MACHINE LEARNING CASE STUDY

Building ML Models with Claude Code for Automation and Efficiency

Roger Estrada · Machine Learning in Business · AI Strategy · 2026

Claude Code lets one operator move from a business question to a tested, documented, deployed model in a short working session, without waiting in a development queue. Cycle time and reach improve. Judgment, validation, and approval stay with people.

What Claude Code is

Claude Code is an agentic development tool that works inside the operator's own files and command line. It reads data and code, writes and runs scripts, checks the results, and iterates across multi-step tasks. The skill is directing it toward a business outcome, not typing code. That is what closes the usual gap between executive vision and engineering capacity: the operator can prototype the system being proposed, not only describe it.

The workflow: what Claude Code does and what a person owns

Stage	What Claude Code does	What the human owns
1. Frame the decision	Drafts the problem statement, success criterion, and cost of errors from the operator's brief	The decision, the metric, and the threshold
2. Prepare the data	Profiles the data, flags missing values, outliers, and leakage, and writes the cleaning code	Imputation and exclusion rules, data access, and privacy
3. Build and compare models	Trains a baseline and challengers, tunes with cross-validation, and compares them on AUC or RMSE	Model choice against the cost of errors, and bias review of the features
4. Explain the results	Produces feature attribution, ROC and diagnostic charts, and plain-language findings	Confirming the interpretation with the business owner
5. Deploy	Wraps the model in a Streamlit app or a scheduled script and writes tests	Release approval, monitoring, and retraining triggers
6. Document	Writes assumptions, limits, and handover notes	Sign-off for governance and independent validation

The models in this library follow that sequence: business framing first, success criteria set before training, several models compared against a baseline, and the metric matched to the cost of errors.

Where it drives automation and efficiency, and the controls

Efficiency. The same workflow produces use-case scoring models for an AI and automation intake queue, risk and propensity models for lending and member retention, anomaly detection on

operational data, search over policy and procedure documents, and internal dashboards and decision-support tools. Each replaces a manual review step or a slow handoff and can be prototyped in days instead of a project quarter.

Fit in an enterprise AI function. The function uses this capability to prototype quickly and put working tools in front of business owners early. That reserves scarce engineering capacity for production hardening and integration.

Controls.

- Every model is validated on held-out data against a success criterion set before training.
- Code is reviewed and tested before use, and the work stays under version control.
- Sensitive member data is handled only in approved environments.
- Prototypes stay labeled as prototypes until they pass the same validation and change control as any model that influences a decision.
- Model documentation covers assumptions, limits, and bias exposure, and is reviewed by risk and compliance partners.