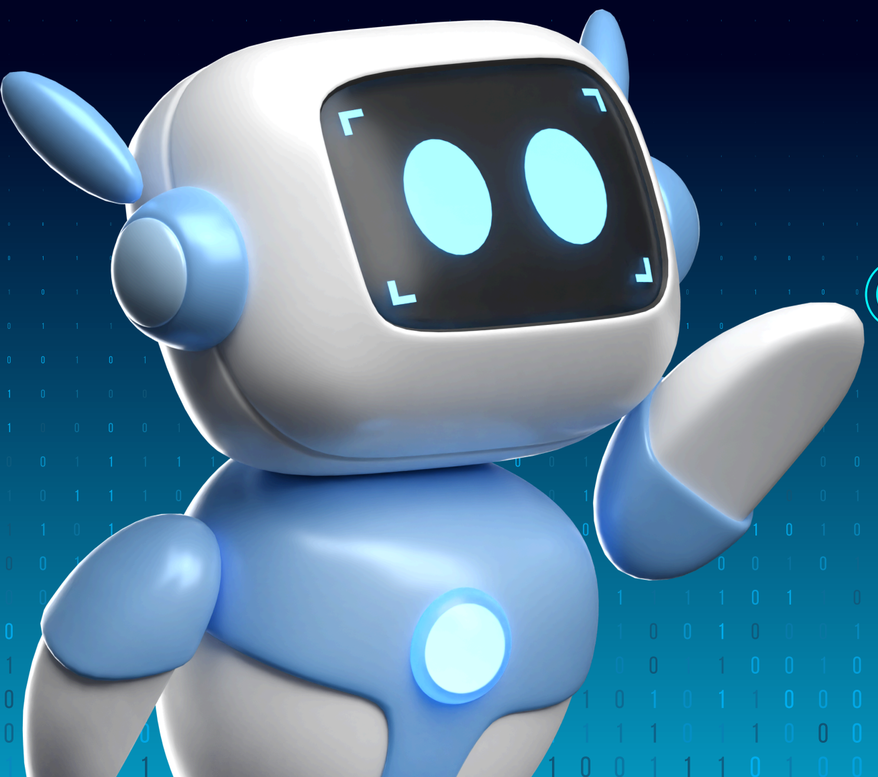
RETENTION RADAR - TELECOM CUSTOMER INSIGHTS



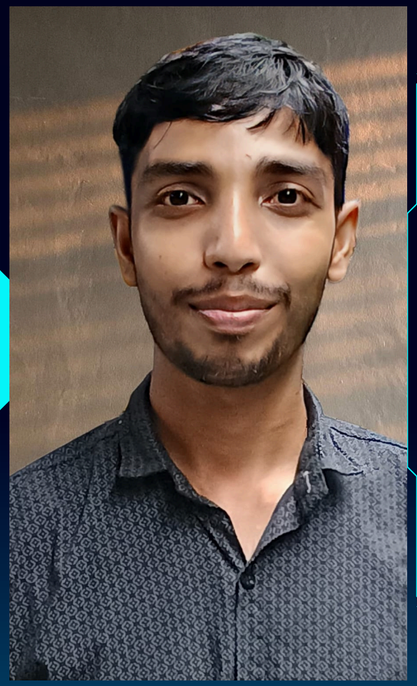
BY ANALYTICS ALLEY

TEAM MEMBERS

Name: Anas Zaki

Email: ayanzaki1947@gmail.com

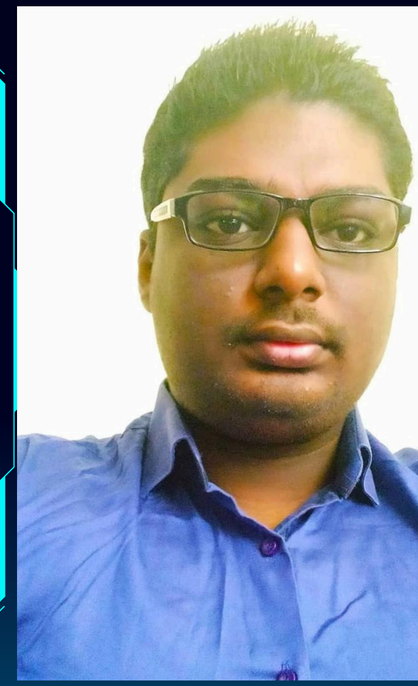
Student Code:pb01_52019



Name: Mahesh Mahato

Email: zatchbell1291514@gmail.com

Student Code: pd01_52015



Name: Ashwin Kumar

Email: ashwinkr676@gmail.com

Student Code: pds25_05004

OBJECTIVE

- This project aims to solve a data-driven problem using end-to-end data science techniques, from cleaning to model training and visualization.
- The goal was to extract insights, build a machine learning model, and present findings using an interactive Power BI dashboard.
- We focused on teamwork, using tools like Python and Power BI for real-world applications.

