PROJECT SERIES OVERVIEW

🧩 How I Built a Complete Credit Rating and Monitoring Solution for a Utility Company (Using ML and Power BI)

This article is a walkthrough of my 3-part project series on Credit Rating Modelling and Monitoring—detailing how I tackled the challenge of early credit risk detection and intervention in utility operations by first generating a simulated dataset of customer behavior, then training a predictive model, and finally building an interactive dashboard for business teams.

By Jase Tran.

---

1. Introduction

Motivation: Predicting Risk Before It Becomes Loss

In utility operations, financial losses don't appear overnight—they accumulate gradually. A missed payment becomes a pattern, which leads to penalties, then service suspensions, and eventually, write-offs.

While working on a credit risk analytics project at **Manitoba Hydro**, I experienced firsthand how impactful early detection can be. The ability to anticipate risk—rather than merely react to it—can mark the difference between financial recovery and revenue loss.

That insight led to this project series: a full-cycle solution designed to simulate, score, and surface customer credit risk in a way that's realistic, explainable, and operationally useful.

What This Article Covers

This recap brings all three parts of the project together. We'll walk through each phase of the pipeline, demonstrating how they connect and build on one another:

- **Part 1: Data Engineering**
  Modeled realistic utility billing behavior across 14,000 accounts over 6 years—capturing usage, payments, penalties, and lifecycle events.
- **Part 2: Risk Classification**
  Engineered behavioral features from the billing data and trained a machine learning model to predict delinquency risk before escalation.
- **Part 3: Power BI Report**
  Designed an interactive report that helps business teams track risk, explore account history, and prioritize interventions.

Whether you're a data analyst, operations leader, or aspiring modeler, this walkthrough offers a clear view into how behavioral data, predictive modeling, and storytelling can work together to address real-world challenges.

---

2. Part 1: Engineering Utility Billing Behavior

Understanding the Data Requirements

The goal of this project was to identify risky accounts before they escalate into financial loss. To do that, we needed a detailed understanding of each account's billing and payment behavior over time—how charges accumulate, how customers respond, and when financial deterioration begins. This behavioral history is critical for quantifying the impact of key risk events like missed payments, service suspensions, or account closures.

Ideally, the dataset should include:

- **Month-by-month account activity** for a large and diverse customer base
- **Complete billing amounts,** payment actions, and outstanding balances
- **Clear escalation signals**—penalties, service suspensions, closures
- **Final outcomes labeled** in a way that's traceable to the behavior that caused them

However, in practice, utility billing data often falls short. It can be noisy, incomplete, or inconsistently labeled—making it difficult to analyze or use for predictive modeling. To solve this, we took a different approach – by simulating our own data.

Why Simulate?

Instead of relying on historical data with known limitations, we created the dataset from scratch—simulating utility account behavior using a rule-based engine. This gave us full control over data structure, behavioral logic, and label quality.

Simulating the data offered several major advantages:

- **Full Behavioral Coverage** – We could observe accounts from activation to closure, capturing the full risk lifecycle.
- **Clean, Traceable Labels** – Escalations like penalties or closures were always tied to specific, explainable behaviors.
- **Customizable Scenarios** – We could model a wide range of risk profiles by adjusting reliability scores, usage patterns, and penalty logic.
- **Experiment-Ready Format** – The output was structured and consistent, making it ideal for downstream machine learning and dashboarding.

This approach ensured we had a dataset purpose-built for understanding—and predicting—credit risk.

How Our Simulation Works

Our simulation engine models each **customer as an autonomous agent** interacting monthly with the billing system. During each billing cycle, the customer:

- Consumes electricity (based on a usage profile)
- Receives a bill based on usage and outstanding balances

- Makes a payment decision governed by a **latent reliability trait**, which influences:
  - **Payment timing** (early, on time, or late)
  - **Payment amount** (full, partial, or none)
  - **Likelihood of skipping payment entirely**

The billing system evaluates accounts at the start of each new billing cycle using the Balance Snapshots table. If payments fall short, it imposes penalty points based on the severity of non-payment. These drive automatic status escalation:

- **5.0 points** → Suspended
- **10.0 points** → Closed

The Balance Snapshots Table

This is the centerpiece of the dataset and risk logic. Each monthly snapshot captures the account's financial state, including:

- **Delinquent Amount / Ratio** – How much of the prior bill remains unpaid, in absolute and percentage terms
- **Delinquency Type / Penalty** – Severity tier (Minor, Partial, Major, Full), assigned a penalty score from **0.5 to 2.0**
- **Delinquency Score** – Cumulative penalty points since last full recovery — a running indicator of financial risk
- **Delinquency Status** – Operational state based on score: **None**, **Delinquent**, **Suspended**, or **Closed**

