Predicting Customer Churn

SyriaTel • Phase 3 (Business Stakeholder Summary)

Goal: identify customers at high risk of leaving so retention teams can intervene early and reduce lost revenue.

Output: a churn risk score per customer + a recommended outreach threshold.

Prepared for business stakeholders (plain language, action-focused).

Why churn prediction matters

Churn is expensive. Retention is usually cheaper than acquisition, but retention capacity is limited.

What the model enables

- Prioritize outreach to the right customers
- Reduce missed churners
- Avoid wasting offers on low-risk customers
- Measure lift with A/B testing

How to think about success

We focus on catching churners (recall), while keeping the outreach list manageable (precision + volume).

A probability score lets the business choose: “target top 10%” or “target all above threshold X”.

Retention decision = model score + business capacity.

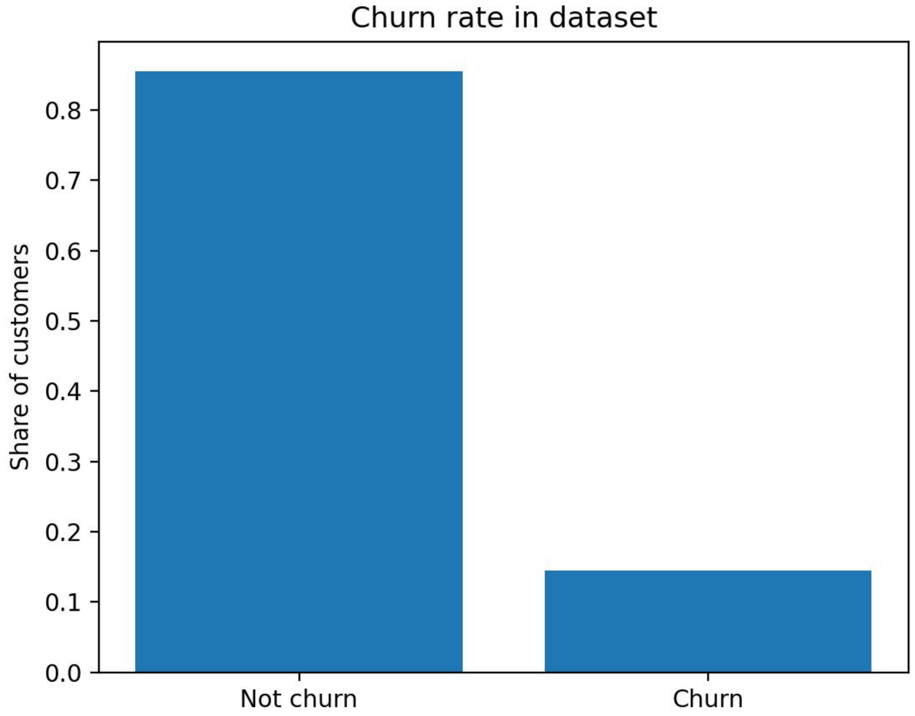
Data used (what we know about each customer)

Dataset: 3,333 customers with usage + service signals and a churn label.

Feature groups

- Call usage (day/eve/night/intl minutes + calls)
- Plans (international plan, voicemail plan)
- Account tenure (account length)
- Customer service calls (service friction signal)
- Geography (state, area code)

