Ravenstack Churn Prediction: Low Activity Is the Strongest Leading Indicator

1. Executive Summary

Key Finding:
96.8% of churned accounts had fewer than **6 active days** (on a 30-day scale) in the full subscription month **preceding their churn**.

Secondary Signals (in that same pre-churn month):

- Downgrade occurred: **5.71%**
- Median Ticket Response Time (TTR) > 48 hours: **5.71%**
- Customer Satisfaction Score (CSAT) < 3.5: **3.53%**

False Signals Removed:
Accounts that were still in *trial* mode at the time of churn: **0%** — indicating trial status is **not** a risk signal.

---

2. Recommended Business Actions

Generate a High-Risk Customer List based on simple, interpretable rules:

- If usage_days_m < 6 in the past 30 days → mark as **At Risk**
- If **any** of the following also occur:
  is_downgrade = 1 OR TTR > 48h OR CSAT < 3.5 → mark as **High Risk**
  Otherwise → **Medium Risk**

Trigger CSM Playbooks:

- High Risk: Prioritized human outreach, CSM call queue, resource allocation
- Medium Risk: Focus on product onboarding, activation and feature engagement

Embed into Weekly Automation Loop:

- Run this rule set weekly
- Refresh high-risk list
- Auto-distribute to the CSM team as a rolling priority list

3. Key Risk Signals (Chart Summary)

Signal Type	% of Churned Accounts
Active Days < 6/30	96.81%
Downgrade (is_downgrade)	5.71%
Median TTR > 48h	5.71%
CSAT < 3.5	3.53%
Still in Trial	0%

Part 2 – Technical Implementation & Reproducibility

4. Methodology Overview

To eliminate hindsight bias, we built a **monthly account panel** using the 5 raw files provided:

- subscriptions.csv used as the **monthly time anchor**
- Merged with: feature_usage.csv, support_tickets.csv, churn_events.csv, and accounts.csv

Observation Windows:

- For churned accounts: The **full subscription month BEFORE churn**
- For all accounts: The **latest available full month**

Key Feature Definitions:

- usage_days_m: Count of distinct active days in feature_usage
- is_downgrade: Flag from subscriptions.csv
- ttr_p50_m: Median ticket response time per month (from support_tickets)
- csat_m: Average CSAT in that month

5. Risk Scoring Logic (Deployment Ready)

```
if usage_days_m < 6:
  if is_downgrade == 1 or ttr_p50_m > 48 or csat_m < 3.5:
    risk_score = 'High'
  else:
    risk_score = 'Medium'
```

```
else:
  risk_score = 'Low'
```

Interpretation:

- **High Risk**: Requires CSM urgent attention
- **Medium Risk**: Focus on activation/product education
- **Low Risk**: No action needed; continue monitoring

---

6. Data Traceability & Deliverables

- high_risk_accounts.csv: Full list of Medium/High Risk accounts with fields:
  account_id, mrr, usage_days_m, downgrade, ttr, csat, risk_score
- high_risk_top20.csv: Top 20 high-risk accounts ranked by MRR × Risk Level
- **Data validation**: ARR ≈ 12 × MRR was 100% consistent in subscriptions.csv
- **Code reproducibility**: Core logic implemented in reproducible pandas snippets (attached in README)

---

7. Next Questions for Real-World Deployment

If this were a real SaaS client scenario, we would recommend further investigation into:

Root Causes of Low Activity:
Is it due to poor onboarding, insufficient product value, seasonal drop-off, or UX friction?

Feature Usage Segmentation:
Can we build deeper metrics like "sticky features" or "module engagement" to improve signal precision?

Segment-Level Risk Factors:
Are certain industries, countries, or referral channels more prone to low usage AND downgrade?