
Final Recommendation

1. Project Objective Recap

The objective of this project was to build a predictive model that can accurately identify telecom customers who are likely to churn. This allows the business to proactively engage at-risk customers and reduce overall churn rates.

---

2. Summary of Approach

- The dataset contained 7,043 entries with 21 features.
- Data cleaning addressed nulls, duplicate records, and formatting issues (e.g., converting TotalCharges to float).
- Exploratory Data Analysis (EDA) was performed to understand the distribution and relationships of features with churn.
- Categorical variables were label encoded.
- Class imbalance was handled using **SMOTE (Synthetic Minority Oversampling Technique)**.
- Multiple machine learning models were tested: **Decision Tree**, **Random Forest**, and **XGBoost**.
- Evaluation metrics included accuracy, precision, recall, F1-score, and confusion matrices.

---

3. Best Performing Model

The **Random Forest classifier** delivered the highest performance based on evaluation metrics, showing strong ability to generalize and correctly classify churners.

---

4. Business Insights

- Customers on **month-to-month contracts** and those using **electronic check** as payment were more likely to churn.
- **Senior citizens** and customers with **higher monthly charges** also showed elevated churn rates.

- Services like **tech support**, **online security**, and **longer tenure** correlated with customer retention.

---

5. Recommended Actions

- **Deploy the XGBoost model** to production, integrating it with the CRM system to flag high-risk customers in real time.
- **Automate alerts and retention offers** for customers predicted to churn.
- **Monitor performance** monthly and retrain the model quarterly using new data to maintain accuracy.
- **Enrich the dataset** with additional behavioral features (e.g., call center interactions, usage patterns) to further improve model performance.
- Consider implementing **A/B testing** on retention strategies guided by model predictions to assess impact on churn rates.

---

6. Future Enhancements

- Apply **SHAP or LIME** to improve interpretability and explain model predictions to business users.
- Develop **customer segmentation** using clustering to tailor marketing and retention efforts.
- Explore **time-series churn modeling** to predict churn risk over time instead of as a binary classification.