Snapshot ID	**Billing ID**	**Account ID**	**Check Date**	**Total Due**	**Date Due**	**Delinquent Amount**	**Delinquent Amount Ratio**	**Previous Unpaid Balance**	**Previous Delinquency Score**	**Is Delinquent**	**Delinquency Type**	**Delinquency Penalty**	**Delinquency Score**	**Delinquency Status**	**Account Action**
1	…	…	2020-03-01	141.42	2020-03-01	141.42	1.0	0.0	0.0	True	Full	2.0	2.0	Delinquent	Account Marked Delinquent
2	…	…	2020-03-01	68.78	2020-03-01	0.00	0.0	0.0	0.0	False	None	0.0	0.0	None	No Action
3	…	…	2020-03-01	71.66	2020-03-01	71.66	1.0	0.0	0.0	True	Full	2.0	2.0	Delinquent	Account Marked Delinquent

Simulated Dataset Overview

For this project series, we simulated a starting population of **3,000 customers** over **72 billing cycles** (6 years). Over time, these customers opened multiple accounts, consumed energy, received bills, made payments, and experienced consequences for delayed or missed payments.

The simulation generated:

- **14,000+ accounts**
- **580,000+ records each** for usage, billing, and payments
- **A complete history** of balance snapshots across the lifecycle of each account
- **500+ final closures**, representing accounts written off due to unresolved delinquency

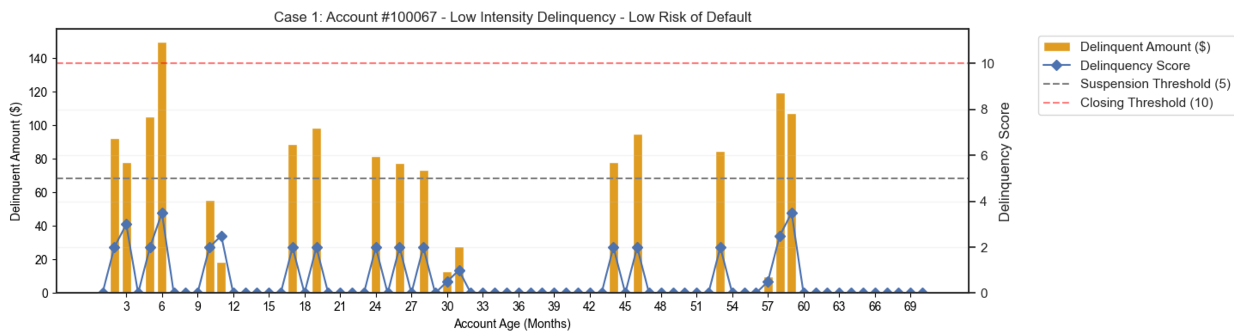
With this volume and richness, the dataset provides a robust base for both modeling and evaluation.

Sample Accounts: How Risk Patterns Manifest

To illustrate how this simulation produces distinct behavioral patterns, here are two example account history profiles, tracked by their monthly **delinquent amount** (bar) and **delinquency score** (line).

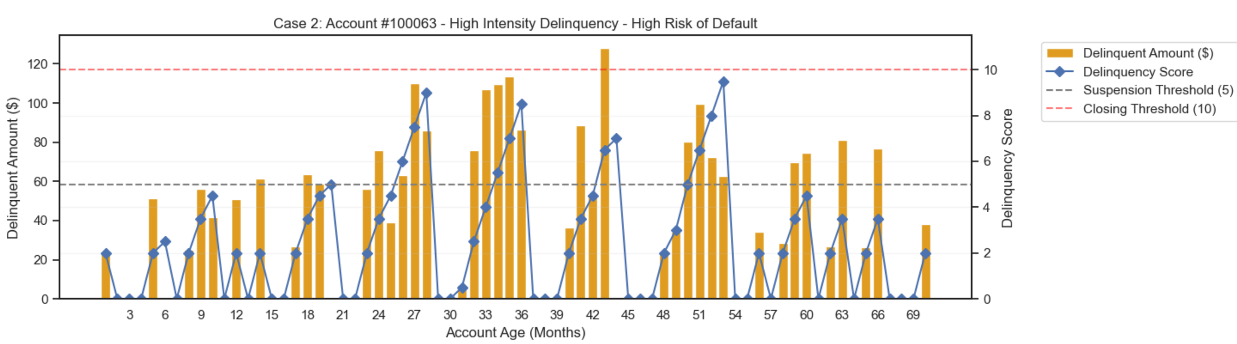
An account becomes **suspended** once the delinquency score crosses the 5-point threshold (grey line) and is **permanently closed** and marked as bad debt once it crosses the 10-point threshold (red line).

Case 1 – Low-Intensity Delinquency (Low Risk):



- **Observation**: This account shows occasional missed or late payments early on but demonstrates recovery and consistent stability thereafter. Delinquency scores remain well below escalation thresholds. As a utility provider, this customer's behaviour is acceptable in the grand scheme.

Case 2 – High-Intensity Delinquency (High Risk):



- **Observation**: This account exhibits repeated and prolonged payment issues. Delinquency points accumulate steadily, leading to multiple service suspensions and near-closures due to sustained non-payment. This customer presents a high risk of defaulting in the future and requires special attention from the utility company to prevent them from becoming permanently closed and written off.

