Agentic Reinforcement Learning for Competitor-Aware Customer Retention

verizon✓

BUSINESS PROBLEM

With over 97% U.S. adult mobile penetration, wireless carriers compete primarily on retention, where acquiring a new customer costs 5-25x more than keeping an existing one. At Verizon, campaign creation is a manual, weeks-long process coordinated across 45+ stakeholders, and once live, campaigns have no mechanism to adapt to customer outcomes or competitor moves. At Verizon's scale, even a fractional reduction in churn preserves hundreds of millions in annual revenue, creating urgent need for an automated, continuously learning system.

DATA SOURCES

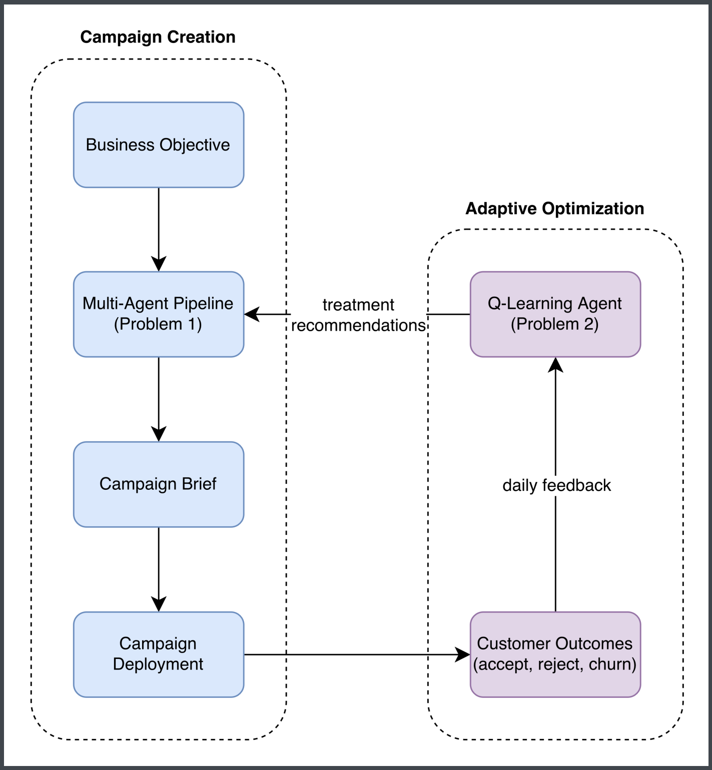
(1) Historical offer data, capturing promotional offers extended to customers with acceptance/rejection outcomes and churn decile scores
(2) Customer churn data recording disconnections with port carrier information
(3) Competitor offer data scraped from rival carrier websites capturing plan pricing and features.

DATA TYPES AND FORMAT

Boolean, Categorical (nominal), Numeric, Text

APPROACH

I designed a closed-loop system with two components. A multi-agent AI pipeline automates campaign creation (audience selection, treatment recommendation, content generation, brief assembly). A Q- learning agent I built from scratch performs dynamic treatment selection over a 12-state MDP (churn risk x competitor attractiveness) with 15 actions across three cost tiers. The reward function grounds decisions in economic value (LTV minus cost, weighted by take rate). An offline pipeline trains daily on outcomes.



Benjamin Maltbie
MIT Leaders for Global Operations, Class of 2026

MIT Leaders for Global Operations

IMPACT

The system demonstrated measurable improvements across multiple dimensions. The Q-learning agent achieved correct treatment preference learning in all 24 experimental conditions tested across 8 validation experiments, a 100% correct preference learning rate. In the standard simulation with 5,500 customers, the agent converged in 161-186 days, developing Q-value separations of 5.8x to 9.4x between optimal and second-best treatments, indicating decisive confidence in recommendations. The agent proved robust to 10% label noise, incurring only a 13-day average convergence delay while still identifying correct treatments. It converged across all acceptance rate scenarios with less than 6 days of variation, and behaved predictably across learning rates spanning a 30x range.

On the operational side, the multi-agent campaign creation pipeline compressed what was previously a multi-week, multi-team stakeholder process into an automated workflow. The architecture I designed was validated when a Verizon engineering team successfully built the remaining agents based on my design. The system enables same-day competitive response, where a rival carrier's new promotion could trigger an automated counter-campaign rather than requiring weeks of cross- functional mobilization.

At Verizon's scale of tens of millions of subscribers, the precision of the RL agent's treatment differentiation, combined with faster campaign cycles and continuous learning from observed outcomes, positions the system to preserve hundreds of millions of dollars in annual revenue through improved retention and more efficient allocation of marketing spend.

DRIVERS

Similar to business problem, Verizon is facing a competitive telecom market that it needs to be able to deploy new marketing offers quicker and retain customers more efficiently. No closed-loop mechanism existed to adapt treatments based on customer outcomes or competitive dynamics. At Verizon's scale, even fractional churn reduction compounds into

BARRIERS

I faced a few barriers during the project. There was fragmented data across multiple internal systems requiring joins across customer, offer, and churn datasets. Campaign creation involved many stakeholders with no single end-to-end owner. LTV 3.0 API was still under development during the POC phase, requiring mock integration. Verizon internal lay

ENABLERS

All stakeholders I encountered during my project were supporting and willing to help. This included my manager, leadership, and related teams. I also relied on the Verizon engineering team to implement the majority of the agentic marketing campaign creation system (I primarily did design for this). Challenges I faced were largely related to working within a large company with many processes and outside factors.

ACTIONS

I designed the multi-agent campaign creation architecture and built the Treatment Selection Agent with segment-level cost-aware scoring. I also built the Q-learning system from scratch: model, 4-step inference workflow, and offline training pipeline. In this, I formulated the 12-state MDP with an economically grounded reward function and conducted simulation experiments validating convergence, noise robustness, and hyperparameter sensitivity.

INNOVATION

This system is novel in combining agentic AI for campaign creation with reinforcement learning for continuous treatment optimization into a single closed loop. The joint state representation encoding both customer churn risk and competitor attractiveness enables the agent to adapt treatment strategy in response to competitive dynamics, not just customer behavior, and is interpretable while still capturing the essential decision structure.

IMPROVEMENT

This was a proof-of-concept so I don't have production metrics to show improvement. From the simulations, we saw 100% correct treatment preference learning across all 24 experimental conditions. Campaign creation cycle compressed from multi-week manual process to an automated pipeline.

BEST PRACTICES

Start with a compact, interpretable state space before scaling to complexity. Ground the reward function in real economic value rather than proxy metrics. Design for modularity and validate extensively through simulation before production deployment. Maintain human-in-the-loop oversight at the final validation step.

OTHER APPLICATIONS

The framework generalizes to any enterprise with sequential customer treatment decisions (e.g.
insurance policy retention, subscription service churn prevention, financial services cross-selling).
The multi-agent campaign creation architecture is applicable to any marketing organization with complex, multi-stakeholder workflows.