Owen Du, Philipp Barski, Georg Ye, Fabian Amherd

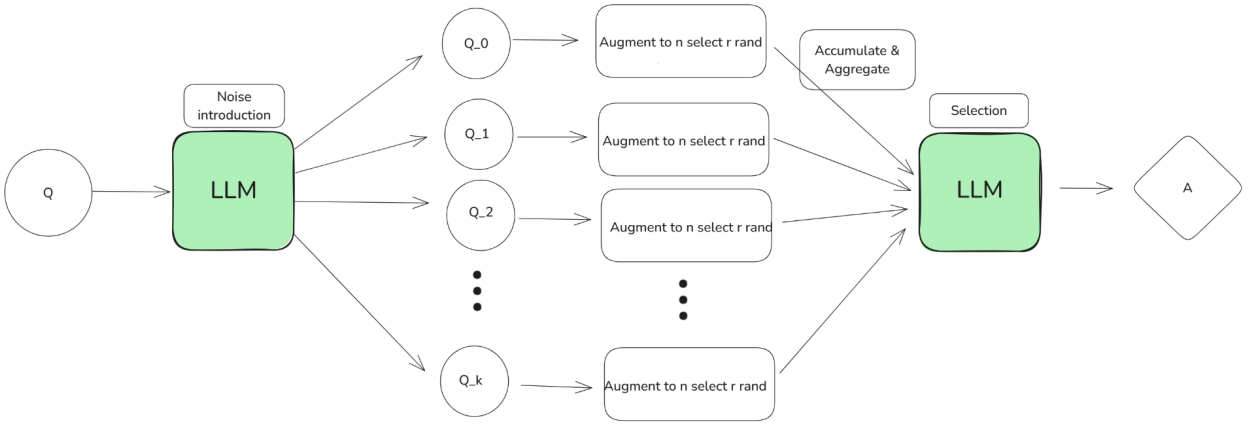
Final Report Datathon 2025 Sentient Challenge

Our implementation

Main contributions: fixed bugs in *crawl4ai_scraper.py* (missing playwright install caused the exception block to activate for every call, deprecated call to markdown_v2 instead of markdown in *extract()*, we did NOT change the behavior of the scraper but just fixed the bug so the normal pro workflow (which we should use) works as intended) that allows the calls for *filter_content_content* to go through. Experiments with a reasoning agent that provides the CodeAgent with a coherent chain-of-thought to guide in hard queries. Prompt engineering for reasoning agent system prompt (performance was similar without reasoning agent, we found that CodeAgent sometimes ignored instructions from reasoning agent). Usage of LLM extractor to extract information from scraped raw text.

Changes to the model & Future directions

We noticed some performance barriers in the model concerning “local” accuracy and exhaustive use of memory resources in the Open Search Tool. In the current implementation we do for each query to the Open Search Tool *num_results* (8) requests from Serper and choose the first *max_results* (2) requests for further processing. The current method does not make use of all requests and is too constrained to the top requests, so we propose a better way of dealing with them (See next image).



Instead of directly transforming the query Q into n requests and further processing only r of them, we propose to get more accurate results (in overhead just introduced by the LLM model if parallelized correctly!) by producing k + 1 “similar” queries Q_0,..., Q_k and performing the upper process separately for each of the queries Q_0,...,Q_k by further *randomly* selecting r request to process introducing a wider range of different internet sources and selecting the dominant results (rather an aggregation of the dominant results) using a final LLM layer to get a more accurate answer in the end.