

CASE STUDY

ShareChat's Machine Learning Team Grows Engagement, Inclusivity

Social media giant with over 400 million monthly active users partners with Arize to help surface and resolve model performance issues faster, freeing up the team to focus on core tasks

Introduction

ShareChat is a rapidly-growing social media unicorn headquartered in Bengaluru, India. In all, over 400 million monthly active users rely on ShareChat along with its popular video apps Moj and MX TakaTak to share and consume videos, audio, photos and text in over 15 languages.

For ShareChat, fairness – particularly for low-resource languages and communities with low connectivity – is core to why the company was founded. With the goal of creating an “inclusive community that encourages and empowers each individual to share their unique journey and valuable experiences with confidence,” ShareChat’s social mission extends to its pioneering use of machine learning in areas like [multimodal and multilingual abuse detection](#). This commitment is also attractive to a growing array of advertisers.

To help in the critical task of better monitoring and troubleshooting its models at scale across both structured and unstructured data, ShareChat’s monetization AI team selected Arize as its ML observability partner earlier this year.



Key Facts



Industry: Social Media

About: Social media giant with over 400 million monthly active users across ShareChat, Moj, and MX TakaTak

Models In Production: >200

Machine Learning Team Size: >100

Primary Use Cases: advertising optimization, click-through rate, content intelligence, computer vision (CV), natural language processing (NLP), recommender systems, more

Challenges

Delays in detecting and diagnosing model performance issues lurked as a potential problem given the high importance of machine learning (ML) to ShareChat’s advertising optimization and user engagement. Before implementing ML observability, the team faced:

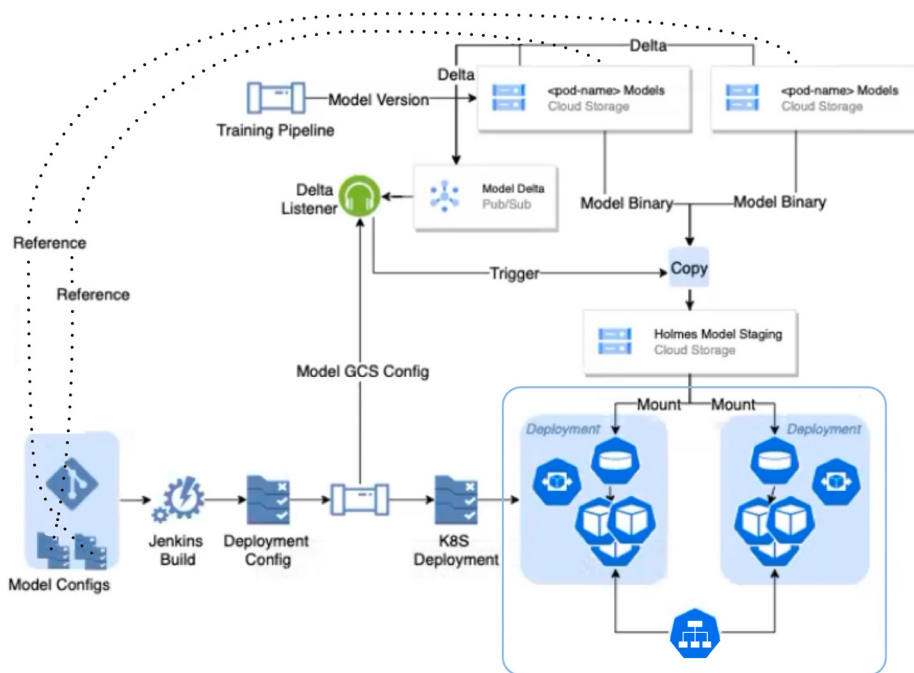
- ~ 24 hour delays to detect performance issues due to the limitations of existing internal dashboards and alerts
- A time-consuming ML troubleshooting workflow involving querying, calculating metrics, and writing ad-hoc scripts to slice data across hundreds of models
- Product and ad sales teams sometimes catching model performance issues before the ML team due to blindspots

“Because we are growing quickly both as a team and in terms of the number of models in production, it would be really tough without ML observability in place. Since implementing Arize, there have been multiple occasions where we were able to detect issues from a monitor and resolve them quickly. One recent example was a bug causing nulls to get created, impacting a key feature for a click through rate model’s performance. Since we had the monitor setup with Arize, it fired and we were able to fix the issue quickly. Because we get alerted proactively as soon as there are issues, we can act sooner and avoid a scenario where we see business losses or impacts to our users.”

