Customer Churn Prediction (Telecom Industry)– Report

Two machine learning models were trained and evaluated to predict customer churn: Logistic Regression and Random Forest Classifier. Both models were assessed using accuracy, precision, recall, F1 score, ROC-AUC, and confusion matrices.

Logistic Regression – Performance Summary

- *Accuracy*: 0.81
- *Precision* (Churn = 1): 0.67
- *Recall* (Churn = 1): 0.52
- *F1 Score*: 0.59
- *ROC-AUC*: 0.846

Interpretation:
Logistic Regression performed strongly, achieving the highest overall accuracy and ROC-AUC among the models. Its recall of 0.52 means it successfully detected slightly more than half of all customers who churned.

The model can correctly predict customer churn 80.5% of the time. the ROC-AUC is 0.846, which means it does an excellent job separating the churners from non-churners. Non-churners were correctly identified with high recall (91%) but performed moderately on detecting churners (recall = 52%) and missed churners 48% of the time. Although logistic regression provides a solid baseline, it misses some churners (false negatives). Therefore, a more advanced model, such as Random Forest, is recommended to improve churn detection.

Random Forest Classifier – Performance Summary

- *Accuracy*: 0.77
- *Precision (Churn = 1)*: 0.59
- *Recall (Churn = 1):* 0.46
- *F1 Score*: 0.52
- *ROC-AUC*: 0.819

Interpretation:
The Random Forest model correctly predicted churn/no churn 77% of the time. Although the model performed reasonably well, its ability to identify customers who are likely to churn was limited. 54% of the customers who left, the model missed them. The model's ROC-AUC score of 0.819 indicates good discrimination between churners and non-churners.

Random Forest provided valuable feature importance rankings, revealing which variables had the strongest influence on churn. The top predictors included:

- MonthlyCharges
- Tenure

- Contract type
- Number of services
- Online security and tech support
- Payment method (electronic check)

Model Comparison & Final Selection

Metric	Logistic Regression	Random Forest
Accuracy	0.81	0.77
Precision	0.67	0.59
Recall	0.52	0.46
F1 Score	0.59	0.52
ROC-AUC	0.846	0.819

Key Insights from the Churn Analysis

1. Contract type: Customers on month-to-month contracts are far more likely to churn than those with one-year or two-year contracts.
2. MonthlyCharges: Customers paying higher MonthlyCharges tend to leave more often. Customers may seek cheaper options when they become dissatisfied with the value received.
3. Tenure: Newer customers (low tenure) are significantly more likely to churn, while long-term customers rarely leave. This indicates that the company must focus more on onboarding and early engagement.
4. Number of Services: Customers who subscribe to fewer services are more prone to churn. Offering package deals might reduce churn.
5. Internet Service: Customers using fiber optic internet churn more than DSL users. This may reflect issues related to pricing, service quality, or expectations.
6. PaymentMethod: Customers paying via electronic check consistently displayed high churn