

Finance Data in Near-real time

Client: Financial research company signalling investment options

Domain: Finance

Challenge:

The client wanted to blend market and social data in order to build a financial recommendation engine that could deliver proprietary signals for systematic equity investing. It had a subscribed audience base to which it would send these trade signals on a daily basis. In order to get this solution running, the company wanted all the possible data from finance that was hot on news, blogs, articles and social media. Such data would facilitate the intelligence that the client was building internally in order to generate the final output on a customized basis. Scale was the primary attribute in picture because the number of sources involved was huge- in the range of 30,000 + websites. Availability of the data and its coverage were important too given the dynamism in the finance market.

The Solution:

PromptCloud's DaaS platform along with its mass scale and low-latency crawl offering was used to address this problem. The system was tuned to fit into this scale and adaptively crawl sources based on the ones that were active vs. ones that were dormant. Alerts were employed in order to notify about dead sources so that crawl results were accurate and the whole system was more efficient. In order to address the low latency requirement of 5-10 minutes, few components were added that could live up to such computation power. The crawled data was indexed using hosted indexing component and a search API was provided that the client could query every few minutes to get the results. Final results were in JSON format.

Benefits to the client:

- 100% API availability and continuous data feeds
- Dynamic list of keywords and sources
- Zero data processing efforts at client's end
- Scalable infrastructure reduced client's costs
- Client's analysts only focused on querying final datasets and running analyses