

Bank Increases Trust, Reduces Risk and Solves the Laborious DQ Problems Automatically with AI/ML

Customer: Financial Services

A Global 1000 Financial Services organization moves credit risk information from 26 different systems to AWS. Data goes through a number of normalization and transformation before being pushed AWS S3 based data lake. A reporting solution is then used to provide various regulatory reports to various groups.



Challenges

Customer wanted to track over 2200 segments of credit risk data based on the country, product type and risk code combination.

The data had to be validated as soon as it was ingested into S3.

They wanted to flag the following types of data errors stemming from either the source system changes or transformation errors.

- Invalid Country-Product-Risk score combination. If a new combination is created (because of source system changes or transformation issues), downstream stake holders need to be altered. In addition, customer wanted to flag any new country code, product code or risk code.
- Abnormal changes in the number of transactions in each of the 2200+ Country-Product-Risk Code based segments. Anomalous credit risk transactions within each of the 2200+ credit risk segments.
- Anomalous changes in the sum of the credit risk in each segment .
- Technical Data Validation Checks (Null value percentage for each segment), duplicate records etc.

Outcomes



DataBuck's ML algorithms autonomously identified over 2200 credit risk subsegments from the banks transaction dataset and created data fingerprints to track each risk segment individual.



DataBuck autonomously constructed detailed data quality fingerprints for monitoring and validating transactions. This was equivalent to over 22,000 traditional Data Quality rules. These DQ rules were self-learning and evolved with the evolution in data over time.



Within 2 weeks of deployment, DataBuck identified 31 errors that the SME's had not anticipated ("unknown-unknowns") and so the existing data quality solution did not catch. The hidden risk of bad and inconsistent data propagating downstream was minimized.

DataBuck – Benefits



People Productivity
Boost >80%



Reduction in
Unexpected
Errors: 70%



Cost
Reduction
>50%



Time Reduction to
Onboard data
set ~90%



Increase in
processing
speed >10x



Cloud
Native