Energy Customer Churn & Price Sensitivity

One-page case study | EDA → Feature Engineering → Modeling → Threshold Decisioning

Author: Owen Nda Diche

ROC-AUC	F1 (tuned @ 0.194)	Recall (Churn, tuned)	Churn rate (test)
0.705	**0.349**	**0.363**	**9.72%**

Situation & Question

Situation: The client is losing customers (churn). A working hypothesis is that customers are becoming more **price sensitive** and switching to lower−cost options.

Question: Can we predict churn risk and quantify the role of price−related behaviour to guide retention?

Data Snapshot

Datasets: (1) Client usage/tenure/value data + churn label, (2) Historical pricing data.
Join key: id (pricing covers **100%** of client IDs).
Split: Train **11,684** | Test **2,922** (stratified).
Imbalance: Churn rate in test set is **9.72%** → accuracy alone can be misleading.

EDA Takeaways

Types: Date fields parsed successfully (e.g., date_activ, date_end, price_date).
Missingness: Largest gap is **channel_sales ~25.5%**; most other fields are complete.
Distributions: Heavy right−skew/outliers in consumption & margin fields → scaling/outlier−aware analysis needed.
Categoricals: channel_sales contains a dominant level + missing category; origin_up is concentrated in top levels.

Method

Pipeline
• Clean & validate types/missingness
• Engineer features: tenure, usage/value, pricing aggregates
• Train Random Forest with *class_weight='balanced'* (imbalance-aware)
• Evaluate using ROC−AUC + Precision/Recall/F1
• Tune the decision threshold for retention objectives

Model Performance (Test Set)

Baseline (all non−churn): Accuracy 0.9028 | Recall 0.0000 | F1 0.0000
RF @ 0.50: ROC−AUC 0.705 | Precision 0.8889 | Recall 0.0845 | F1 0.1543
RF tuned @ 0.194 (max F1): Precision 0.3355 | Recall 0.3627 | F1 0.3486

Takeaway: AUC shows useful ranking power. Threshold tuning increases churn capture from **8.5%** to **36.3%** (~4×), trading off precision for recall.

Drivers & Price Sensitivity

Top drivers (feature importance)
• **Customer value:** margin_gross_pow_ele, margin_net_pow_ele, net_margin
• **Consumption:** cons_12m, cons_last_month, imp_cons, forecast_cons_12m
• **Tenure/changes:** months_activ, months_modif_prod
• **Price sensitivity proxies:** spreads/variability + month−to−month price differences

Insight: Pricing matters, but churn is also strongly linked to customer value, usage behaviour and tenure.

Recommendations & Next Steps

Recommendations
1) Use the model as a **risk−ranking** tool for retention targeting; start with threshold **0.194** as a balanced operating point.
2) Choose threshold based on business costs (cost of offers vs cost of missed churn).
3) Run an **A/B uplift test** on flagged customers to measure ROI and refine segmentation (including a price−sensitive segment).

Note: F1 and Recall are reported at tuned threshold = 0.194. Artifacts: 01_EDA, 02_Feature_Engineering, 03_Modeling, 04_Threshold_Tuning.