These examples reflect the realistic progression of customer risk—and show how trackable patterns form the basis for model classification.

With this comprehensive dataset in place, we were ready to move on to the next phase: designing a classification model that predicts risk based on these behavioral signals.

---

3. Part 2: Creating A Robust Credit Rating System With ML

Goal & Challenge

To detect risky accounts early enough for intervention, we need to predict their future outcome—and to do that, we need to learn from their past behavior. The core challenges were twofold

The first challenge is data structure: we need to capture and aggregate each account's full behavioral timeline. While the original dataset provides rich monthly snapshots, these need to be transformed into features that summarize trends, frequency of delinquency, recent windows' performances, and more.

The second challenge is labeling: how do we decide which accounts are "risky"? Simply labeling those that were eventually closed might ignore a broader set of accounts that are still active but trending poorly—or are bad debts in the making. Defining risk in a way that balances accuracy with coverage was key.

Feature Engineering

To translate each account's timeline into structured, meaningful features, we engineered behavioral indicators that reflect how risk builds up over time.

For every modeling snapshot, we looked back at the account's history and calculated trends such as:

- **Rate of Offense:** How often did the account miss, delay, or get suspended?
- **Streak of Offenses:** Were there clusters of repeated misses or prolonged delinquency?
- **Peak of Offense:** What was the highest delinquency score the account reached?
- **Recent Trends:** How had the account been performing over the last 3, 6, and 12 months across key signals?

To ensure fairness across accounts at different stages of their lifecycle, we normalized many of these features based on account age.

We also intentionally **excluded monetary values** from the training process to ensure the model focused on behavior—not just bill size or usage level. This helped create a model that generalizes across customer tiers and plan types.

Final Engineered Dataset

Snapshot Key	Account Age	Max Delinquency Score	Avg. Penalty Per Incident	Deliqnuency Rate	Suspension Rate	Active Delinquency Rate	Avg. Penalty	Delinquency Streak Length 1M Rate	Delinquency Streak Length 2-3M Rate	Delinquency Streak Length 4M+ Rate	Payment Full Miss Rate	Payment Major Miss Rate	Payment Partial Miss Rate	Payment Minor Miss Rate
216621	16	4.5	1.722	0.562	0.000	0.562	0.968	0.250	0.125	0.000	0.375	0.125	0.000	0.062
281483	26	3.5	1.650	0.384	0.000	0.384	0.634	0.230	0.076	0.000	0.269	0.038	0.000	0.076
50048	1	0.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
313125	52	7.5	1.592	0.519	0.038	0.480	0.826	0.096	0.134	0.019	0.269	0.115	0.096	0.038
56279	9	3.0	1.750	0.444	0.000	0.444	0.777	0.222	0.111	0.000	0.333	0.000	0.111	0.000

Labeling Policy

With features in place, the next step was defining what we mean by *"risky."*

A simple binary classification—labeling only closed accounts as high-risk—misses important nuance. Some accounts show prolonged signs of trouble without formally closing, while others may stabilize after early issues.

To reflect this, we defined a three-tier labeling system:

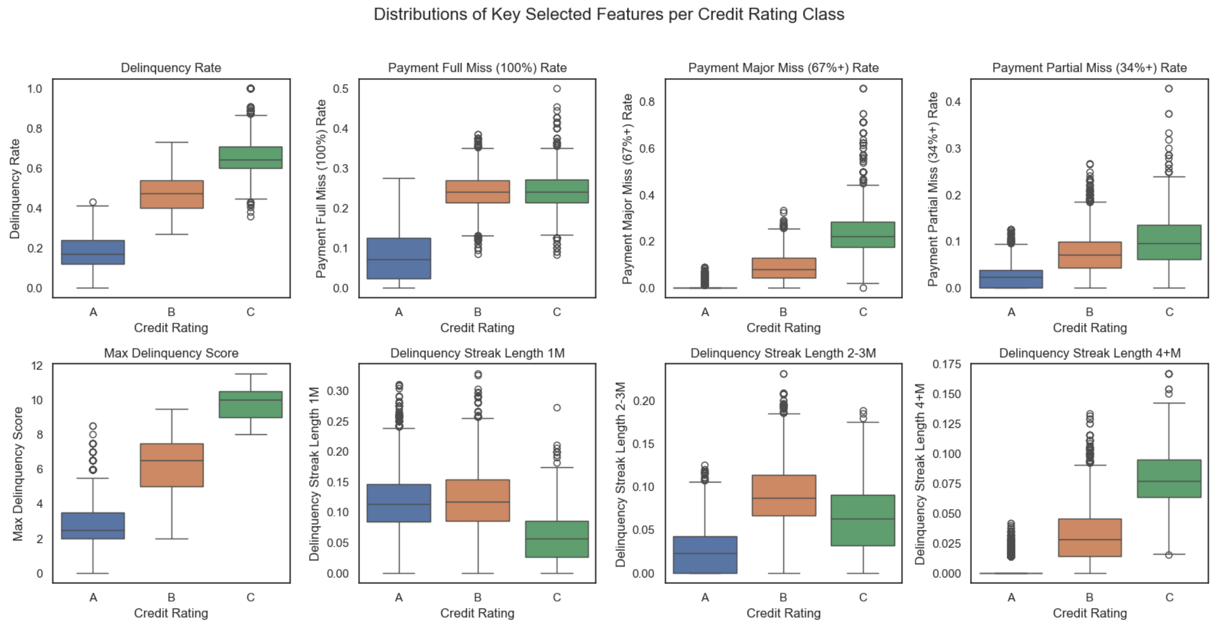
- **Class C (High Risk):** Accounts that were eventually closed due to delinquency, or those with persistent risk patterns—such as extended suspension or repeated penalty spikes.
- **Class B (Medium Risk):** Mature accounts (24+ months) that show recurring issues like partial payments or elevated delinquency, but have not escalated to closure.
- **Class A (Low Risk):** Mature accounts with minimal delinquency and consistent payment history over time.

