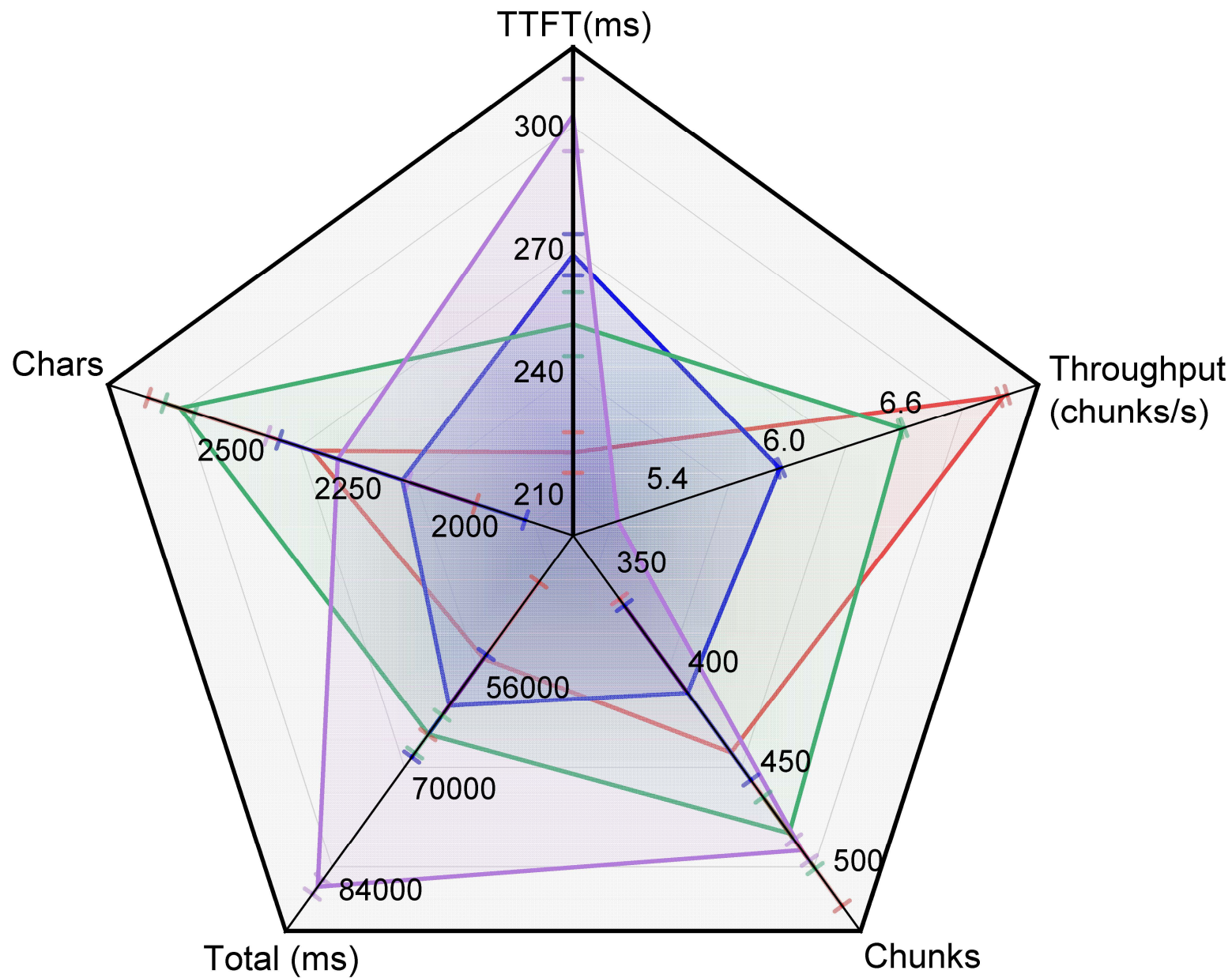
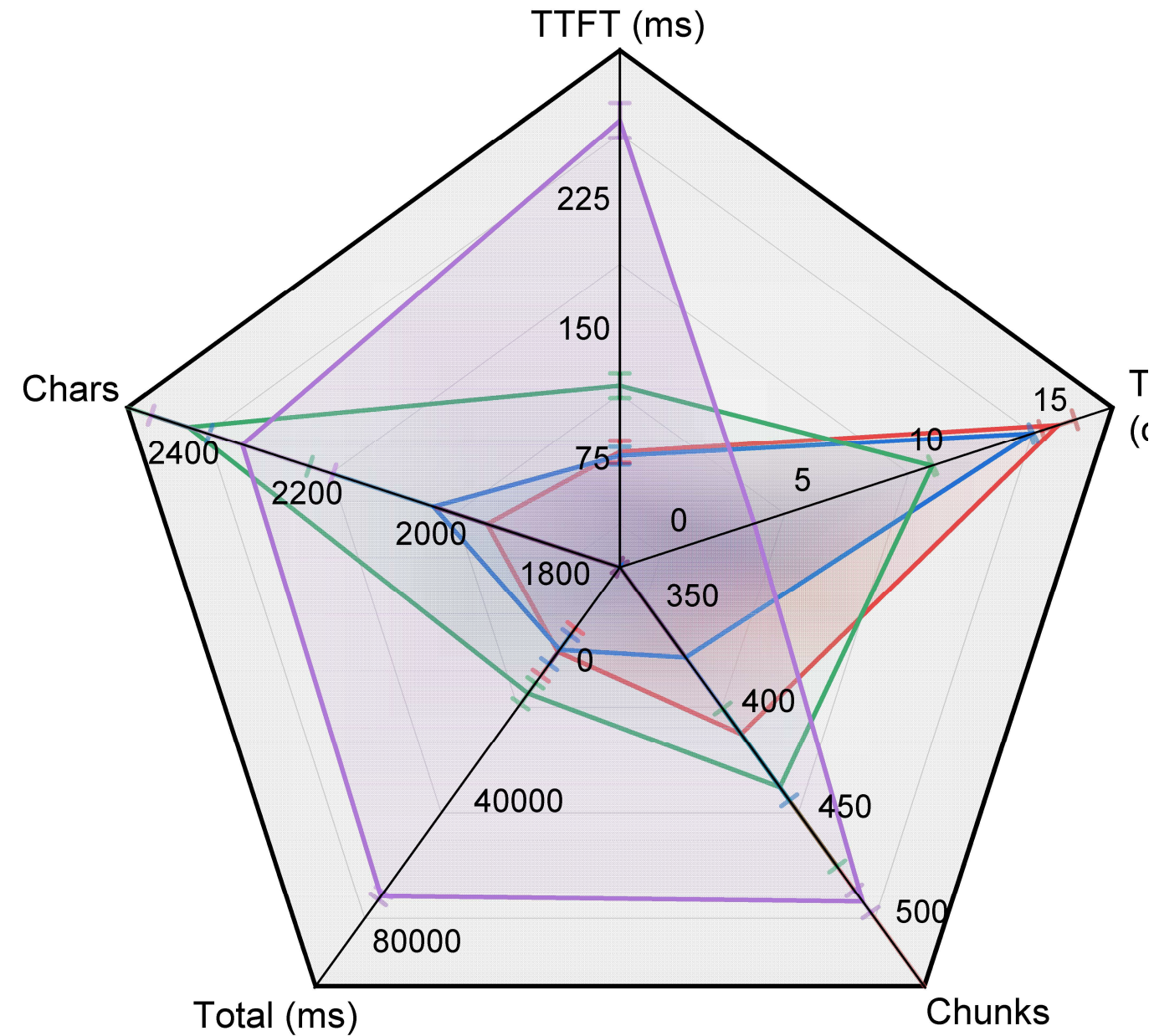


■ Gemma-2-2B 
 ■ Llama-3.2-3B 
 ■ Qwen-3.5-3B 
 ■ Phi-3.5-mini



Evaluation on 2B-3B Scale Models

■ Gemma-3-1B 
 ■ Llama-3.2-1B 
 ■ Qwen-2.5-1.5B 
 ■ DeepSeek



Evaluation on 1B-1.5B Scale Models