Sourav Maitra, Technical Lead, AI - Ad Relevance, ShareChat

At ShareChat, machine learning models touch nearly every part of the business, including:

- Delivering consistent results on these use cases across 400 million monthly active users speaking and writing in over 15 languages is challenging. To help, the team at ShareChat has invested in a core ML stack to enable model training and serving and consistent service at scale.



"One of our main motives at ShareChat is to build a cohesive, inclusive community. For the ML team, that starts with content intelligence and creating a safe space free from things like hate speech or unwanted harassment. It's also about making sure we are connecting people across events and channels to lessen the distance between cultures and languages."

Sourav Maitra, Technical Lead, AI - Ad Relevance, ShareChat

- ### Solution

- Pre-launch model validation
- Automated monitors based on predefined thresholds
- Monitoring of both structured and unstructured data
- ML performance tracing to quickly pinpoint the source of model performance problems and map back to underlying data issues
- Concept and feature drift monitoring and troubleshooting to compare across training, validation, and production environments
- Data integrity checks to ensure the quality of model data inputs and outputs with automated checks for missing, unexpected, or extreme values
- Bias and fairness tracing to ensure models are not generating potentially biased or unfair outcomes for protected segments of interest

ShareChat's monetization AI team is seeing improved model performance and significant time savings since deploying Arize. That translates to:

- Hundreds of extra hours freed up per year across the team
- A payback period of under a year; >100% ROI
- Improved model performance from proactively surfacing feature drift and performance impact score at a cohort-level
- Robust drift monitoring for structured data, with the plans to implement embedding drift monitoring for NLP models
- Immediate visibility when issues arise based on predefined and automated thresholds, maximizing internal visibility and speeding up mean time-to-resolution
- Positive impacts on AI fairness goals

Challenges

ShareChat faces problems that are common at many companies growing an ML practice. Several pain points stand out.

Troubleshooting ML Is Distracting and Time-Consuming

“Even though we had multiple dashboards in place to catch issues, the process of troubleshooting was exhausting and not the best use of everyone’s time,” recalls Sourav Maitra. “Turnaround was often slowed by needing to go through a lot of tables and then do a bunch of queries. To really get to the bottom of an issue, you also need expertise – often, the person who created the model.”

Building More Robust Model Monitoring In-House Is Costly

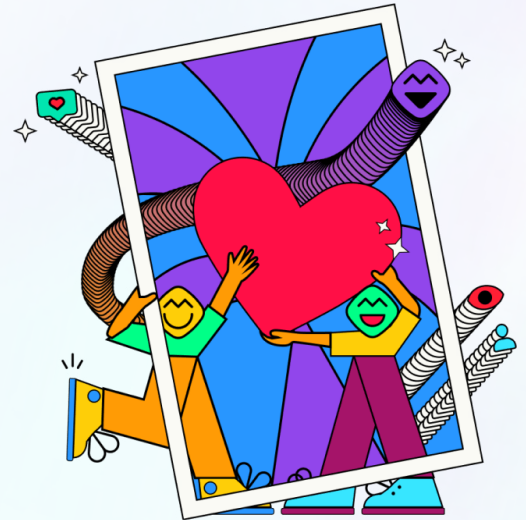
Despite having a sizable ML platform team well-equipped for such a task, every initiative comes with tradeoffs. “If we had created ML observability as part of our ML platform team that cut across the feeds, ads, and everything it would likely take at least three or four people along with a product person to understand the problem and do the orchestration,” Sourav Maitra notes. “No matter how good an internal solution we stand up, it’s likely Arize would outrun us because it’s totally focused on ML observability,” he adds.