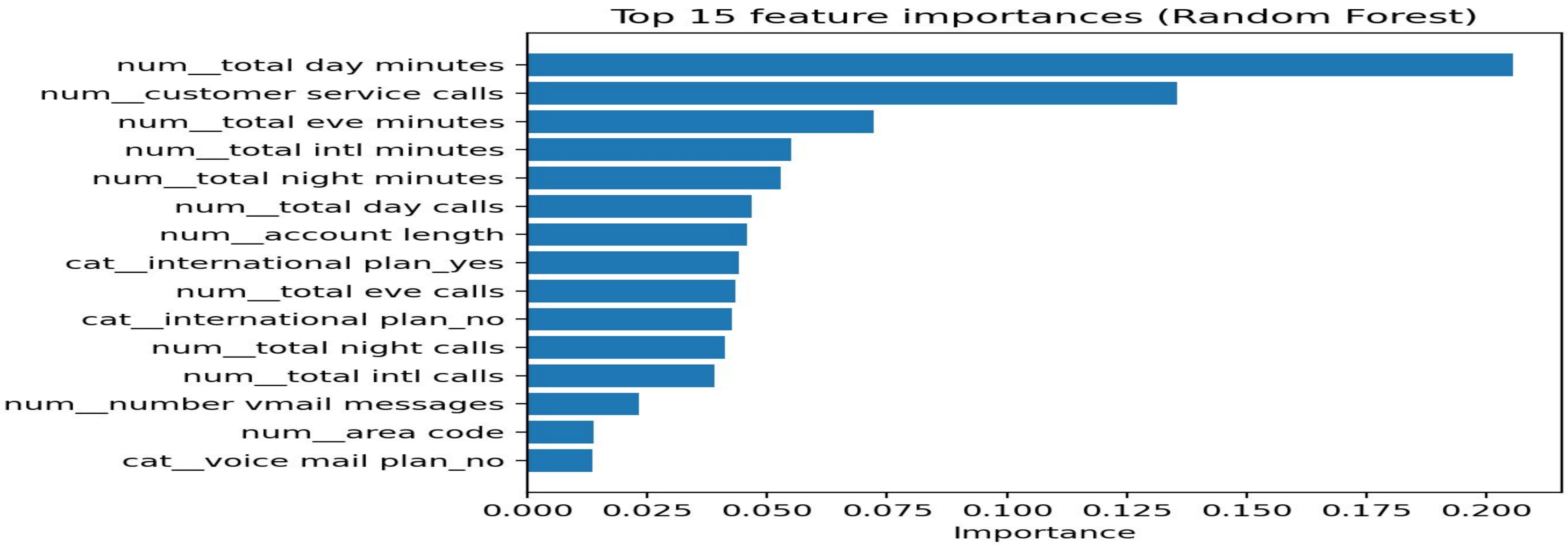
Churn rate



We removed obvious identifiers (phone number) and redundant charge fields.

Predictable patterns (what tends to signal churn)

The model learns patterns that correlate with churn. These are signals to investigate—not proof of causation.



Business interpretation (examples):

- Many customer service calls can indicate unresolved issues
- Heavy day/evening usage may reflect price sensitivity or plan mismatch
- Plan features (international/voicemail) can segment risk profiles

Method (in plain terms)

We train a model on past customers: it learns which combinations of signals tend to precede churn.

Customer data

Usage, plans,
service calls

Model

Learns patterns
from history

Risk score

Probability of churn
per customer

Decision rule (business-in-context)

Choose a threshold to fit capacity. In this project we use 0.35, which flags ~14% of customers for outreach.

If capacity changes, move the threshold up/down (or target top N% of highest-risk customers).

Key point: the model scores risk; the business chooses the outreach cut-off.

Results (test set)

At the chosen threshold, the model catches most churners while keeping the outreach list focused.

