Legend:
- LTFilter+LTTopK (RankZephyr)
- LTFilter+LMPQ (RankZephyr)
- LTTopK (RankZephyr)
- LMPQ (RankZephyr)
- Pairwise (RankZephyr)
- LOTUS (QWEN3-8B)
- LOTUS (Zephyr-7B)
- LMPQ (Qwen3-8b)
- LMPQ (Zephyr7b)
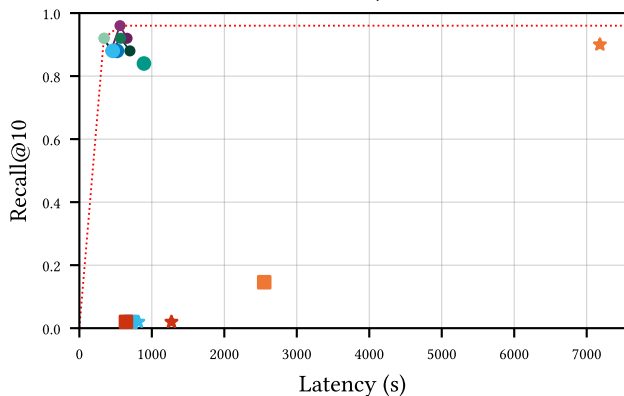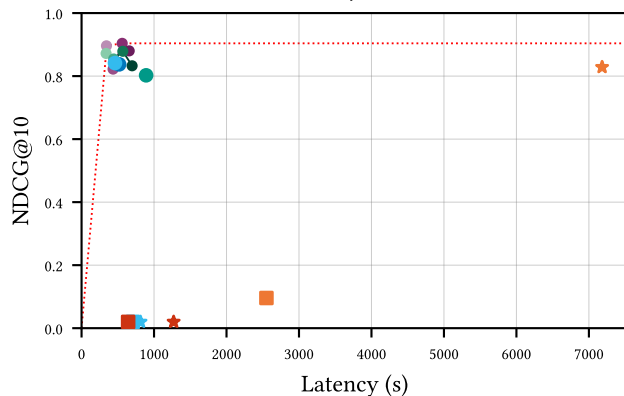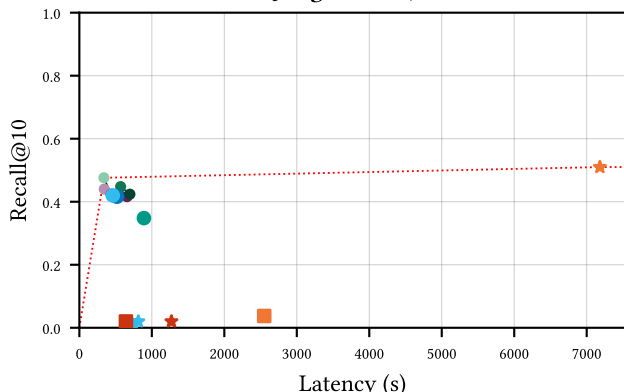- Pointwise (QWEN3-8B)
- Pointwise (Zephyr-7B)

Human Labels, Selection

Human Labels, Selection+Sort

LLM-as-Judge Labels, Selection

LLM-as-Judge Labels, Selection+Sort