TOOLS & TECHNOLOGIES

- Python for data analysis and ML
- Pandas and NumPy for data cleaning and manipulation
- Seaborn and Matplotlib for visualization
- Scikit-learn for ML model training and evaluation
- Power BI for dashboarding
- GitHub for version control and collaboration

METHODOLOGY

- Started with raw data cleaning, fixing null values, and converting categorical to numerical
- Visualized the dataset using plots to find trends and correlations.
- Scaled numerical data for machine learning compatibility.
- Trained and evaluated the model using sklearn.
- Built an interactive Power BI dashboard for visualization.

TEAM CONTRIBUTIONS

- **Mahesh Mahto** handled data cleaning, encoding categorical values, scaling numerical data, and identifying relationships using correlation heatmaps
- **Ashwin Kumar** trained the machine learning model, selected key features, tuned parameters, and evaluated accuracy to ensure reliable predictions.
- **Anas Zaki** designed the Power BI dashboard with interactive visuals to highlight key insights. He also managed the GitHub repo and handled project documentation.

KEY INSIGHTS

- **Insight 1: Model 1 — Logistic Regression:** We used Logistic Regression as a baseline classification model due to its simplicity and high interpretability. It achieved an accuracy of 80% and allowed us to understand feature impact by analyzing the model's coefficients.
- **Insight 2: Model 2 — Random Forest:** We implemented Random Forest as our second model because it handles non-linear relationships effectively and reduces overfitting by averaging multiple decision trees. It achieved a high accuracy of 99%, with strong precision, recall, and F1-score values. The model also offers good interpretability through feature importance analysis.

DASHBOARD

Retention Radar - Telecom Customer Insights

5626	1495	26.57%	$64.96	$12.95M
Total Customers	Churned Customers	Churn Rate (%)	Avg MonthlyCharges	Total Charges

Gender: Female | Male

Contract: Month-to-month | One year | Two year

Write a caption.

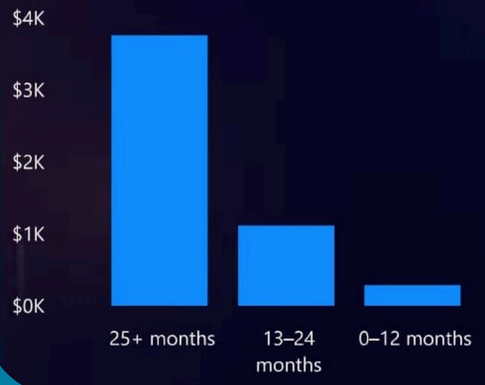
No. of customers by PaymentMethod and Churned

Total Revenue by Contract

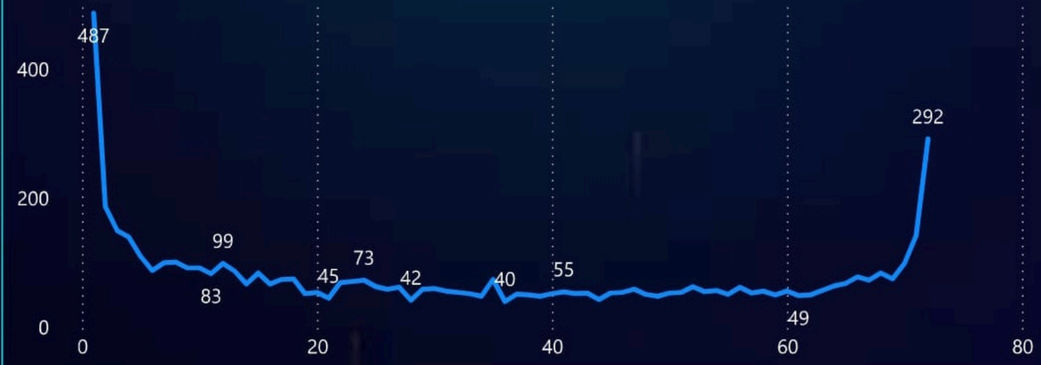
CLTV by gender and Churned

Top 20 Customers

Average CLTV by Tenure Group



No. of Customers by tenure



CHALLENGES & SOLUTIONS

- **Categorical Data:** Several columns were in text format. We used Label Encoding to convert them into numeric values suitable for model training.
- **Null Values:** The dataset had missing values, which we handled using pandas by applying imputation techniques or removing incomplete rows.
- **Dashboard Slowness:** The Power BI dashboard was slow due to heavy visuals. We optimized performance by simplifying charts and limiting slicers.

CONCLUSION & LEARNINGS

- Understood the complete data science lifecycle, from preprocessing to modeling and visualization.
- Learned hands-on usage of Python libraries, model evaluation, and dashboarding.
- Improved teamwork, version control with GitHub, and communication.
- Project reflects how data science solves real-world business problems.

FUTURE SCOP

- **Use Advanced ML Models:** Plan to improve accuracy by implementing models like Random Forest and Gradient Boosting for better performance on complex data.
- **Live Data in Power BI:** Aim to connect Power BI to live data sources, enabling real-time updates and dynamic dashboards.
- **Model Explainability:** Use tools like SHAP to explain model predictions and understand which features impact results the most.
- **Web App or API Deployment:** Deploy the model as a web app or API to make it accessible for real-time use by other users or systems.

Q&A

“Thank you for your attention. We’re happy to take your questions!”