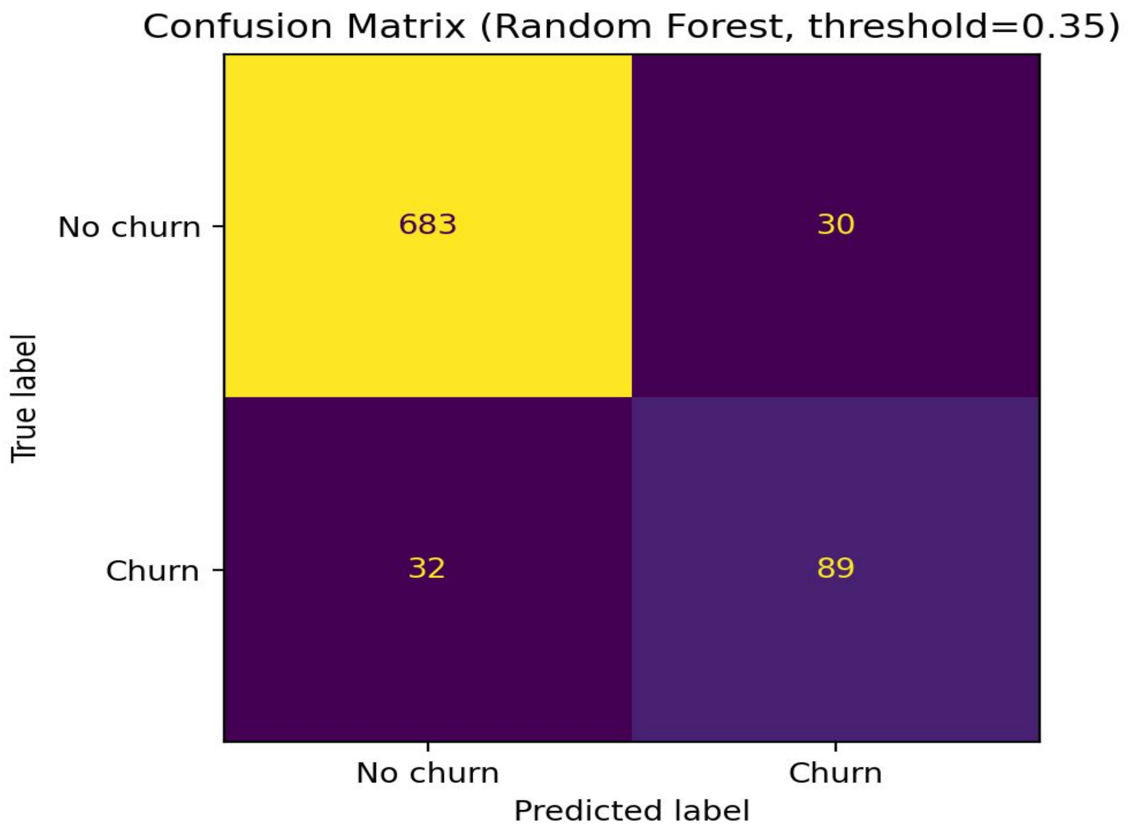
Recall (catch churners)	Precision (low waste)	ROC-AUC (ranking quality)	Flagged for outreach
73.6%	**74.8%**	**0.885**	**14%**

How to interpret these numbers

- If 100 customers would churn, the model flags about 73.6% of them in time to intervene.
- Of the customers we flag, about 74.8% truly churn (so offers aren't wasted).
- Only ~14% of customers are flagged at this threshold (manageable outreach).

What errors look like (confusion matrix)

This shows how many customers we correctly/incorrectly flag for retention outreach.