We generated these labels using a policy that combines rule-based logic (for class C) with behavior-driven clustering (to separate classes A & B)—to capture both clear-cut cases and nuanced in-between states.

Class Feature Analysis

The resulting label tiers align strongly with behavioral differences:

- **Class A:** Low delinquency, short offense streaks, and minimal penalties.
- **Class C:** High missed payment frequency, long delinquency streaks, and peak penalty accumulation.
- **Class B:** Transitional behavior that blends some risk signals with signs of recovery or stability.



This confirms that our labeling policy effectively separates risk levels—creating a structured foundation for the model to learn from and supporting actionable, explainable classification downstream.

Model Selection and Training

To train the model, we focused on mature account snapshots—taken 6+ months into the account's lifecycle—and sampled them every 3 months to reduce redundancy. Final outcome records were excluded to prevent label leakage. We used a Random Forest Classifier, well-suited for structured behavioral data, and split the dataset by account to ensure a clean separation between training and validation.

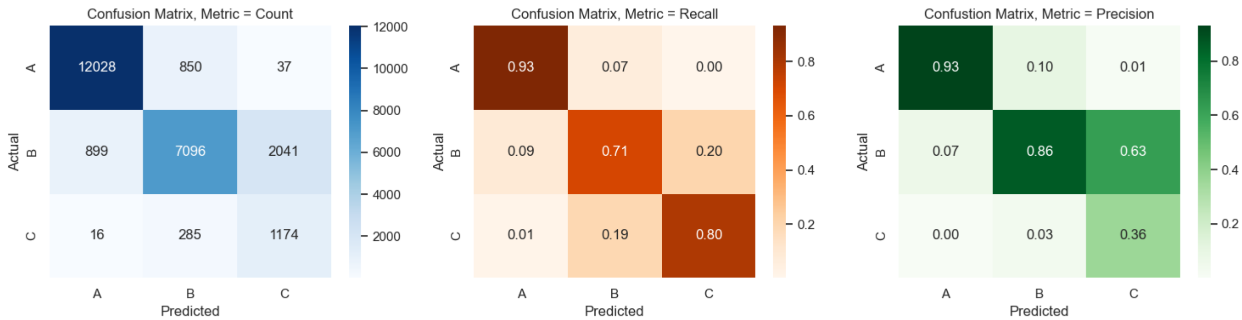
Evaluation Focus

Model performance was assessed across three key areas:

- **Class Separation:** Clear distinction between low (A), medium (B), and high (C) risk groups
- **Confidence Calibration:** High-risk predictions showed strong probability support
- **Early Detection:** Risk signals emerged well before suspension or closure

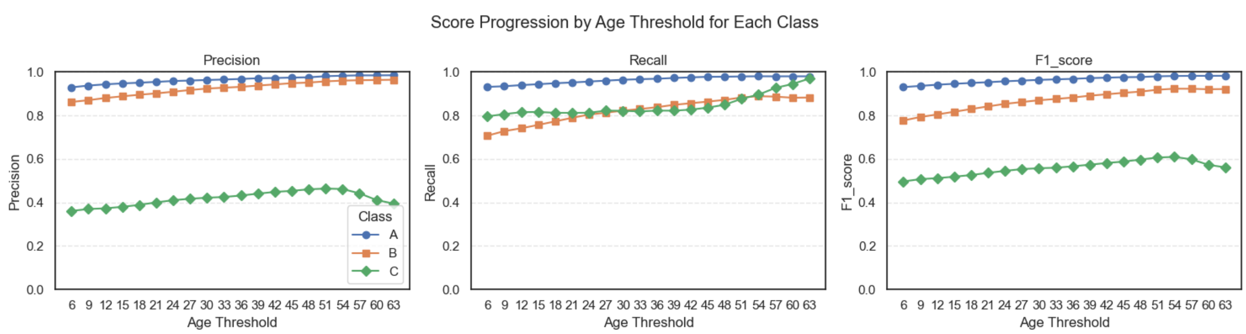
Clear Risk Separation

The model performs strongly across all classes—especially at the extremes, correctly identifying over 90% of low-risk (Class A) accounts. As expected, medium-risk (Class B) accounts are more ambiguous, with some misclassified as high risk due to overlapping patterns.



