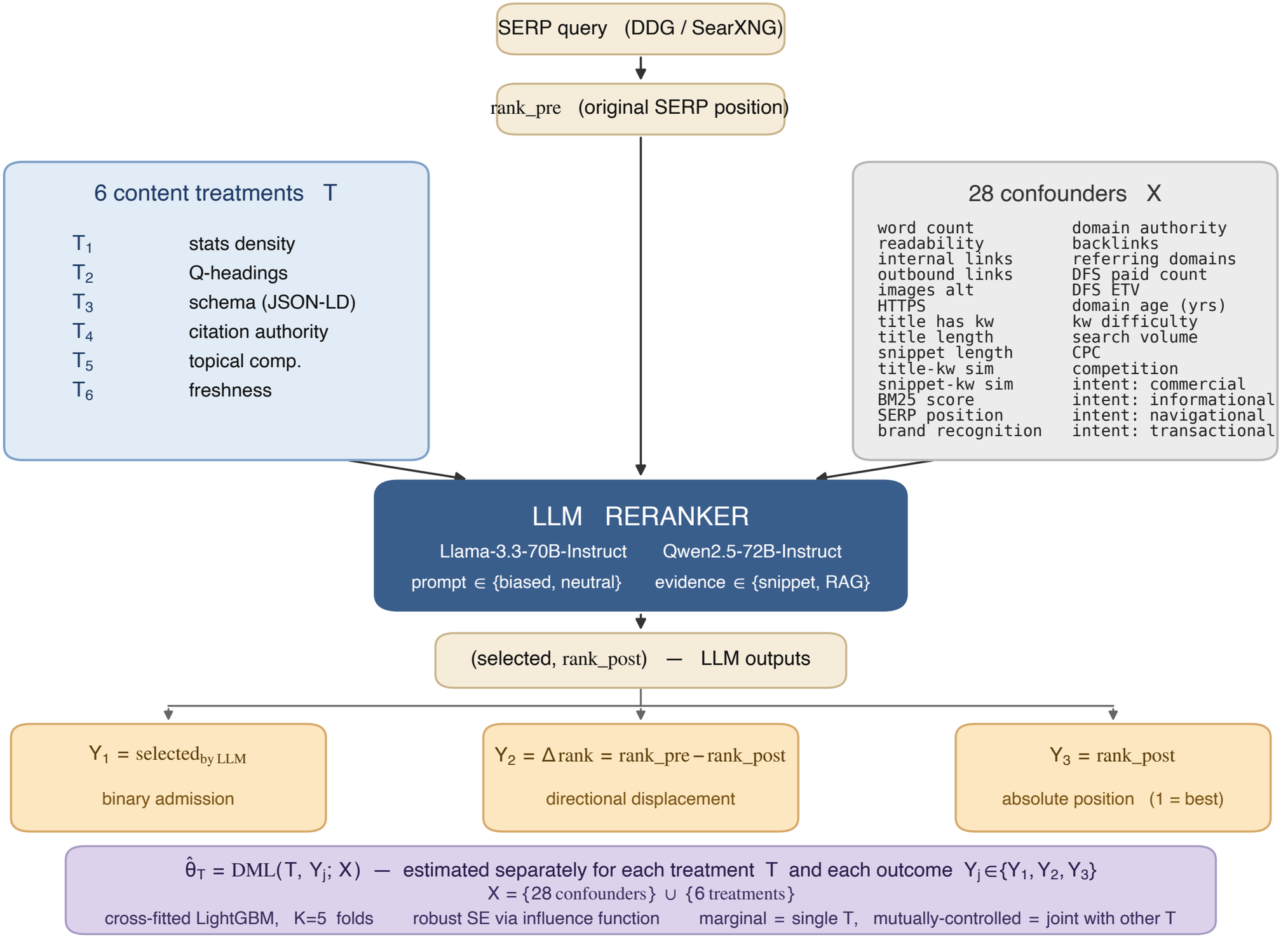
SERP query (DDG / SearXNG)
rank_pre (original SERP position)
6 content treatments T
T1 stats density
T2 Q-headings
T3 schema (JSON-LD)
T4 citation authority
T5 topical comp.
T6 freshness
28 confounders X
word count
readability
internal links
outbound links
images alt
HTTPS
title has kw
title length
snippet length
title-kw sim
snippet-kw sim
BM25 score
SERP position
brand recognition
domain authority
backlinks
referring domains
DFS paid count
DFS ETV
domain age (yrs)
kw difficulty
search volume
CPC
competition
intent: commercial
intent: informational
intent: navigational
intent: transactional
LLM RERANKER
Llama-3.3-70B-Instruct
Qwen2.5-72B-Instruct
prompt ∈ {biased, neutral}
evidence ∈ {snippet, RAG}
(selected, rank_post) — LLM outputs
Y1 = selected by LLM
binary admission
Y2 = Δrank = rank_pre − rank_post
directional displacement
Y3 = rank_post
absolute position (1 = best)
θ̂T = DML(T, Yj; X) — estimated separately for each treatment T and each outcome Yj ∈ {Y1, Y2, Y3}
X = {28 confounders} ∪ {6 treatments}
cross-fitted LightGBM, K=5 folds
robust SE via influence function
marginal = single T, mutually-controlled = joint with other T