

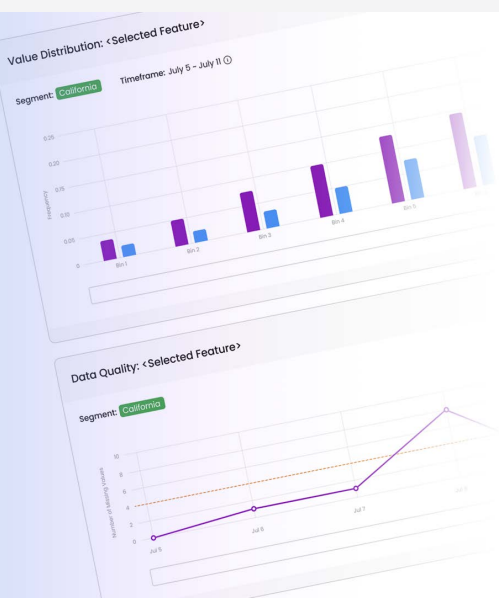
High-Performance ML Operations at Scale

VIANOPS uniquely brings together observability across layered, high-volume, complex dimensions, with root cause analysis to identify high-risk hotspots that jeopardize model behavior, and the flexibility to drive continuous ML operations across any cloud and any data source.

Today, organizations struggle to balance the need for simplicity in operationalizing models, flexibility in operationalizing different types of models, diverse use cases, various data sources and clouds, and the need for scalability while managing costs.

Alert fatigue, inability to drill down into massive amounts of data to pinpoint and address the cause of model drift, difficulty trying to automate retraining, and a lack of confidence that models are ready to be redeployed are all barriers to keeping models performing in production – VIANOPS simplifies these complexities for the user, enables the user to drill into the important details leaving aside the noise, and solve model performance problems immediately, across all data sources and any cloud provider.

With VIANOPS, we remove today's limitations and bring a comprehensive, high-scale, and flexible approach that puts the power into the hands of ML operations teams that need to operationalize machine learning models at scale, over a long period of time, to support reliable business outcomes.

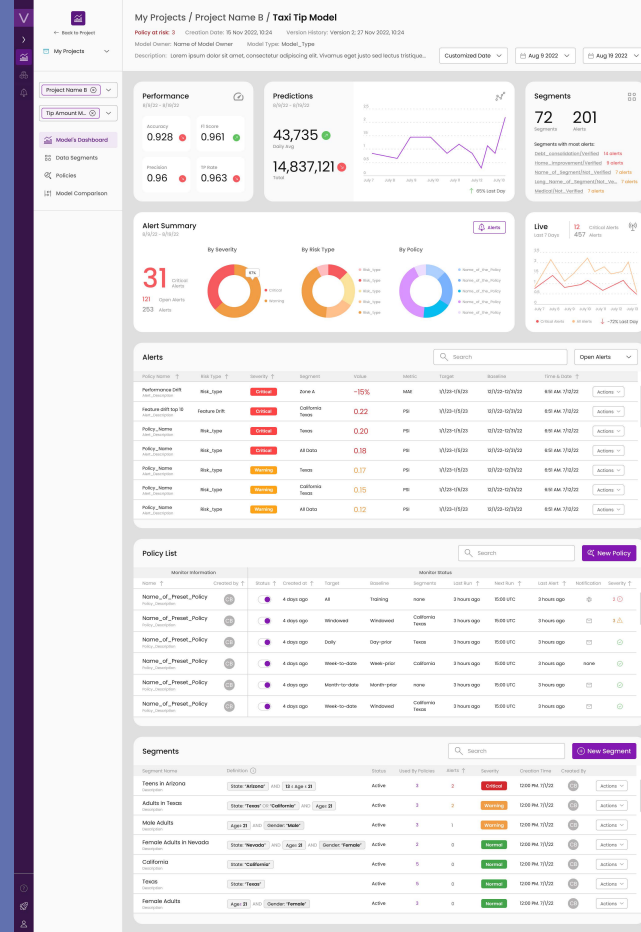


VIANOPS empowers data scientists, ML Engineers, and others responsible for managing AI and ML models in production with reliability and explainability, that starts from contextual monitoring, to root cause analysis of model behavior change, to recommended or auto-triggered best actions such as retraining, to model validation of new candidates of models before promoting to production.

4 Steps to Operationalize ML Models in Production – At Scale

- ▶ Know when models degrade
- ▶ Identify root cause
- ▶ Define what actions to take
- ▶ Validate the fix addresses the problem

As more models with more features and more inferences get into production, do this at scale - while **minimizing costs**.



