Tiger Analytics helped Fortune 500 P&C Insurer reduce 70% efforts in real-time model and data monitoring

Tiger Analytics built a comprehensive Model Management Framework that enabled real-time model and data monitoring. The solution went a long way to reduce manual model retraining efforts and minimize performance degradation. It ensured the smooth monitoring of variable and model performance and proactively retrained models, reducing the training time from 4-5 weeks to 2-3 weeks.

The Background

Our client is a Fortune 500 P&C Insurer with 200+ models deployed in production across LOBs and functions. The current system could not handle repetitive tasks around model monitoring and retraining. So, the client needed a referenceable single source of truth for model information and monitoring. They also required the ability to proactively identify issues such as model failure in the production environment, rather than an analyst to do so reactively. It would help significantly reduce the time taken to retrain models in production.

Key Challenges

No real-time monitoring: The lack of real-time monitoring made it impossible to track model execution and failure statistics daily.

Absence of early warning systems: There were no early warnings around variable distribution, leading to degradation in model performance.

Lengthy manual model retraining: Manual retraining tasks took more than 5-6 weeks of manual effort.

Our Solutions

Tiger Analytics created and implemented a framework for ongoing model monitoring in the client's production environment. This solution approach was divided into four core components, which were as follows:

The first component was Model Inventory. The team worked with the client to collect and document relevant data on models like base model info, model execution schedule, etc., by developing a web-based application to host the information.

In Data and Model Monitoring Framework, the second component, the team defined custom components to measure data drift, data quality and volume computation. They built an automated system to capture model execution statistics and designed a scalable solution, allowing quick new model onboarding.

The third component was Dashboards for BI Reporting, in which an interactive dashboard was developed to act as a single source of truth. It enabled the client to monitor model execution trends for the selected period and to drill down on model failure details.

The fourth and final component was Model Retrain Framework. The team developed pipelines that automated areas like Data Extraction, Feature Engineering, EDA, Model Retrain, and Result Validation. Similarly, a web application was also created to orchestrate retraining process automation.

Tech Stack

|	
AWS (SageMaker and CloudWatch)	Snowflake/ Oracle
AirFlow	Kibana

Solution Architecture

Model Management Framework

Orchestrator – Airflow
- Configure automate retries
- Track Individual Modules for failure
- Detailed log analysis
- Introduce containerization
- Integration with other Cloud services

Alerts and Notification – Service Now
- Breaches trigger ServiceNow Incidents
- L1 (Data Scientists) and L2 (MHM) Alert Levels
- Drift and DQ Dashboard in OpenSearch

Schedule Job run

Raise incident

Update/Assign/Close Incident
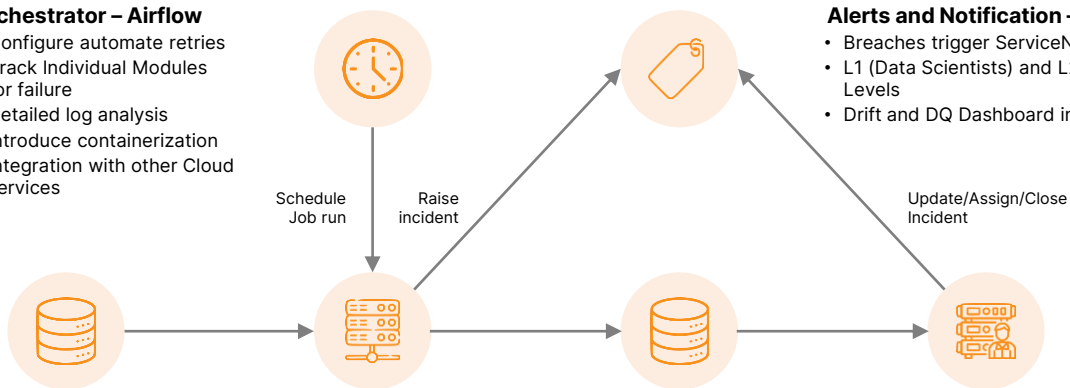
Model Data – Oracle/Snowflake
- Snapshot/Data source of each model.
- For cloud models, can easily fetch data from the respective source

Monitoring Framework – Cloud
- As workloads are dependent on the number of features and cardinality in addition to other factors, we can easily configure compute instances to scale accordingly using services like AWS SageMaker.
- Parallel processing can be introduced on module level.

Monitoring Data – Oracle
- Tables that have metric data (Drift, Data Quality and Model Performance) for each model at different frequencies.

Reporting – Excel and Dashboard
- Automatically process the monitoring data to create excel based reports with breach data and visuals which will
  - Reduce manual error and efforts in creating data and visuals for reporting.
  - Allow additional bandwidth to investigate breaches in depth.

Value Delivered

More than 70% reduction was achieved in data scientist efforts, allowing them to focus on strategic activities.

The capability to consistently monitor variable and model performance and proactively retrain models was enabled to avoid any negative business impact.

The model training time was reduced from 4-5 weeks to 2-3 weeks.

About Us

Tiger Analytics is a global leader in AI and Analytics, helping Fortune 500 companies solve their toughest challenges. With over 4000 data technologists and consultants spread across offices in the US, Canada, UK, India, Singapore, and Australia, we help our customers accelerate their AI and Analytics journey in sectors like CPG, Retail, Insurance, BFS, Manufacturing, Life Sciences, and Healthcare. Tiger Analytics is a

Great Place to Work Certified and a 'Leader' in the Forrester Wave: Customer Analytics Services Report 2023.

Visit https://tigeranalytics.com. to see how Tiger Analytics provides certainty for a better tomorrow.