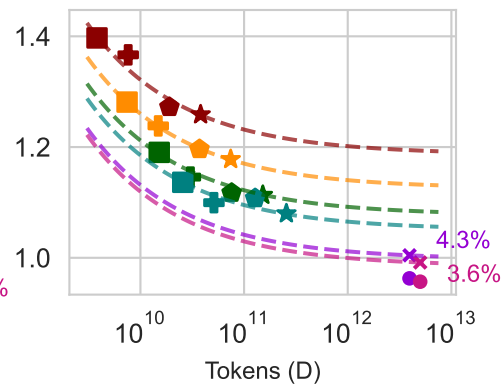
PIQA (Fitting error: 0.31%)



CommonsenseQA (Fitting error: 1.05%)



Social IQa (Fitting error: 0.54%)



OpenBookQA (Fitting error: 0.59%)