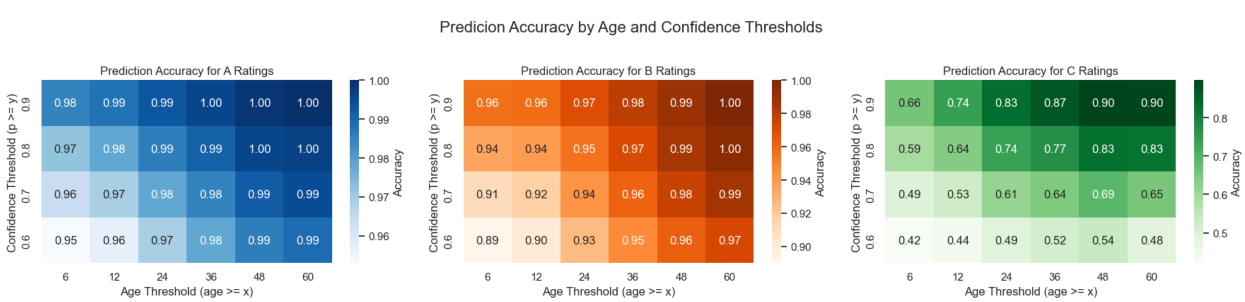
Performance Improves with Age

As more history becomes available, the model's performance improves. Precision, recall, and F1 scores all rise with account age, confirming that behavior over time helps the model separate stable vs. deteriorating accounts.



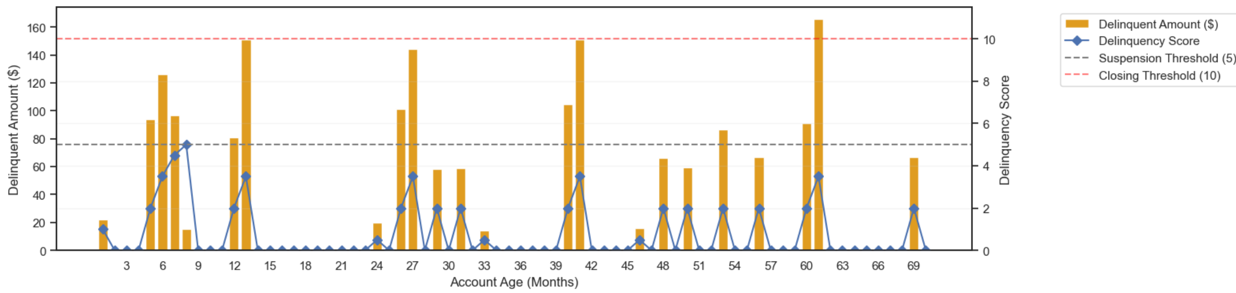
Confidence Aligns with Accuracy

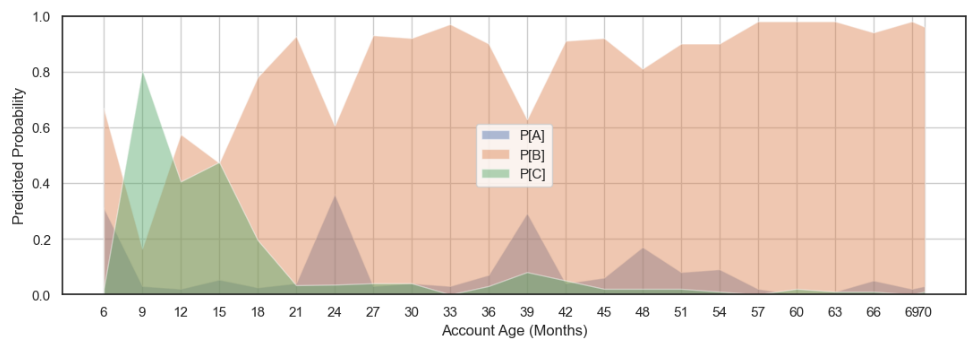
Predictions with ≥70% confidence are highly reliable for Classes A and B. For Class C, accuracy improves significantly at higher confidence and later lifecycle stages—making the model practical for early but cautious intervention.



Example Prediction Trajectory

To bring our prediction to life, let's take a look at how the model tracks this example account over time—and why it ultimately lands on a medium-risk rating:





This account provides an illustration of how the model adjusts its risk assessment over time in response to evolving behavior patterns.

- **Early Trouble:** The account accumulates unpaid balances and hits suspension thresholds early, triggering a high-risk (Class C) prediction.
- **Signs of Recovery:** Around months 24 and 39, improved payments lead to brief dips in risk, showing potential stabilization.
- **Persistent Instability:** These gains fade as inconsistent behavior returns, leading the model to settle on medium risk (Class B).

