

# Conversation History Task

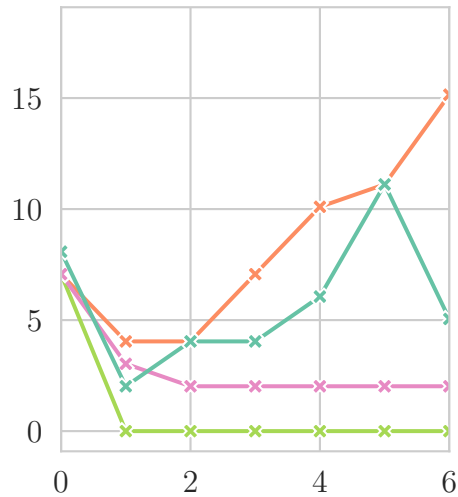
—x— Gigaword

—x— Rotten Tomatoes

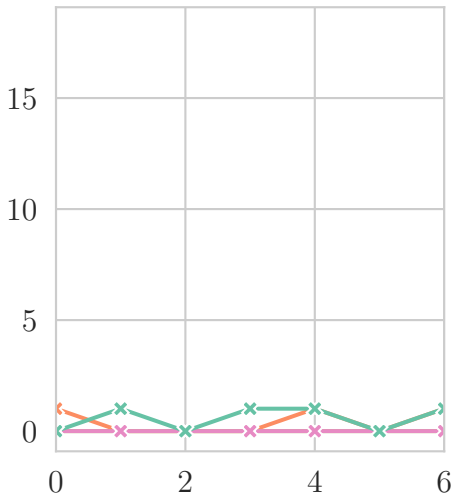
—x— MMLU Abstract Algebra

—x— TweetQA

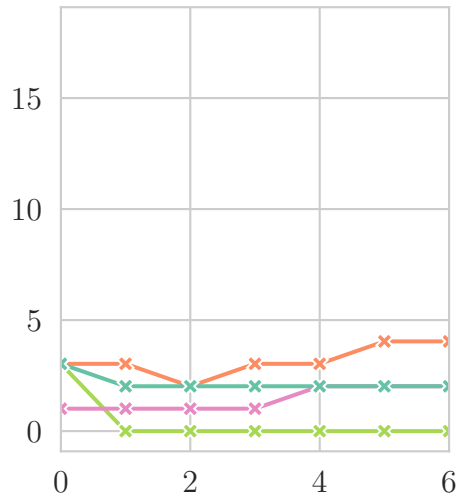
## GPT-3.5



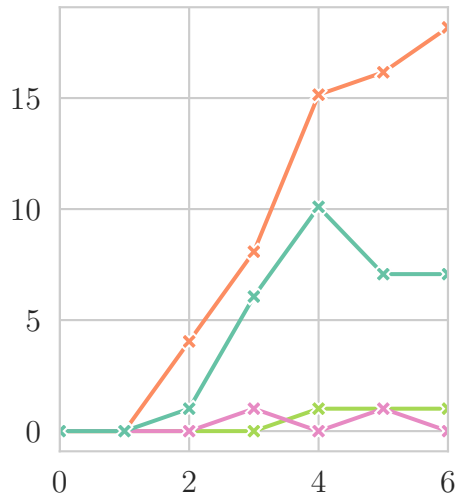
## GPT-4



## LLAMA-7B



## Mistral-7B



Conversation History Length  $L$