Monitoring

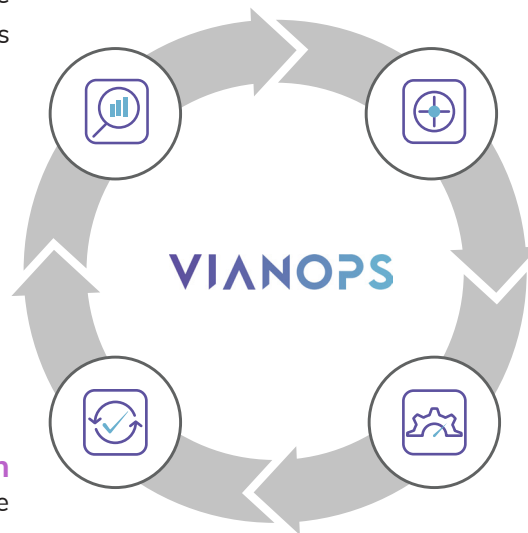
Actionable, unambiguous metrics to answer specific questions

Eliminate alert fatigue with multiple levels of warnings & critical alerts

Validation

Confidence in model performance before redeployment e.g., Champion/ Challenger

Back-testing and understanding risk levels before deployment



Root-Cause Analysis

Dashboards for analysis - across time & data segments

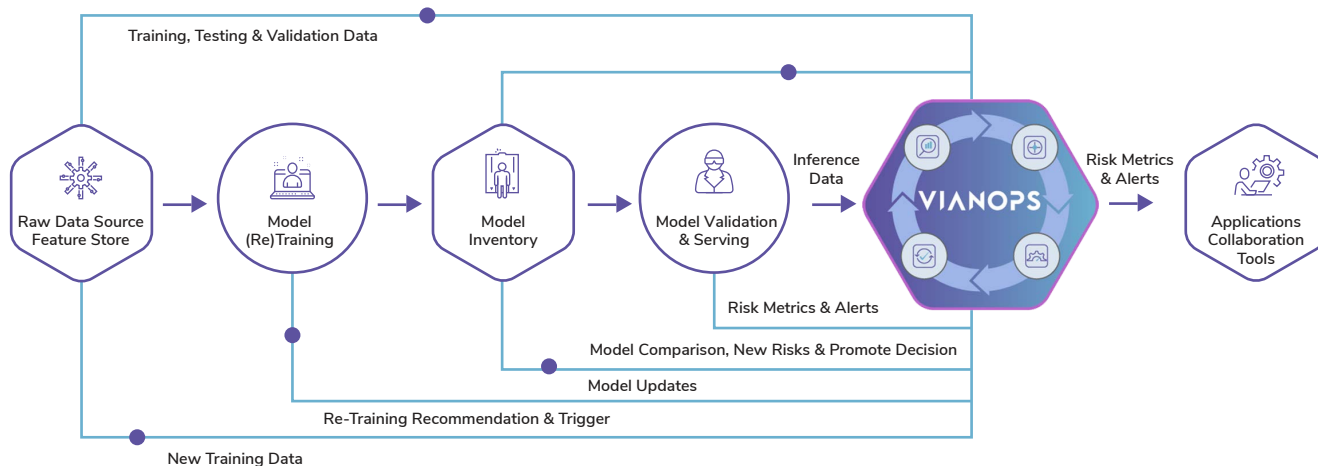
On-the-fly drill-down analytics across hundreds of features for real-time what-if analysis

Mitigation

Recommendations on actions e.g., data segments to retrain on


Automated and manual triggers

VIANOPS enables data scientists, ML engineers and ML operations teams to observe model behavior changes across multi-faceted dimensions in a simple user experience, gives users unparalleled ability to drill into root causes and identify the most under-performing or drifting subsegments, without impacting the cost or efficiency of these systems and operations. We do this through our unique approach to enabling high-performance ML operations at scale.



Extreme Scalability

 **Tens of thousands of inferences per second**

 **Hundreds of features per model on commodity hardware with no need to deploy expensive clusters**

Get started fast:

Call API to send production data.

Create policies to start monitoring model.

Operationalize your high-transaction, feature-rich model within minutes.

Enabled through:

- Kubernetes based architecture, agnostic to any cloud.
- API-first for easy integration with workflow or CI/CD pipelines.
- Data Connectors with out-of-box connections to hundreds of data sources including Snowflake and Databricks.
- MLflow as model repository and model version management.
- Channel integration with emails and Slack for alert notification.

