Respectful Offers: keep subscribers without pestering them (simulated results)

PR/FAQ · data: WSDM Cup 2018 KKBox churn prediction challenge (Kaggle), used for non-commercial research and education · offer responses simulated

Press release

Simulated results. A retention policy that decides which subscriber gets a content nudge, a cheaper plan or a discount, and that never repeats an offer a member declined, spent **97 NTD** of discount per extra member kept, against 211 NTD for a rule-following baseline and 289 NTD for today's habit of discounting everyone at risk. It was evaluated on 200,000 real members of the KKBox music service, whose responses to offers came from a pre-registered simulator: KKBox's data holds no offers.

Respectful Offers predicts each member's churn risk, removes every offer the two respect rules forbid (a declined offer stays declined; members at risk see content before commerce), and lets a Thompson sampling bandit choose among what is left, within a monthly discount budget. An independent auditor recomputed every rule from the decision log: the respectful policies broke none, the status quo broke 31,490 (central variant).

The pre-registered claim was that the bandit keeps more members per 1,000 NTD of discount than the rule-following baseline. It held in all three simulator variants, bandit minus baseline with 95% intervals: conservative +1.74 (+1.28 to +2.22), central +5.57 (+5.06 to +6.10), optimistic +9.36 (+8.73 to +10.02) members per 1,000 NTD. The bandit spent 33% of the baseline's discount for 73% of its extra members kept (7.8 against 10.8 per 1,000), leaning on free nudges and the cheaper plan. Read every number as conditional on the simulator's assumptions.

FAQ

Are these real results? The members, their behaviour, their churn labels and the churn model (AUC 0.84 against 0.72 for a two-feature baseline) are real. Every outcome that depends on an offer is **simulated**: the simulator's values were frozen by their sha256 before the real run, and every result is reported under a conservative, a central and an optimistic variant.

What would a real A/B test add? The one thing a simulator cannot give: how real members respond to offers. A live test, randomised at the member level with the same holdout, would measure acceptance and the lift in renewal directly, check whether sleeping dogs exist at the assumed share, and catch effects the simulator leaves out (offer fatigue across months, word of mouth, a discount teaching members to wait for the next one). Until then, the conclusion is: under these assumptions, this policy is worth testing.

Why a holdout? A randomly chosen 10% of members are never offered anything. Their retention, against the no-offer policy on the evaluation members, checks that the lift is not a measurement artefact: the difference was +1.0 per 1,000 (95% interval -1.8 to +3.8), consistent with none.

Why a doubly robust estimator? A live system would only have its own logged decisions. Doubly robust off-policy evaluation combines a model of the reward with importance weights, so it stays unbiased if either one is right. Here it was validated against the simulator's exact answer: its 95% interval covered the true value in 12 of 12 policy and variant cases.

Why do the rules sit outside the learning? A bandit that learns its own respect would trade it away whenever pestering pays. The filter removes forbidden actions before the bandit chooses, so no reward can buy a violation, and the auditor checks the log independently of the code that made the decisions.

What does respect cost? The status quo, which ignores both rules, kept 11.7 extra members per 1,000 against the rule-following baseline's 10.8 (difference +0.88, 95% interval +0.43 to +1.35), and paid 289 NTD per extra member against 211 (central): a little more retention, at a higher price per member kept and with members asked again after saying no.