

AI USE CASE ASSESSMENT

Analyst: Tom Becker, AI Governance Lead
Organization: Corvana Bank
System: LimitWise

Instructions

A use-case assessment is the structured evaluation of whether and how an AI use case should proceed. It is conducted before any data is touched or model trained. Each of the eight sections below opens with an objective that explains what the section is for and why it belongs in the assessment, followed by the guiding questions to answer. Work top to bottom.

SECTION 1: BUSINESS PROBLEM & CONTEXT

Objective: Establish the real-world problem the use case is meant to solve. State the problem in business terms and avoid referencing AI. The goal is to produce a clear-eyed account of what is going wrong today, what the status quo costs, and why acting now is justified. This way, the case rests on a genuine need rather than technology enthusiasm. If the problem cannot be articulated without mentioning the model, that is a signal to pause.

What is going wrong today?

Corvana's warning systems see trouble too late. The incumbent scorecards are backward-looking at the moment that matters most: the months before a customer stops paying, when bills climb and payments shrink. In the last credit cycle the bank charged off 6.8 percent against a 4.9 percent plan, and the board asked why nobody saw it coming. The honest answer is that the instruments could not.

What does the status quo cost?

Missed early interventions and avoidable losses. Customers who could have been helped by a hardship plan reach crisis before anyone calls them. The bank absorbs charge-offs it might have prevented, and it manages its book with a forecast that learns about trouble after trouble has arrived.

Why act now?

The board has directed earlier warning, a prototype exists and back-tests about one statement cycle ahead of the incumbent scorecard, and the next loss forecast is a stated business objective. Acting now also means the action question gets decided deliberately, on the record, rather than under cycle pressure later.

SECTION 2: DECISION(S) THE USE CASE SUPPORTS

Objective: Pin down the specific human decision or process the AI system is meant to inform. List who owns the decision and/or process. This section fixes the system's scope and posture: determining whether the system offers decision-support with a human in the loop or automates the decision outright. This section also sets the purpose limitation by drawing the red lines (i.e., the uses the output must never be put to). Everything downstream to the decision or process should inherit these boundaries.

What does the decision change?

Which accounts receive an intervention, and which intervention they receive. The candidate menu runs from proactive outreach and hardship program offers to credit line decreases, authorization blocks, and account closures. For the customer, these range from a helpful call to a materially harmed account. For the bank, they change loss outcomes, complaint volumes, and the fair lending record.

Who is the human decision-maker?

Today, account actions are decided by credit operations under the Chief Credit Officer's policy. Under this assessment's recommendation, a human reviewer decides every action in the program's first phase. The score brings an account to a person's attention. A person decides what happens.

Is the AI system decision support, or will it make decisions automatically?

Recommendation: decision support, preventive-first. In phase one, scores inform proactive outreach and hardship offers only, each reviewed by a human before contact. No account intervention of any kind, and no automation of any action. The model is unvalidated, its reason coverage is unproven, and a line decrease is adverse action under Regulation B with notice obligations the program cannot yet meet. Account interventions should be considered only as a later phase, gated on validation, proven reasons, and measured outcomes from phase one.

What is the AI system NOT to be used for?

LimitWise output must not trigger any account intervention without human review, and in phase one must not trigger account interventions at all. It must not be used for marketing, new-account underwriting, collections prioritization, or the pricing of other products. No action should be taken on the score during a customer's declared hardship. The sex, education, marriage, and age fields in the extract must not be used as decision inputs pending the fair lending review. Any new use requires a new assessment.

SECTION 3: AI / ML FRAMING

Objective: Translate the business problem into a precise technical problem type. Name the unit of analysis, the target being predicted, and the baseline the model must beat. Then, test honestly whether a simpler non-AI method would do the job. The objective is to keep the team disciplined about whether machine learning is actually warranted, and to make the model's added value (and added risk) an honest comparison rather than an assumption.

What is the technical problem being addressed?

Binary classification at the account level, monthly. Given an account's profile, credit limit, and six months of statement, payment, and repayment-status history, estimate the probability of a payment default in the next statement cycle. The output is a score used to rank accounts for attention.

What is the unit, target, or baseline?

The unit of analysis is an account in a given statement month. The target is payment default in the following cycle, as labeled in the modeling extract. The baseline to beat is the incumbent behavioral scorecard, on lead time and on detection at matched false-alarm rates. If LimitWise cannot beat the incumbent's practical performance, the bank should improve the incumbent instead.

Could a simpler, non-AI/ML solution solve this problem?

Partially. Simpler delinquency rules and the existing scorecard already catch late-stage trouble, and better outreach playbooks need no model at all. What the simpler tools demonstrably lack is lead time. A learned model is a reasonable candidate for that gap, provided the comparison to the incumbent stays honest and the added risk is governed. Section 5 defines the terms of that comparison.