VIANOPS Key Capabilities

Monitor Different Types of Model Risks

Enable data scientists to monitor all types of risks matters to model behavior: data quality, feature drift, prediction drift, concept drift, bias & fairness, outlier, and uncertainty.

Observe Risks from Different Dimensions

Enable data scientists to observe risks in the context of business from all different dimensions like day to training data, weekday to last 10 weekdays, week to week, month to month at desired sensitivity of drift and schedule.

Analyze Root Cause and Explain Model Behavior

Automatically spot under-performing or prediction drifting subsegments attributing biggest impact to the model behavior change, enable data scientists to explain model predictions and do what-if analysis to better understand model.

Improve Model Quality and Build Trust

Give data scientists action recommendations like fixing data quality issues of specific segments, adding more data to a minority class, retraining model with more recent data etc., to improve model quality more purposefully.

Validate Candidate Model before Promotion

Enable data scientists to do back-testing and compare new candidate model to production model to ensure better performance and no new risks are introduced.

Boost Efficiency with Affordable Cost

Allow data scientists to analyze massive amounts of data without worrying about the speed and cost of processing data. Allow data scientists to operate tens to thousands of models at scale and prioritize models with highest risk.

“The breakthrough for us is the sheer scale at which we can now monitor our models, with customized monitoring plans across a very large number of features. This includes tens of thousands of events per second that we can monitor for potential risks around accuracy, drift and other issues very quickly, and then retrain models if needed.”

Financial Services Customer

Choose Policy



Performance

Monitor model performance over time.



Feature

Monitor self distribution



Label Drift

Monitor model ground truth for a distribution drift.



Conc

Monitor to mod

Select Policy Type



Distanced

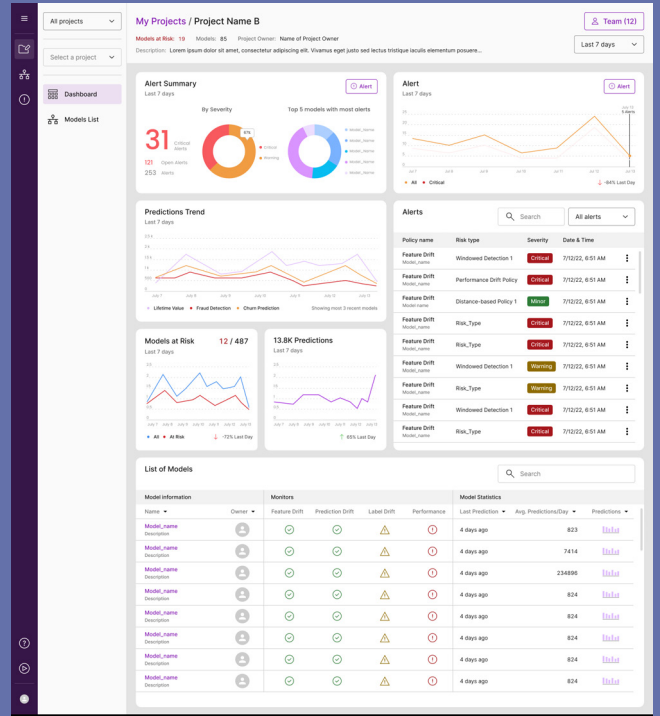
Detect and measure changes in your model's inputs

High-Performance ML Models at Scale

Monitor and observe across varying, complex and high-volume sets of dimensions: Observe models for risk metrics, moving windows, segments, multi-variance and more to help you best understand your model behavior.

Cut through the noise: Identify root causes and spot segments and sub-segments influencing model behavior across infinite possibilities at high speed.

Run on any cloud and integrate any data source: Agnostic to all clouds and seamlessly integrates with all types of data sources / feature stores, workflows, collaboration tools, as well as Notebooks and MLOps platforms.



KEY CAPABILITIES

- **Model Risk Metrics Store**
 - ▶ Performance Metrics
 - ▶ Feature Drift
 - ▶ Prediction Drift
 - ▶ Label & Concept Drift
 - ▶ Data Quality
 - ▶ Bias & Fairness
 - ▶ Uncertainty & Outlier
- **Inference Data Store and Analytics Engine**
- **Root Cause Analysis Engine**
- **Dashboard**
- **Alert Management**
- **Policy Creation and Management**
- **Model Repository and Version Management**
- **Channel Integration and Notification**
- **Sharing & Collaboration**

Software Stack



Cloud Agnostic



Let's Work Together!

✉ info@vian.ai