Telecom Churn Prediction Report

Project Goal: Predict customer churn for a telecom company using behavioral data and machine learning, and provide actionable insights for reducing churn.

1. Project Overview

Customer churn is a major business challenge in the telecom industry, often causing significant revenue loss. This project builds a predictive model that classifies whether a customer is likely to churn based on their usage patterns, service interactions, and demographics. Using real-world data from a telecom provider, we applied supervised machine learning to identify at-risk customers and understand the behavioural factors behind churn.

2. Key Insights & Behavioural Patterns

From exploratory data analysis and modelling, we discovered:

- Shorter subscription length, recent complaints, and lower usage levels are the strongest churn indicators.
- Account status flags (e.g., Status_1) were the most important predictors in the model — likely representing payment or account issues.
- Customers with fewer social interactions (low SMS, low call diversity) and low overall engagement were more prone to churn.
- Younger age groups tend to send more SMS than make calls — signalling that communication style is a relevant dimension.

These insights allow targeted intervention strategies based on usage behaviour.

3. Model Summary & Results

Two models were evaluated: Logistic Regression (baseline) and Random Forest (final model).

Metric (Churn Class)	Logistic Regression	Random Forest
Precision	51.0%	88.2%
Recall	89.9%	81.1%
F1-score	65.0%	84.5%
Accuracy (overall)	84.9%	95.3%

Final Model: Random Forest

- Achieves high precision and recall, making it ideal for real-world deployment.
- Confusion matrix analysis confirms it minimizes false positives while catching most actual churners.
- Feature importance plot enables transparency and trust in the model’s decision-making.

4. Strategic Business Recommendations

Targeted Retention Strategy

- Prioritize retention efforts for customers with:
    - Short subscriptions (<3 months)
    - Complaints in the last 30 days
    - Low recent call/SMS activity
    - Flags in account status

Operational Impact

- Focus retention resources on the high-risk segments identified by the model to maximize ROI.
- Minimize customer loss while avoiding over-offering incentives to loyal users.

5. Deployment & Next Steps

Deployment Plan

- Export the trained Random Forest model via joblib
- Deploy via a Streamlit dashboard or REST API to score customer churn weekly
- Integrate into CRM or marketing automation tools for real-time retention alerts

Future Improvements

- Add temporal features (e.g., usage trend over time)
- Incorporate textual feedback or complaint logs
- Tune the model using GridSearchCV or XGBoost for enhanced accuracy

6. Conclusion

This project demonstrates how machine learning can move beyond experimentation and into strategic impact. By identifying churn drivers and targeting the right customers at the right time, the business can reduce revenue loss and build stronger customer relationships.

Result: A highly interpretable, deployable churn model that bridges data science and business action.