Lack of Tooling for Monitoring Unstructured Data In Production

Despite a decade of industry investment in deep learning, to date there has not been a great way in the industry to monitor computer vision and NLP models as performance shifts in production. Performance degradation caused by new patterns that models did not encounter in training are difficult to catch with (often costly) human labeling teams. ways AI can solve business problems.

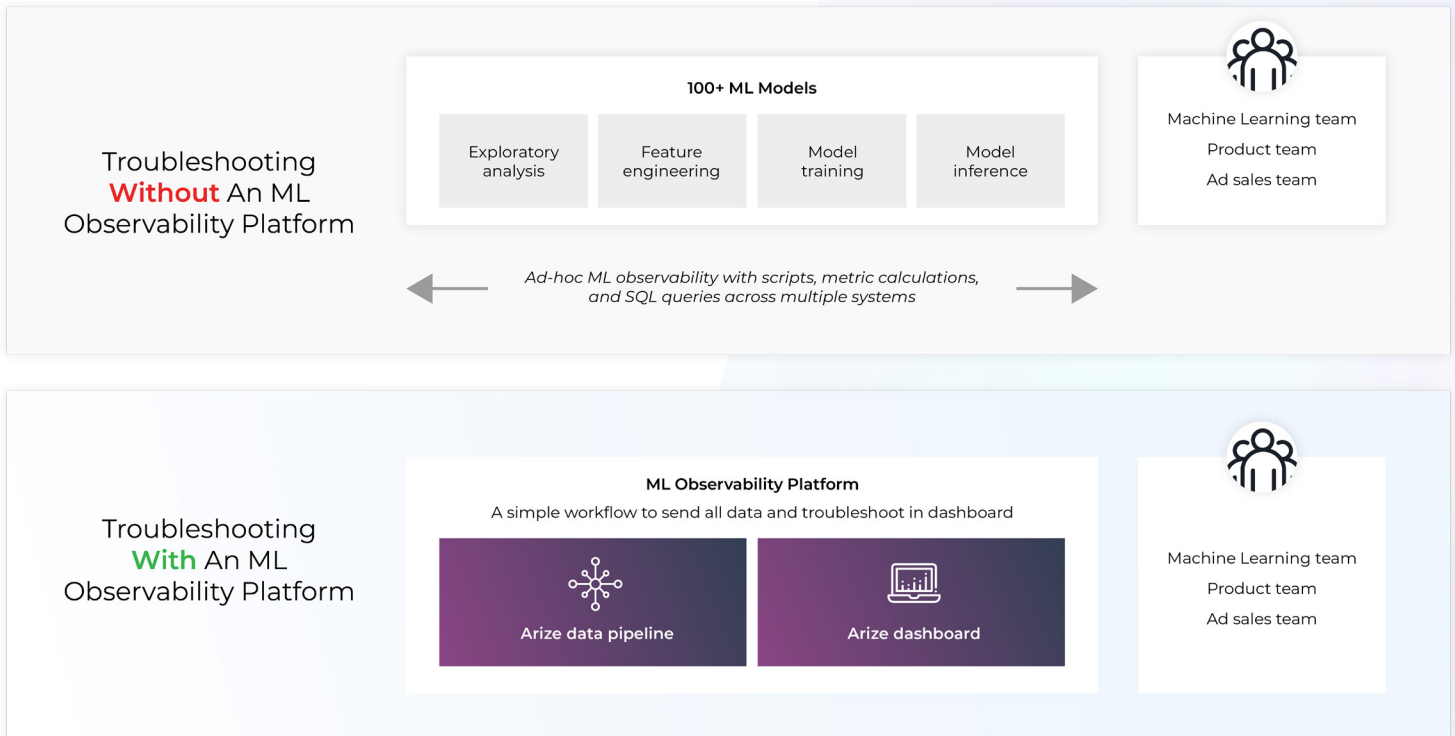
Analytics Complexity

A platform’s architecture matters, particularly when supporting analytics at enterprise scale; without it, ML teams might be delayed or lose insights at the worst possible time. With over 400 million monthly active users and billions of inferences a month, any potential ML monitoring solution would need to be able to expand query semantics to petabytes of data.



