

Distribution of lengths of wrong predictions of few shot test prompting

