



## CASE STUDY

# Anomaly Detection with IoT Data

## Executive summary

IoT Data sensors are not perfect and can send in bad data. Odometer data is used to create preventative maintenance plans, so when data is incorrect, it can cause trucks to not go in when they should, leading to money and time lost, or get maintenance before they should, leading to an increase in unnecessary spending. This solution helps showcase how ML can be used to bring awareness to these abnormal odometer readings, allowing our partner to make decisions quicker in order to get preventative maintenance plans back on track.

## The challenge

- **Semi-Automated Pipeline for Evaluating Odometer Data**  
Replace the existing manual process with a more automated way of being alerted to incorrect odometer data coming in from third-party sensors
- **Accessible Results**  
Make results of model easily accessible for both technical and non-technical users
- **Example Project for Future Analytics Projects**  
Implement best practices and documentation for this project to stand as a repeatable example for future projects

## Why choose AWS for your machine learning (ML) needs?

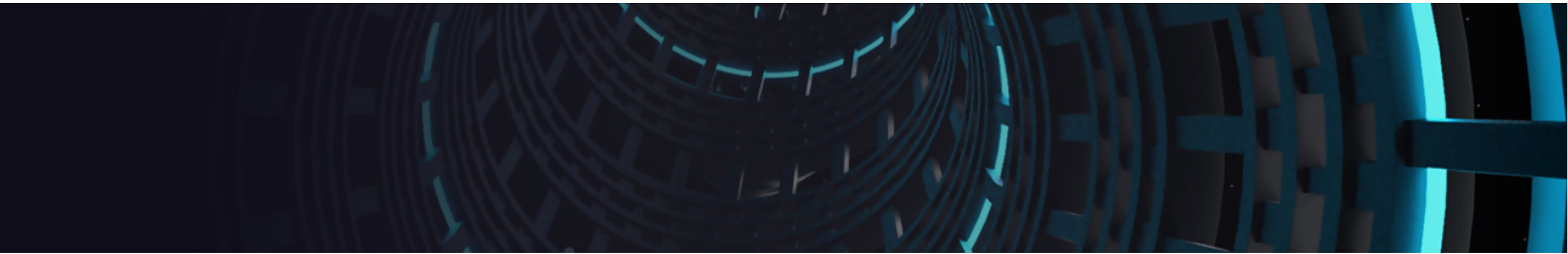
With high-performance compute options powered by machine learning, Amazon Web Services (AWS) enables organizations to undergo broad digital transformations with modern, cloud-native solutions. Offering a broad set of ML services and supporting cloud infrastructure, AWS enables organizations to tailor their ML solution to meet the unique needs of their business.



## About TMC Transportation

As the trucking industry's premier open deck carrier, we are known for integrity, innovation, highly trained drivers, prestigious late-model equipment, superior safety record, and unwavering dedication to quality and customer service.

*"Source Allies employs highly qualified people that have made a positive impact on my team and TMC in general. They know AWS extremely well." - Development Manager, TMC Transportation*



*"The team members from Source Allies have all excelled at getting TMC up to speed with numerous AWS services. They provided the initial guidance as well as ongoing support to help us meet our goals. I appreciated their ability to train us on all AWS services we wanted to utilize. Additionally, Source was able to pull in resources directly from AWS when needed to solve issues with AWS services." - Data Engineer, TMC Transportation*



## Why Source Allies?

Founded in 2002, Source Allies is a premier, team-based, technical consultancy in Des Moines, Iowa. We provide high-performing coaches and teams who work alongside our partners to cultivate a shared ownership environment. We are dedicated to our partner's long-term success and focus on continuous value delivery while remaining flexible to our partners' dynamic business needs. Our mission is to combine the skills of data science/data analytics/statistics with software development skills and best practices (TDD, data engineering, MLOps, cloud architecture) in order to incorporate the right skills into product teams so that we can answer tomorrow's questions without redesign.

## The solution

Third-party sensor data makes its way into the AWS environment via AWS Database Migration Service, landing in an S3 bucket. From here, data is directly ingested into Amazon SageMaker. Scripts are set up in this pipeline to take care of the cleaning and normalizing of this data before training occurs. An Amazon SageMaker endpoint is hosted to allow for easy use of model interaction and prediction. AWS S3 is used to land predictive results where they will be further evaluated and acted on. The model consists of an Isolation forest algorithm that goes through a hyperparameter optimization stage when training. F1, Precision, Recall, Specificity, and False Anomaly Rate are all used to evaluate the model performance.



## Results

Our model and application results in a reduction in reaction time in fixing bad odometer data within the source systems. A pipeline was established to train, test, and validate the model that can help determine bad odometer data. This pipeline will make future iterations 10x faster, reducing the amount of time needed for future development. Best practices have been set up as part of this project including monitoring, logging, and playbooks to act as a sample analytics project template for future projects.

## Benefits

The automation of this project has saved hundreds of hours from critical members' day-to-day jobs, freeing them up to complete other tasks. In addition, the reduction of unnecessary maintenance has saved hundreds of thousands of dollars in maintenance costs. Drivers spend less time on the side of the road waiting to deliver their goods resulting in happier drivers, happier customers, an increase in on-time deliveries, and money saved across the board.

*"Source Allies is very knowledgeable at cloud technologies, infrastructure and Project Management. They are able to level up the entire team on the AWS stack and even helped me view working differently."  
Data Engineer, TMC Transportation*

## Next steps

This process is still manual in some places. The next steps are to create a way for our partner's SMEs to review the outcomes of the model, label them as correct or incorrect and then feed this back into the algorithm for future learning.

## Benefits

### ➤ Fewer Roadside Assistance Calls

Any time a truck ends up on the side of the road is a critical moment for the driver. Any delays drivers face, especially time costly delays like breakdowns can result in unhappy customers and unhappy drivers. This solution helps correct odometer data quicker, allowing for

### ➤ Reduced Unnecessary Maintenance

Any time a truck goes in for maintenance before it needs to is time and money wasted. Each truck is on a predictive maintenance schedule and by correcting odometer fluctuations, we can save the time and money of trucks going in before they were scheduled to.

### ➤ A Machine-Learning Solution Using Best Practices

This project sets up everything needed to deploy a machine learning solution into AWS using best practices like a documented playbook, monitoring on the infrastructure and model, alerting enabled, and a continuous delivery pipeline established.

## About Source Allies

We partner with privately owned small businesses to fortune 100 organizations. We have received the Migration & Modernization Competency and applied for the Machine Learning Competency in 2022. Our teammates are certified in: Data Analyst Specialist, Professional Solutions Architect, Professional DevOps, SysOps Administrator, Machine Learning Specialty.



• Migration Services Competency