Glassbeam White Paper

Machine Learning Capabilities

Contents

Glassbeam Enhances Predictive and Prescriptive Features in IoT Analytics Platform with Machine Learning Capabilities

Virtually every product – whether a storage device, server, medical imaging system, industrial robot, automobile or ATM – generates and communicates information about itself (known as "machine data") that can disclose vital insights into system performance, health, usage patterns and more. These connected products are an integral part of the Internet of Things (IOT) phenomenon that is fundamentally altering how manufacturers build, market and support their offerings.

Gleaning insights from machine data becomes more difficult as the volume, variety and velocity of this data continue to mushroom exponentially. The challenge to extracting insight from machine data is twofold: the amount of machine data can be so large that it defies management; and the analytic tools for generating meaningful insights from unstructured machine data are inadequate.

Glassbeam's unique machine-data analytic capabilities address both these challenges. By leveraging machine learning capabilities, Glassbeam is adding a powerful new dimension to its IOT Analytics platform that will enable product manufacturers and their customers to become proactive about heading off product failures instead of simply reacting to them.

By gleaning insights from historical data, manufacturers can predict future behavior and preempt situations that cause product failure.

Applying Machine Learning To Data Analytics

Any attempt to predict how a product will perform (or fail to perform) in the future requires three things: 1) data about past performance; 2) lessons and rules learned from analyzing that data; and 3) a model of how the product is likely to perform in the future. At its most basic level, machine learning automates this process, giving computers the ability to learn from data, instead of being explicitly programmed. An example is an email spam detector, which is trained with spam and non-spam emails. Once trained, it can automatically classify new emails as spam or not. The logic to detect spam is not explicitly programmed.

Machine learning relies on "training" a device to apply a model of actions to be taken when a set of conditions is observed. Obviously, the more data gathered about past performance, the more accurate the prediction model will be. So while it is possible to perform these three steps just described, it is inevitable that manual efforts will be very limited in scope compared to what can be accomplished using automated methods. Modern tools allow for the collection, ingestion and analysis of very huge amounts of data; thereby enabling the creation of more accurate models that fully utilize the powerful capabilities inherent in predictive analysis and machine learning.

From Predictive To Prescriptive Analytics

Machine learning is closely tied to predictive analytics. Predictive analytics draws upon machine learning, data mining and modeling techniques in an effort to predict future events based on models informed by past experience. Well-known implementations of predictive analytics include credit scoring, fraud detection, and insurance or securities underwriting.

The next frontier in the Internet of Things will involve real-time, predictive analytics incorporating sophisticated machine-learning algorithms. This will move predictive analytics (forecasting what might happen) to prescriptive analytics (suggesting what should be done to produce a better outcome).

Integrating Apache Spark with Glassbeam's SCALAR Platform

This breakthrough in predictive analytics is accomplished by integrating Apache Spark with Glassbeam's SCALAR data transformation platform, which takes any type of data – be it structured, unstructured or multi-structured - from any type of machine and prepares it for analysis. The combination of Glassbeam's distributed machine-data processing architecture and the Spark distributed in-memory computer architecture, and its MLLib library, yields a fast, scalable analytics solution for processing large-scale IoT machine data.

For example, using historical machine data, Glassbeam can build a model that analyzes incoming data to determine device health. This prediction engine is integrated with Glassbeam's rules and alert engine, allowing customers to create rules that will automatically notify them when the prediction engine anticipates a device might fail in the near future. Similarly, a model can be built to predict when a system will run out of capacity and cause performance problems or a service interruption. Advance notification of a likely interruption or failure will enable customers to take proactive steps to prevent the problem from happening.

Core Benefits of Machine Learning

Applying machine learning to predictive analytics has two benefits. First, it automates the process of identifying potential problems, reducing the need for manually specifying thresholds. Second, it expands the number of factors that can be considered in assessing product performance or health. This makes predictions of future problems more accurate by drawing upon the largest possible pool of factual information.

When a product fails or its performance degrades, there's almost always a trail of evidence leading to the problem. But that evidence usually is obvious only in hindsight.

The possible causes of a product failure or poor performance can be numerous, and can involve combinations of factors that might never be discoverable by manual means. So, the ability to analyze large amounts of machine data – and then perform regression analysis to assess the relationships between various factors – can be extremely useful in understanding the cause(s) of product failure or poor performance. When this knowledge is gathered in the context of machine learning, rules can be created to alert users to situations that may lead to failures, allowing them to take actions that will prevent such failures.