SECTION 4: AFFECTED PARTIES

Objective: Identify everyone the use case touches, not just its users. Identify and list the AI subjects (those being scored or acted on), the decision-makers, and the parties affected downstream of the system's deployment and decision-making. Pay particular attention to those affected who have no say. Surfacing protected groups here is what makes fairness testing possible later. This section is the foundation for the success criteria in §5 and the harm inventory in §7.

Who are the AI subjects?

The bank's roughly 4.6 million cardholders, whose accounts would be scored monthly. They are the people ranked by the system, and they bear the consequences of its errors. The near-prime customer, often building or rebuilding credit, has less cushion against a wrongful account action than a prime customer.

Who are the decision-makers?

Human reviewers in credit operations, as the deciding humans in phase one. The Chief Credit Officer, as owner of credit policy. The Chief AI Officer, as approver of record for program decisions.

Who is affected downstream?

Customers in early difficulty, who may be reached in time or missed. Customers in good standing, who bear the cost of false alarms if interventions ever sharpen. Collections and customer service teams whose queues the score reshapes. Customer Advocacy. Compliance and legal. OCC examiners. Investors who watch the charge-off number.

Which protected groups are present?

The extract contains sex, age, education, and marital status as separate fields. Sex, age, and marital status are prohibited bases under the Equal Credit Opportunity Act. Their presence in the modeling data is a documented fact requiring a documented decision, and per-group measurement of the model's errors is required before any account intervention phase.

Who is affected but has no say?

The customer who never learns a model flagged them. In phase one, a missed customer simply is not called, and a falsely flagged customer receives a call they did not need. Neither appears in any complaint queue. The invisible population is why outcome measurement by group is a condition of this assessment rather than an aspiration.

SECTION 5: DEFINITION OF SUCCESS

Objective: State what success means concretely across three lenses: the business outcome, the model metrics that fit the data (not a default like accuracy), and the fairness criterion with its named groups and threshold. Defining failure just as concretely is part of the job. The goal is a target that is measurable and testable rather than aspirational. With this, later stages can verify whether “success” was met.

What is the specific business outcome the solution will address?

Detect payment trouble a statement cycle earlier than the incumbent scorecard, shift the intervention mix toward preventive contact, and hold charge-offs to plan without harming customers in good standing.

What metrics will be measured to determine success?

Lead time versus the incumbent at matched false-alarm rates. Detection metrics fit for an imbalanced outcome: precision, recall, and area under the precision-recall curve, not accuracy. Preventive-to-punitive mix of actions taken. Hardship program uptake following outreach. Complaint rates among contacted customers. Reason coverage for any future account action.

What fairness criterion will be used to identify whether the solution performs equitably?

Per-group comparison of error rates across sex, age bands, and marital status, using the fields the extract already contains for measurement purposes only. The working criterion is that contact and error rates should not differ materially across protected groups. The precise gap measure and thresholds are finalized during evaluation and recorded before any phase-two consideration.

What thresholds will determine the AI system's success?

To be pre-registered before deployment and graded at retirement. Working targets: measurable lead-time gain over the incumbent at matched false-alarm rates. Outreach precision high enough that reviewers trust the queue. No material adverse per-group gap in contact or error rates without documented justification. One hundred percent reason coverage before any account intervention phase.

What does failure look like?

A repeat of the holiday cuts at model speed would be failure regardless of every other number. So would detection gains that never translate into earlier help, contact patterns that skew against a protected group, or a queue so noisy that reviewers stop trusting it. Failure is defined now so nobody redefines it later.

SECTION 6: CONSTRAINTS

Objective: Catalogue the boundaries the use case must operate within before design begins. Naming the legal/regulatory, ethical/fairness, data/privacy, operational, and resource constraints up front prevents late-stage surprises and defines the guardrails the build must respect. Each constraint named here should have a visible home in the design and, where relevant, in the compliance record.

What legal and regulatory considerations inform the AI system's design?

The Equal Credit Opportunity Act and Regulation B, including the definition of adverse action reaching unfavorable changes to existing accounts (12 CFR 1002.2(c)) and the notice requirements (12 CFR 1002.9). The Fair Credit Reporting Act's adverse action provisions (15 U.S.C. 1681m). The CARD Act and Regulation Z on changes to account terms. UDAAP exposure under 12 U.S.C. 5531. SR 26-2 model risk expectations, implemented by the bank's own policy: high-materiality classification, independent validation, documented rationale.

What ethical and fairness principles inform the AI system's design?

Interventions that help before they harm. No use of prohibited-basis fields or close proxies as decision inputs. Fairness measured and monitored rather than assumed, with special attention to the customer in good standing. The bank's stated commitment not to lead with the stick.

What data protection and privacy considerations inform the AI system's design?

