

Conversation History Task

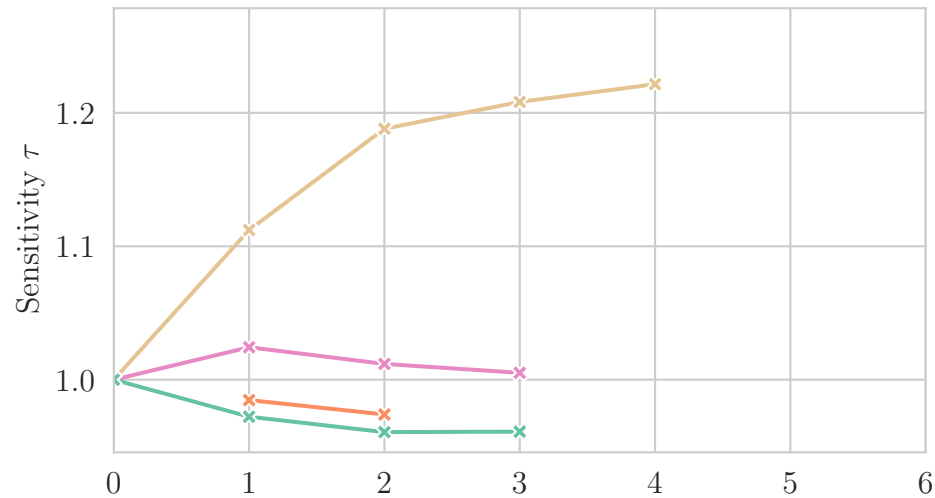
—x— Gigaword

—x— Rotten Tomatoes

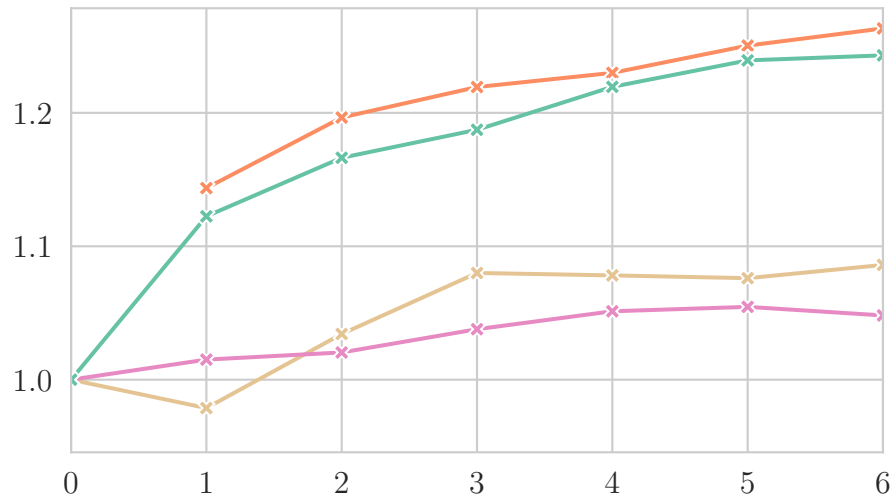
—x— MMLU Human Aging

—x— TweetQA

LLAMA-7B



Mistral-7B



Conversation History Length L