Overall, the model reads this as a borderline case—too erratic to be low risk, but not severe enough for closure—making it a strong candidate for monitoring.

Prediction Explanation

Finally, to support explainability, we used SHAP values to show which behaviors—missed payments, rising penalties, or suspensions—influenced each prediction. This built transparency and made the model ready for operational use.

Next Step

With a validated and explainable risk model in place, we shift to operationalizing it. In Part 3, we integrate these predictions into a Power BI dashboard, enabling real-time monitoring, account investigations, and data-driven credit decisions.

4. Part 3: Report and Monitoring Using Power BI

Report Requirement

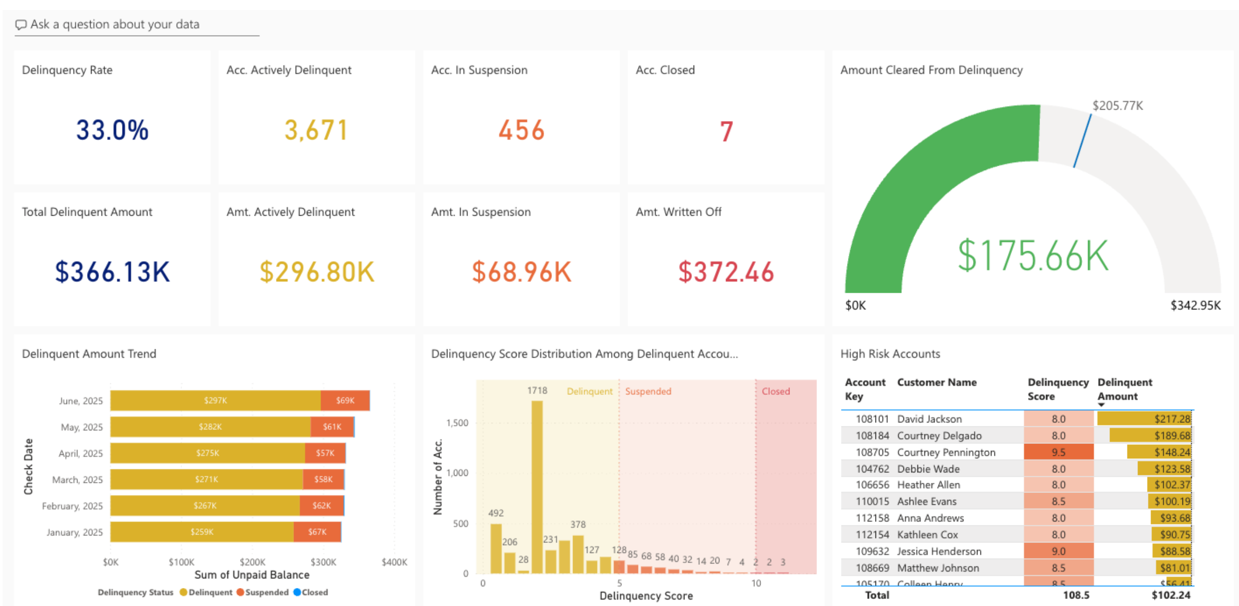
Once we had our machine learning insights, the next step was to make them operational. The goal was to turn complex, time-based model output into a format that's easy to explore and understand—especially for users without a technical background. Risk ratings and SHAP values need more than static tables; they must support fast, informed decisions.

The solution was to build an interactive Power BI report that combines model predictions, behavioral trends, and feature-level explanations—all in one place. The report lets users explore risk from multiple perspectives: portfolio-level summaries, segment breakdowns, and account-level drilldowns.

The Power BI report includes three interactive pages, each designed to support a specific stage of risk monitoring and decision-making. Together, they help users—from executives to analysts—move from portfolio-wide insights to individual account action.

Portfolio Overview Dashboard

This view is built for executives and finance teams who need a quick but complete picture of portfolio health. A finance lead might use it to track changes in delinquency over time, while a risk officer could scan for emerging hotspots or escalating segments.