In business terms

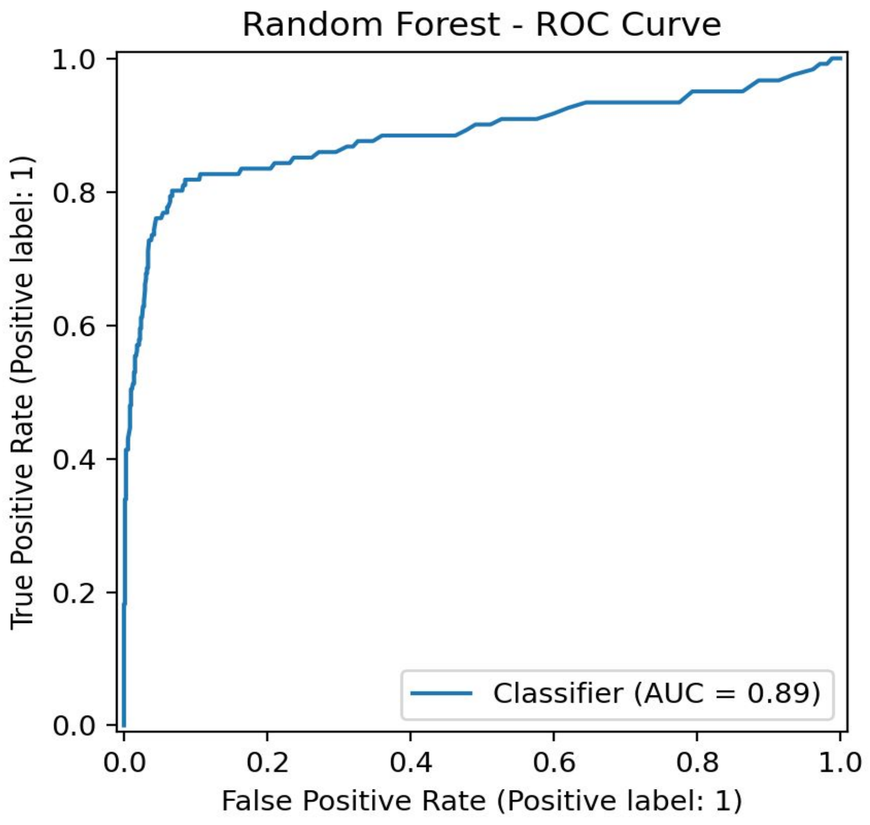
- True positives: churners we catch (high value)
- False negatives: churners we miss (lost revenue)
- False positives: we contact a non-churner (some cost)

- We tune the threshold to reduce missed churners without overloading the retention team.

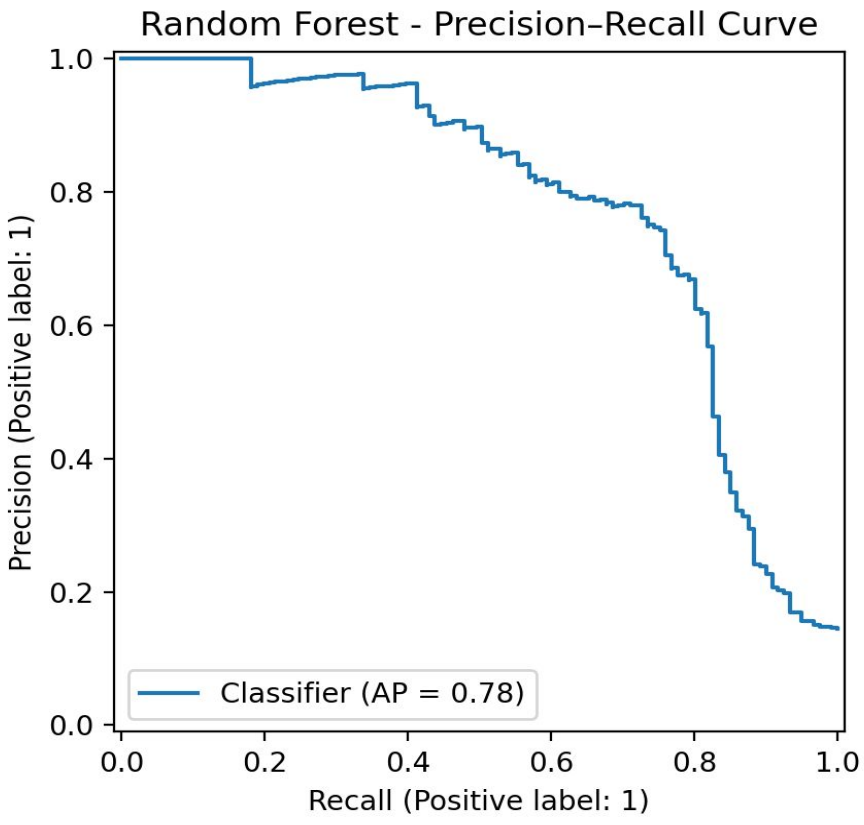
Model quality (ranking performance)

These curves summarize how well the model separates churners from non-churners across all thresholds.

