Legend: LTTopK (RankZephyr), LMPQ (RankZephyr), LMPQ (Zephyr7b), LOTUS (QWEN3-8B), Pointwise (QWEN3-8B), LTFilter+LTTopK (S=1..15), LTFilter+LMPQ (S=1..15), LMPQ (Qwen3-8b), LOTUS (Zephyr-7B), Pointwise (Zephyr-7B), Pairwise (RankZephyr)
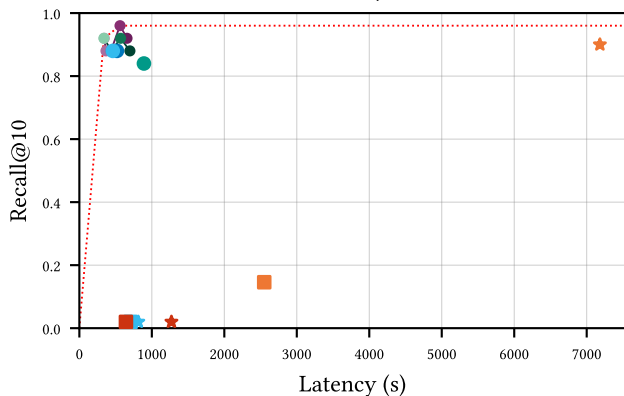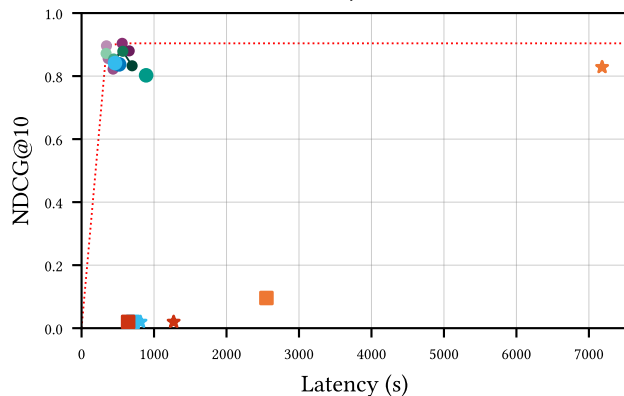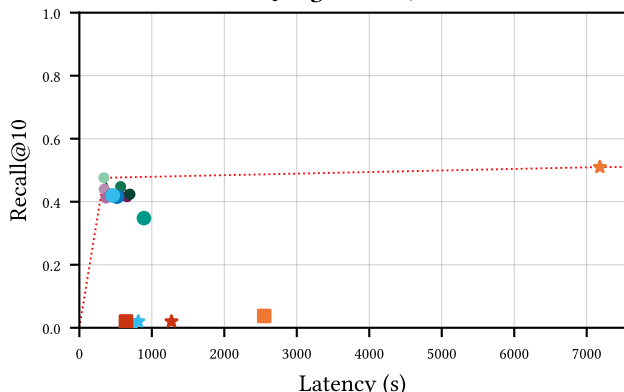
Human Labels, Selection

Human Labels, Selection+Sort

LLM-as-Judge Labels, Selection

LLM-as-Judge Labels, Selection+Sort