Financial Services Firm Heavily Automates Data Mastering & Classification Using Tamr

Traditional Approaches To Data Mastering & Classification Have Speed and Scale Limitations

One of Tamr's most valuable clients is a global financial services firm that provides broad financial solutions to individuals and organizations around the world. The enterprise has numerous divisions from personal banking to investment banking.

The Investment Bank division of the company is a major player in the industry and offers clients a wide range of services - from advisory on mergers and acquisitions to financing consultation. One of the primary services of the division is oriented around investment research - providing clients with investment insight based on structured, quantitative research.

One key output of the investment research group are datasets that inform detailed research reports pertaining to potential investment opportunities. In order to surface these insights, the group pulls copious amounts of information on markets and market participants from a variety of sources - both internal and external to the organization. In order to stitch together this data, the group had been relying heavily on programmatic and rules-based methods. These methods would allow the team to master and classify the entities within the collected data - identifying and categorizing each unique company and product in the data they processed - for consumption downstream by analysts all around the world. Unfortunately, this manual approach was proving to be too slow and not scalable enough to satisfy the business' appetite for complete, clean datasets. Given the volume of research reports being generated, the group needed a more effective and efficient way to conduct ad-hoc mastering and classification to accelerate this research cycle.

A manual approach to mastering key entities was proving to be too slow and not scalable enough to satisfy the business' requirements

Tamr's Human-Guided Machine Learning Automates Data Mastering & Classification

In searching for solutions to their mastering issues, the investment bank evaluated a variety of solutions to resolve speed and scale concerns. The team narrowed down the field of options to traditional, fully featured MDM platforms and Tamr's agile data mastering solution. Ultimately, the bank selected Tamr due to its capabilities around:

+ Automating the mastering and classification process across domains -- Tamr's approach to data mastering and classification leverages human-guided machine learning. This technique would enable an accelerated and scalable way to master and classify entities across any domain (organization, person, product, location) to an internal taxonomy

Ultimately, the bank selected Tamr due to its capabilities around:

Automating the mastering and classification process across domains

Agility and flexibility

Accuracy

+ Agility and flexibility -- In addition to Tamr providing a multidomain mastering and classification solution, the technology enabled the easy addition of datasets from any internal or external source through its schema mapping capabilities that use the same human-guided machine learning process

+ Accuracy -- Because Tamr interacts with non technical experts, primarily analysts from across the globe in this case, during its matching validation process, Tamr's accuracy was proven to be superior to that of traditional technologies

Open, Horizontal Architecture Results In Broad Applicability

Tamr's data mastering solution integrated nicely into the investment bank's existing data infrastructure and processes due to it's unique, open architecture and enterprise-grade capabilities. Specifically, the research group used Tamr to accelerate its mastering and classification capabilities across domains in the following way:

+ Tamr pulled dozens of relevant datasets from the open web into its system in order to stitch together a unified view of a company and the products it sells -- these datasets included information on everything from store opening and closing times, job listings, retail catalogs, pricing details, and any other information that could prove to be a leading indicator of performance

+ The system uses human-guided machine learning to master company and product names -- when it's not sure of a company match (e.g. Is 'Alphabet' the same company as 'Google'?), Tamr will ask an expert to validate the match to ensure accuracy and tune the machine learning algorithms for future use

+ Tamr then classifies the products to an internal taxonomy using the same process as mastering the entities -- for instance, the system will attempt to categorize "Jeggings" as "Jeans" but will ask an expert to validate the accuracy of this if it doesn't meet a certain confidence threshold

The insight generated from Tamr's mastering and classification operations powers analyst insight and redefines what's possible to capture as well as confidence in the results. Downstream users of the research group's dataset will now be able to easily conduct ad-hoc research on new target companies as well as product markets and identify leading performance indicators.

Tamr's ability to quickly master and classify entities across domains using machine learning coupled with expert validation makes the technology attractive and compelling for a virtually limitless number of use cases. This global financial services firm is already undertaking a variety of projects using the technology and will continue to roll it out across the enterprise as it identifies other areas where rapid, scalable data mastering and classification across domains is needed to fuel growth.

Tamr's data mastering solution integrated nicely into the investment bank's existing data infrastructure and processes due to it's unique, open architecture and enterprise-grade capabilities