User Analytics in the Telecommunication Industry

Presented by: Shivangi Chaudhary

Nexthikes IT Solutions — September 2025

Business Problem

- Investor evaluating TellCo (mobile provider) for acquisition.
  - Need to analyze system-generated xDR data to identify growth opportunities.
  - Focus: customers, engagement, experience, satisfaction.

Data Overview

- Source: 1 month aggregated xDR dataset (~55 features).
  - Includes user IDs, handset info, app-level DL/UL, RTT, TCP retrans, throughput.
  - Goal: derive user-level metrics and actionable insights.

Project Objectives

- Task 1: User Overview Analysis
  - Task 2: User Engagement Analysis
  - Task 3: User Experience Analysis
  - Task 4: User Satisfaction Analysis + model deployment

Task 1: User Overview — Top Handsets

- Top 10 handsets by user count (Insert chart).
  - Recommendation: Target marketing to top device models.

Task 1: Top Manufacturers

- Top 3 handset manufacturers (Insert chart).
  - Show top 5 handsets per manufacturer (Insert table).

Task 1: Application Usage

- Most popular applications: Social Media, YouTube, Netflix, Gaming.
  - Insight: Video & social drive highest data consumption.

Task 2: Engagement Metrics

- Metrics used: session count, session duration, total traffic (DL+UL).
  - Top 10 engaged users (Insert table).
  - Normalize metrics and cluster users (k=3).

Task 2: Engagement Clusters

- Clusters: Low, Medium, High engagement (Insert cluster plot).
  - Elbow method used to validate k.

Task 2: Insights & Recommendations

- Small % of users consume large share of data (80/20).
  - Recommendation: Upsell premium data plans to high-engagement users.

Task 3: Experience Metrics

- Metrics: Avg RTT (DL/UL), Avg Throughput, TCP retransmissions.
  - Aggregate per-user metrics and analyze by handset type (Insert charts).

Task 3: Experience Distribution

- Show distributions and top/bottom TCP, RTT, throughput values (Insert visuals).
  - Interpretation: Most users OK, but subset has poor QoE.

Task 3: Clustering & Insights

- Experience clusters: Good / Moderate / Poor (k=3).
  - Action: Prioritize network improvements where experience is poor.

Task 4: Satisfaction — Methodology

- Engagement score = Euclidean distance to least-engaged cluster center.
  - Experience score = Euclidean distance to worst-experience cluster center.
  - Satisfaction = average(engagement_score, experience_score).

Task 4: Top Satisfied Users

- Top 10 satisfied customers (Insert table/charts).
  - Use cases: Upsell, loyalty programs.

Task 4: Regression Model

- Linear Regression to predict satisfaction score from engagement & experience features.
  - Key drivers identified (Insert coefficients table).

Task 4: Clustering on Scores (k=2)

- Cluster users into two groups by engagement & experience scores.
  - Report cluster averages for satisfaction & experience (Insert table).

MySQL Export

- Final table exported to MySQL: telecomdb.final_task4_scores
  - Run: SELECT * FROM final_task4_scores LIMIT 10; (Insert screenshot here)

Model Tracking (MLflow)

- Tracked model runs, parameters, metrics, and artifacts with MLflow.
  - Insert MLflow UI screenshot showing the logged run.

Recommendation & Limitations

- Recommendation: BUY TellCo — focus on upsell and network improvement.
  - Limitations: only 1 month of data; network upgrade costs not modeled; market factors not considered.