Identifying Product Quality Drivers with Machine Learning

TekLink

Case Study

VENATOR

Business Objective

▪ The business objective was to explore the ML and AI opportunities to gather, collate and analyze IoT data from Venator's manufacturing sites to identify the key conditions required to produce the highest quality pigment and highlight any outliers that could proactively enable engineers to take corrective actions.

Client

▪ **Venator** is a global manufacturer and marketer of chemical products that comprise a broad range of pigments and additives that have applications across a wide range of industries, spanning agriculture to cosmetics.

Industry

▪ Manufacturing

Function

▪ Manufacturing / Production

Technology

▪ Microsoft Azure, Data bricks, Data Factory, Power BI

The Solution

By closely collaborating with their business and IT Teams, we provided end-to-end services to help our client establish a framework for a structured data collection and analysis process, based upon the following components:

▪ An MS Azure-based architecture using Data bricks and Azure Data Factory.
▪ Automated data pipelines to collect, integrate, enhance, and store IoT data for future analysis.
▪ Self-service, ad-hoc analysis capabilities of available data using PowerBI.

In an iterative process, the team developed machine learning models based upon the framework:

▪ Exploratory analysis of IoT sensor data to identify statistical relationships and anomalies and kick-off in-depth conversations with the engineering team to help refine the final dataset for analysis.
▪ Feature engineering: Extraction of features from raw data utilizing the deep domain knowledge of process experts to optimize the performance of the machine learning models.
▪ Applying tree-based machine learning algorithms (XGBoost and Random Forest) in close collaboration with process experts to iteratively fit the dataset and train models to predict product quality.
▪ Visualizations of trends and sensor influences from these models prompt further conversations around how to optimize the production process for better product quality.

Outcomes and Benefits

▪ Sensor data from the production plant is continuously captured and consolidated into the Cloud, providing key decision-makers with easy and immediate access to cleansed data in real-time.
▪ Machine learning models confirmed and substantiated the business's understanding of the key drivers influencing dispersion quality and have given pointers for future investigations into key drivers of product quality and into improving the approach to collecting IoT data in order to gain more mileage out of the sensor data.
▪ Efficient data collection enables a deeper understanding of the process and helps retrain and update machine learning models, adjust necessary parameters, and output more accurate analysis results that will help reduce production cost and increase revenue by providing an increased yield of high-quality products for our client.