“It’s better to have ML observability in place, adopting it early before there is a must-have situation. There is always time needed to onboard and for teams to get used to any new system and way of approaching problems. Once your ML organization scales, it becomes tougher to realize that mindset shift.”

Sourav Maitra, Technical Lead, AI - Ad Relevance, ShareChat



Solution

Arize AI processes hundreds of billions of predictions monthly at scale on behalf of clients applying ML across a wide array of use cases. By connecting offline training and validation datasets to online production data in a central inference store, Arize's ML observability platform helps ML teams streamline model performance management, drift detection, data quality checks, and model validation. Arize also enables users to log models with both structured and unstructured data to the platform for monitoring.

ShareChat first selected Arize in early 2022 after a competitive review of model monitoring platforms. Key factors influencing the decision include Arize's robust capabilities to surface model and feature drift and performance issues at a cohort level and broad alignment on values and product roadmap.

"Arize enables us to focus on our core rather than building something from scratch," says Sourav Maitra. "If we're building a recommendation system, that can be our true focus. That's one reason why we thought third-party options would be a good fit. There were several other potential partners, but after doing our due diligence it was clear that Arize was the winner – open to feedback, with good alignment on what we need to provide long-term value."



Results

With robust ML observability from Arize now part of its ML stack, ShareChat's monetization AI team is able to focus on building new models and accelerating time-to-value for product teams. The organization sees a few key benefits from the Arize platform.

Model Performance Issues Are Now Caught Immediately Instead of Taking 24 Hours or More

Lagged dashboards are now supplanted by real-time alerts. When performance falls below a set baseline, a monitor fires and the team can quickly surface the problematic cohort where drift or performance impact score is highest. Compared to a traditional workflow, "Arize is much faster and you get horizontal and vertical cuts along with some unique insights," observes Sourav Maitra. Ultimately, that translates into significant time savings for the team and better business results.



Proactive Drift Detection and Troubleshooting

Models are not static. Especially in hyper-growth businesses like ShareChat where data is constantly evolving, accounting for drift is important to ensure models stay relevant. With Arize, ShareChat has robust tracking for prediction, data, and concept drift across model dimensions and values as well as the ability to compare across training, validation, and production environments. "Detecting drift was our p-zero use case with Arize and has been really valuable," continues Sourav Maitra.

A Design Partnership That Helps ShareChat Hit Its Goals

As an early stage startup, Arize is constantly building and iterating on new features and improvements in close partnership with its design partners. To ShareChat, that means gaining access to an agile team 100% dedicated to building an ML observability platform that fits its vision and needs.

"There are areas where we have made feature requests and the Arize team has been really accommodating in terms of altering internal timelines to help us accomplish our goals," says Sourav Maitra. "Generally, Arize is really open to feedback or even constructive criticism if there is a bug and has been proactive in terms of enabling easier onboarding."



Robust Model Validation

Arize also helps ShareChat validate models and achieve goals around AI fairness. "We're planning to create toy models and then push them to Arize to see if all the features we are using are proper or not – and, if not, then we can pull some of the features down," says Sourav Maitra. "That helps us in building, decisioning, and features."

Conclusion

Since its founding in 2014, ShareChat has made tremendous headway in its mission to help individuals form substantial connections and stay entertained. The company's innovative machine learning team, now aided by Arize for ML observability, is well positioned to help the company achieve that goal well into the future.

To start your ML observability journey, [sign up for a free account today](#) or [request a demo](#).