Machine Learning In the Real World

Let's look at how things change with large-scale machine-data analytics, and then how they change again with the assistance of machine learning.

Case 1: High-Performance Networking

The product involved is a local-area networking device used to direct traffic between hundreds or even thousands of clients. Degraded performance or failure of a device can have negative impacts that are instantly evident to users.

In the past, network managers usually did not know there was a problem with a device until they started receiving complaints about network performance. Then, they would have to track down the source of the problem and pursue their best guesses as to what the cause(s) might be. Depending on the complexity of the problem, it could take hours to identify and resolve the issue.

Once the manufacturer of the network devices started using the Glassbeam service, things changed dramatically. The LAN controllers self-report information about themselves on a regular basis, and the Web-based Glassbeam application gives network managers instant and easy views of the current status and past performance of all devices. When a device fails, or suffers significant performance degradation, it is instantly evident using Glassbeam tools. Mean-time-to-resolution of a problem is reduced dramatically, and in many cases no change in network responsiveness is evident to end-users.

With the ability of Glassbeam to gather, analyze and report on large amounts of machine data, network managers identify factors that have led to product problems in the past, and apply lessons learned. For example, they might decide that any time a controller reaches 80% of bandwidth capacity for more than one minute, an alert should be triggered.

With the addition of machine learning capabilities to the Glassbeam service, even more dramatic changes occur. Through the application of regression analysis, many factors can be considered as possible contributors to a problem in product performance. This is a level of fact-finding and analysis that goes far beyond what humans could do, both in the amount of data to be studied and the number of variables considered. By applying machine learning, much more sophisticated and sensitive rules can be developed to guide network managers. For example, the “80% for one minute rule” could be augmented or replaced with new rules that would trigger an alert – e.g., 75% of capacity for controllers using version 12.0 or higher of the software, and running on specified model numbers.

Case 2: Efficient, Reliable Mass Storage

The products involved are large storage devices used by customers who require reliable, low-latency access to huge stores of information. But while reliable performance is important, so is efficient use of storage capacity, given the capital and operating costs of the devices.

For years, the devices had the ability to gather information about themselves and how they were performing, but there never was a way for this potentially valuable "machine data" to be uploaded and analyzed. Moreover, if a problem with a drive did occur, storage administrators had to engage in time-consuming search efforts to isolate the cause(s) and implement a remedy. These frustrations were resolved when the manufacturer engaged the Glassbeam service, which produces a Web-based, graphical, easily understood dashboard of performance metrics. IT staffs can now see in an instant how all devices are performing, where problems may currently exist, and what factors may have contributed to past problems.

With the addition of machine learning capabilities to the Glassbeam service, the power of machine-data analytics is put to full use, to the benefit of both the manufacturer and its customers.

With machine learning, techniques such as decision-tree learning, support vector machines, reinforcement learning and regression analysis can produce sophisticated and highly nuanced analysis of the various factors that can affect product performance and possibly lead to failures – far more nuanced, and based on far more hard data, than could ever be achieved by human efforts. These factors can then be turned into rules that the machines learn to follow. The result is that IT administrators are alerted when thresholds are reached (e.g., capacity utilization, error messages, spindle failures) and they can act to prevent failures or performance degradations.

Further, customer information from a CRM system can be combined with the machine data like say capacity usage, performance to predict the probability of a customer making an upgrade or a purchase. As an example – Customer X in the past decided to upgrade when the capacity was C1, performance P1 and number of users using the device was N1. Based on the data currently being sent back by the device, the Glassbeam platform can predict that a similar situation will occur in 3 months from now – and this could be used as a the basis for estimating the likelihood of an upsell opportunity.

Summary

When machine learning, predictive analytics and machine-data analysis are combined and delivered in easy-to-use tools, the benefits to both product manufacturers and their customers are clear: proactive identification of potential problems before they cause interruption or performance degradation, along with deeper insight into how products can be best used.

By tightly integrating Apache Spark machine-learning capabilities with its IOT Analytics platform, Glassbeam now offers real-time and predictive analysis of machine data.

Contact us at sales@glassbeam.com

Glassbeam, Inc.
5201 Great America Parkway, Suite 360 • Santa Clara, CA 95054
Phone: 408-740-4600 • www.glassbeam.com