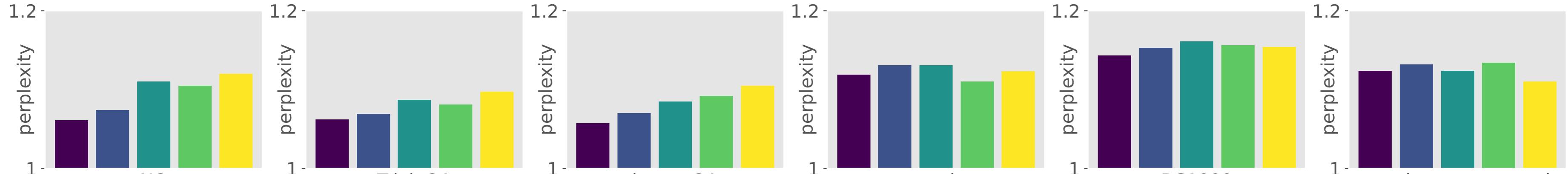


(1). perplexity of GPT-3.5 on six datasets



(2). perplexity of Llama2-13b on six datasets

oracle distracting random irrelevant_dummy irrelevant_diff