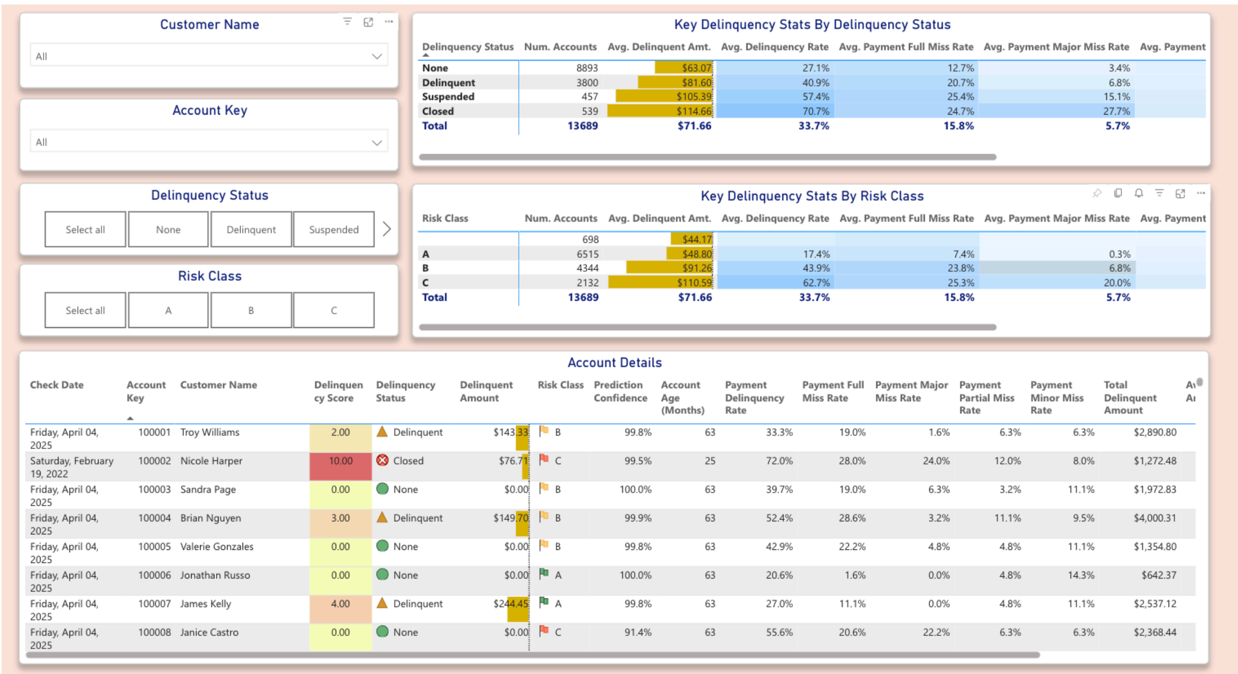
Key Insights Provided:

- *What portion of the portfolio is currently delinquent?*
- *Are risk levels improving or worsening?*
- *Where is risk most concentrated?*

KPIs include delinquency rate, total unpaid balance, recovery progress, and counts of suspended or closed accounts. Visuals like risk distribution charts and top-risk account tables help prioritize attention.

Account Explorer Page

Built for analysts and operational teams, this page enables active management of customer accounts. A collections analyst can filter for high-risk customers, compare behavior across segments, and flag cases for intervention.



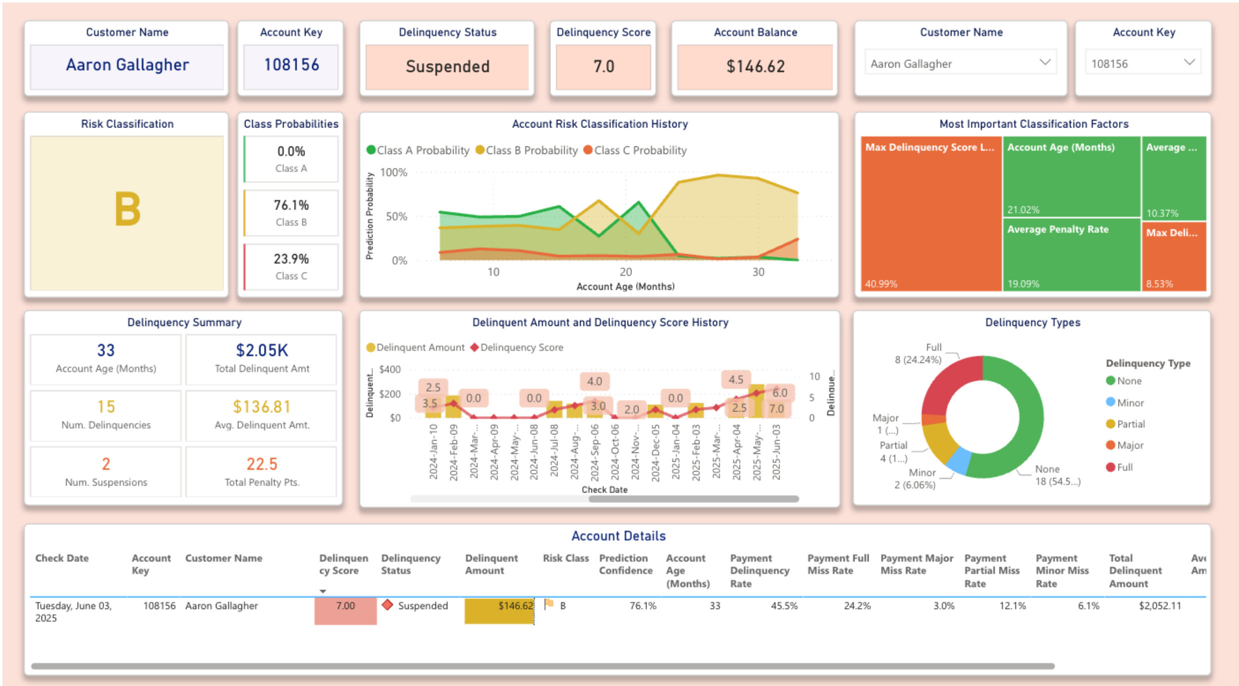
Key Insights Provided:

- *Which accounts are showing risk signals?*
- *What are the historical behavior patterns of these accounts?*
- *How do different risk groups compare?*

Filters include account ID, status, and predicted risk class. Analysts can access delinquency types, frequency, and behavioral averages for selected groups.