The extract is customer financial data governed by the bank's information security program. Use of account history for model development should be confirmed against the bank's privacy notices. Access is limited to the program team, and no customer data leaves the bank's environment.

What operational considerations inform the AI system's design?

Scoring 4.6 million accounts monthly is an operational system, not a report. Human review capacity must match the queue the threshold creates. The model must fail safe: if LimitWise is unavailable, account management continues under current practice. Reviewer training and a documented escalation path are required.

What resource considerations inform the AI system's design?

A modeling team already committed to the credit calendar, a governance office of three, and reviewer capacity that money can expand only slowly. The action mapping must produce a queue the organization can actually work. A backlog of unworked flags is risk, not coverage.

SECTION 7: INITIAL RISK & HARM INVENTORY

Objective: Produce a first-pass catalogue of the positive and negative harms the use case could cause, including who bears each and the worst/best plausible case. This is not the full risk assessment that comes later. The objective is to see the hazards clearly enough in order to judge whether and how to proceed, and to hand the downstream risk assessment a starting list rather than a blank page.

What negative harms could the AI system cause?

Wrongful account interventions against customers in good standing, at scale, if the mapping ever sharpens without validation. Contact and error patterns that skew against protected groups, including through proxies. Actions the bank cannot explain in a compliant notice. A queue that overwhelms reviewers and turns human review into rubber-stamping. Erosion of trust in the segment the bank exists to serve. Regulatory findings under ECOA, Regulation Z, or UDAAP.

What positive harms could the AI system cause?

Customers in difficulty reached while a hardship plan can still work. Fewer accounts reaching charge-off. A preventive intervention mix the bank can stand behind publicly. A governance file that makes the bank's next model easier to do right.

Of the harms listed, which stakeholders are affected?

Cardholders bear the wrongful-action and skewed-contact harms, with near-prime customers in good standing bearing the most from any repeat of the holiday cuts. The bank bears the legal, reputational, and financial harms. Reviewers bear the queue. The benefits accrue to customers reached early, to the credit function's loss line, and to the bank's standing with its regulators.

What is the worst plausible case that could emerge from the AI system's development?

The team retrains on an extract containing undocumented code values and prohibited-basis fields, bakes an unreadable data problem into a high-materiality model, and the defect surfaces during validation or, worse, during examination.

What is the best plausible case that could emerge from the AI system's development?

Development becomes the proof that the bank can do this properly: the codebook gets reconciled to the data, the fair lending review happens before modeling instead of after, and validation receives a documented, readable package.

What is the worst plausible case that could emerge from the AI system's deployment?

An automated or semi-automated intervention wave harms thousands of customers in good standing, concentrated in a protected group, discovered by journalists or examiners before the bank's own monitoring notices. The holiday cuts, at model speed, with the bank's name on a second story it swore never to repeat.

What is the best plausible case that could emerge from the AI system's deployment?

Customers quietly reached a month before crisis, hardship uptake climbing, charge-offs bending toward plan, and a bank that can show its regulator exactly how every action was decided. Early warning that actually arrives early, pointed first at help.

SECTION 8: GO / NO-GO DECISION & DOCUMENTED ASSUMPTIONS

Objective: Make and record the gate decision: proceed, proceed with guardrails, or stop. Document the guardrails that condition a "go" and, critically, the key assumptions that would change the answer if the assumptions turn out to be false. This creates an auditable decision trail and gives later stages a set of assumptions to monitor and revisit.

Should we build this?

Proceed, with guardrails. The business case is strong and the risks are manageable if the sequence is respected: documentation, validation, preventive phase, measured outcomes, and only then any conversation about account interventions. This is a go for the next lifecycle stage, documented data work and review. It is not a deployment decision, and it is emphatically not an approval of any account intervention.

With what guardrails should the AI system be developed?

Reconcile the codebook to the data before any retraining: every undocumented value in the education, marriage, and repayment-status fields gets a documented meaning or a documented exclusion. Resolve how the risk-enriched sample maps to the portfolio's real default rate. Complete the fair lending review of the sex, age, education, and marriage fields. Obtain independent validation as required for a high-materiality model.

With what guardrails should the AI system be deployed?

Preventive-first mapping only, per Section 2, with human review of every contact. Reviewer capacity matched to the queue before go-live. Monthly measurement of contact outcomes by group, with a stop-and-investigate trigger for material skew. Proven reason coverage before any future phase touches an account's terms. A rollback plan that returns account management to current practice.

What assumptions, if false, would change the organization's decision to develop and deploy this AI system?

That the extract's default labels are accurate and consistently defined. That the undocumented code values can be resolved to real meanings. That account history may lawfully be used for this development under the bank's notices. That the risk-enriched sample can be calibrated to portfolio reality. That reviewer capacity can match the queue. If any of these fails, this assessment returns to the gate.