Cleaner data, better metrics, and a simpler model deliver stronger, actionable churn predictions

Before

The inherited design was limited because:

1. Rows with missing info were dropped, and redundant features were kept, leading to messy data
2. Used Gradient Boosting model without testing if other models could perform better
3. Only reported ROC-AUC, a metric poorly suited for imbalanced datasets like this churn dataset

After

Improvements made to address these:

1. Cleaned and standardized data by encoding missing values as ‘Unknown’
2. Compared multiple models – Logistic Regression won on performance and easy interpretation
3. Introduced PR-AUC and Top-10% precision metrics to rank customers by churn risk

With the new pipeline, outreach can focus on the 10% highest risk customers, capturing **~40% of churners**.

See speaker notes for definitions of all metrics.

Elizabeth Wang

1

2

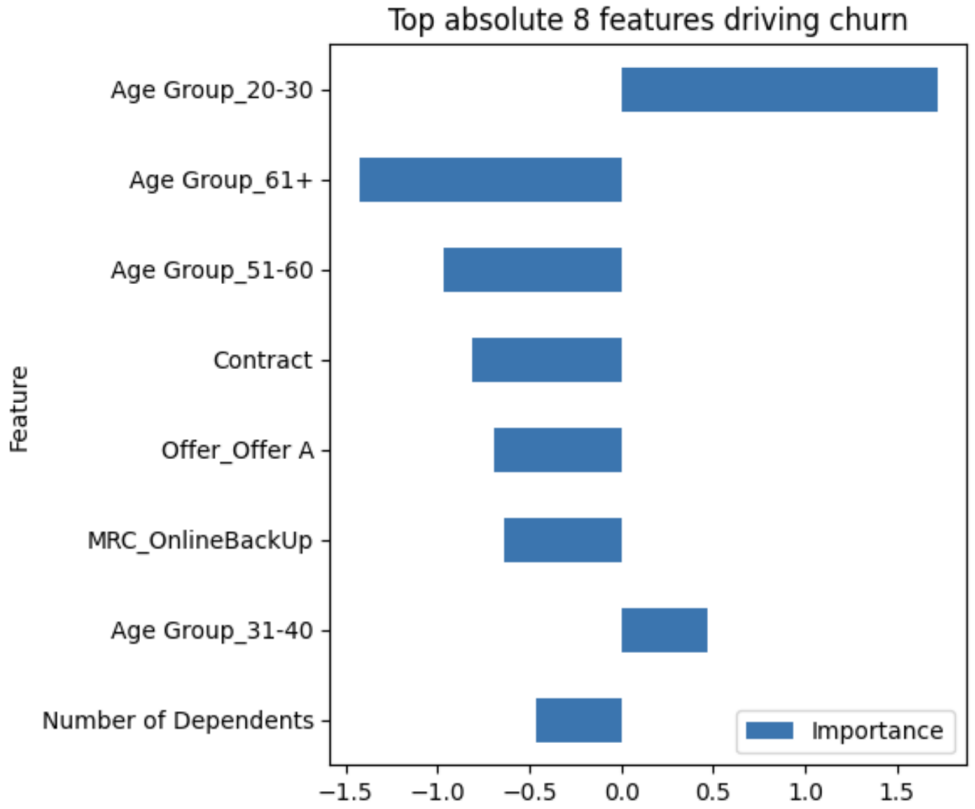
Cleaning messy, redundant inputs and switching models improved performance and revealed key churn risk factors

What changed in the model:

- Dropped redundant features (e.g. duplicate spend totals)
- Encoded missing categorical values as ‘No Offer/Unknown’
- Used Logistic Regression model to provide clearer insights into the key drivers of churn

Key drivers of churn:

- **Younger** (20–40) customers more likely to churn
- **Month-to-Month** contracts are more likely to churn
- Certain offers (**Offer A**) and paying for **Online Backup** reduce churn risk
- Customers with **more dependents** tend to stay



Younger individuals on **short-term contracts are more likely to churn** than those on **longer-term family plans.**

Elizabeth Wang

3

New metrics highlight real performance gains and business impact

What changed in the model:

- Switched to PR-AUC because ROC-AUC is most useful with balanced data (~50/50), whereas churn rate is only 8%
- Ranks customers by churn risk so marketing can target the top-K% given our budget

What improved:

- PR-AUC increased from 0.258 to 0.277, showing **stronger ability to identify churners**
- For the top 10% highest-risk customers flagged by the model, precision rose from 26% to 31% while recall stayed around 40%, meaning **fewer wasted calls without losing coverage**
- Our F1 score improvement reflects a **better balance between catching churners and avoiding over-calling**
- Our lift shows that our outreach is nearly **four times more effective than random selection**

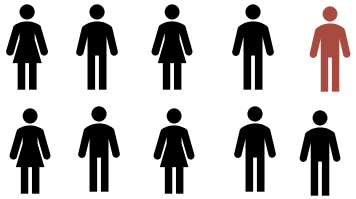
Metric	Original*	New
ROC-AUC	0.814	0.823
PR-AUC	**0.258**	**0.277**
Precision@10%	26.1%	31.0%
Recall@10%	38.7%	38.8%
F1@10%	0.312	0.345
Lift@10%	3.84×	3.87×

Our new model delivers a **ranked churn-risk list,** enabling outreach that's ~4x more effective than random selection, with fewer wasted resources and no drop in how many churners reached.

*Most metrics were added to the original pipeline for comparison

Elizabeth Wang

Next Steps



Step 1 – Generate ranked list and align on K. Train the model on historical data, score current customers, and rank them by churn risk. Pick our K cutoff based on budget. If possible, perform cost sensitivity analysis weighing the cost of missed churners against the cost of contacting customers.

Step 2 – Launch targeted outreach. Engage our top-K risk group, capturing a large share of churners while minimizing wasted efforts.

Step 3 – Personalize offers using churn drivers. Target key churn risks (short contracts, individual plans) or, if time, use SHAP to show the 1–2 reasons each flagged customer is high-risk. Tailor offers to customer needs (discounts, loyalty perks, onboarding support).

Step 4 – Validate through A/B testing. Hold out a control group within the top K% to measure true churn reduction and track KPIs (churn rate delta, net customers retained).

Step 5 – Re-score monthly, watch for drift, and retrain if churn patterns or key features shift.

Optional / Further Exploration

- Test alternative methods for missing demographic data (beyond dropping nulls)
- Run deeper hyperparameter tuning, calibration, and alternative models (e.g., XGBoost, CatBoost)
- Explore additional features such as recency, usage deltas, or service interactions

Elizabeth Wang