Account In-Depth Analysis Page

This diagnostic view is for model validators and support teams. A data analyst can trace which features influenced a score, track changes in confidence, and assess whether signals align with real behavior.



Key Insights Provided:

- *Why is this account labeled as high, medium, or low risk?*
- *How did the risk prediction evolve over time?*
- *What features contributed to the model's decision?*

Visuals include prediction timelines, SHAP feature impact charts, and delinquency histories.

Report Workflow

A typical usage scenario might begin in the Portfolio Overview, where a risk analyst spots a spike in delinquency among commercial accounts. They move to Account Explorer, filter down, and spot trends in penalty scores. Finally, they investigate a specific account via In-Depth Analysis to validate the prediction and identify intervention strategies.

This flow transforms data into insight, and insight into targeted, timely action.

---

5. Case Studies: A, B, and C Accounts in Action

To illustrate how the system works, here are three real-life inspired account examples, each representing one of the risk classes.

🟩 Class A – Low Risk: Stable and Reliable

Sharon's account maintains a clean record across its full 63-month history. Delinquency score stays low, all payments are timely, and no suspension or penalty patterns emerge. The model confidently predicts a Class A rating with 97%+ probability across all periods.

🔎 *Insight*: A textbook low-risk account—no flags, no intervention needed.

🟧 Class B – Medium Risk: Unstable But Recoverable

Brett's account shows recurring partial payments and some missed months, yet avoids reaching the suspension threshold. Penalties accumulate slowly, and the model fluctuates between Classes A and C before converging on B.

🔎 *Insight*: A mixed account—still active, but showing repeated financial strain. Prime for monitoring or early outreach.

🟥 Class C – High Risk: Delinquent and Deteriorating.

Ashley's account builds up high unpaid balances, frequent full misses, and crosses the delinquency threshold by month 24. The model quickly assigns a Class C rating, supported by escalating penalty points and prolonged instability.

🔎 *Insight*: High-risk trajectory confirmed. The model flags this well before write-off, creating a window for early intervention.

---

6. Final Summary and Reflections

This 3-part project series was built around a clear operational requirement: identify risky customer accounts early enough to enable timely intervention—before they escalate into suspension or write-off. Meeting that requirement meant building a system that could track customer behavior, recognize patterns of risk, and surface insights in a way that business teams could use.

Through simulation, machine learning, and interactive reporting, we delivered on this requirement: an end-to-end solution that connects behavioral signals to actionable outcomes.

- In **Part 1**, we generated a utility billing dataset that mirrors real-world account behavior over time.
- In **Part 2**, we trained a machine learning model to detect emerging risk using engineered features and historical patterns.
- In **Part 3**, we built a Power BI report that transforms predictions into practical insights—ready for real-time use by analysts and business teams.

While this was framed for utility services, the framework is adaptable. Whether it's telecom, SaaS, insurance, or lending—any domain where customer behavior unfolds over time can benefit from this kind of predictive monitoring.

This project also demonstrates a full stack of **data and analytical skills,** including:

- **Data simulation and synthetic data generation**
- **Feature engineering and behavioral modeling**
- **Supervised classification and model evaluation**
- **Explainable AI techniques with SHAP**
- **Power BI reporting and stakeholder communication**

By combining these elements into a cohesive pipeline, the project highlights how to design data solutions with tangible operational impact.

Whether you're a data analyst, ML practitioner, or product owner, I hope this walkthrough provides valuable insights on connecting data-driven insights to real-world action.

Linked Series in This Project

This article wraps up a three-part project series focused on modeling, predicting, and monitoring customer credit risk in utility operations. Each phase built on the one before it—moving from raw behavioral logic to real-time insights. You can find the articles for each component here:

- **Part 1 – Data Engineering**
  Developed a simulation engine that generated realistic utility billing behavior across 13,000+ synthetic accounts. The dataset captured usage, payments, penalties, and lifecycle events, providing a rich and controlled foundation for modeling.
- **Part 2 – Risk Classification Model**
  Engineered time-aware features from each account's behavioral history and trained a supervised model to classify risk levels (A, B, C). Predictions included confidence scores and SHAP-based explanations to support transparency and trust.
- **Part 3 – Power BI Dashboard**
  Brought the model output into Power BI—building an interactive dashboard that helps business users monitor trends, explore account-level predictions, and make informed decisions grounded in behavioral data.

If you're interested in building similar solutions—whether for credit risk, churn prediction, or customer lifecycle monitoring—the methods in this series are adaptable across industries.