ROC curve



Precision–Recall curve

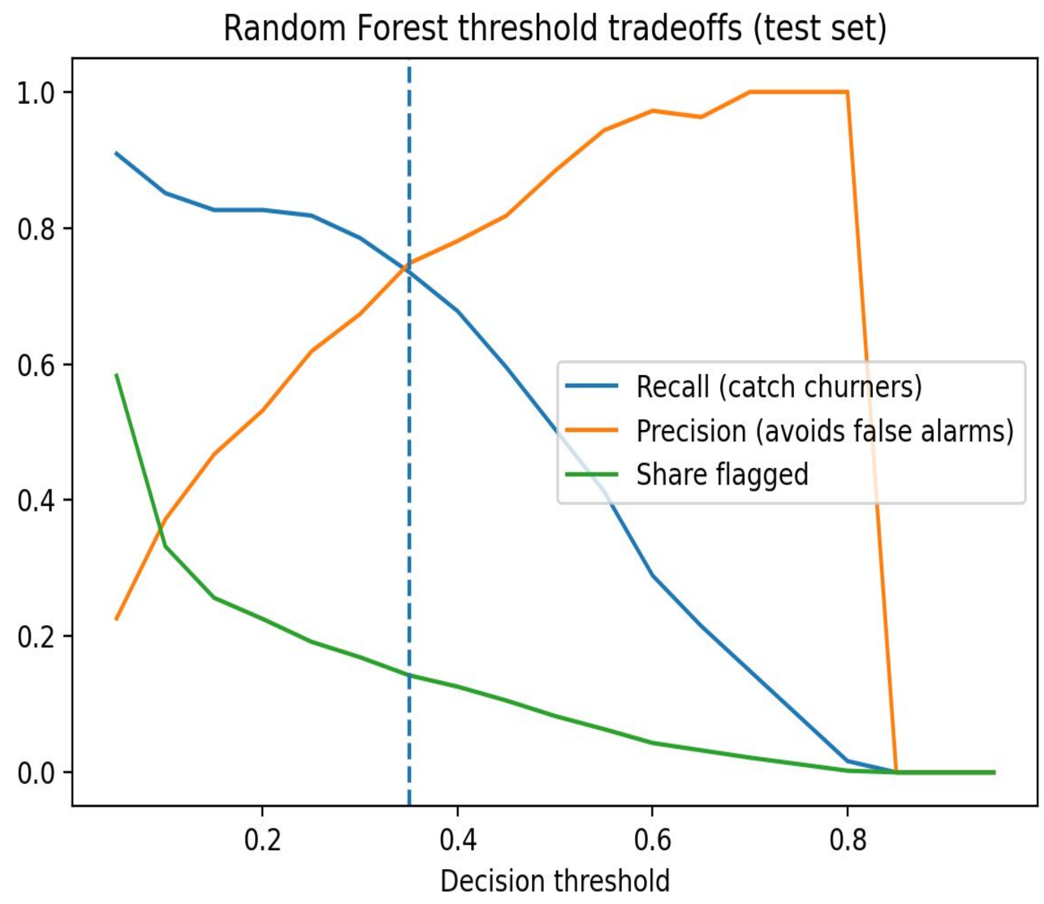


How to operationalize (simple rollout plan)

Recommended workflow

- Score customers weekly/monthly
- Rank by churn probability
- Select top group (threshold/top N%)
- Route to retention playbook
- Track outcomes (saved vs churned)
- Re-train + monitor drift quarterly

Threshold tradeoff



Conclusion & next steps

Bottom line

This model can flag ~14% of customers for outreach while catching ~73.6% of churners (test set).

Next steps

- Pilot the retention list for 4–6 weeks
- Run an A/B test to measure churn reduction
- Build a simple dashboard: flagged volume, conversion, churn saved
- Retrain periodically and monitor performance drift

Limitations (important)

This dataset is a snapshot. Real-world churn changes over time, so production performance should be monitored and the model retrained.

Thank you — questions?