Customer Churn analysis

Table of

Content

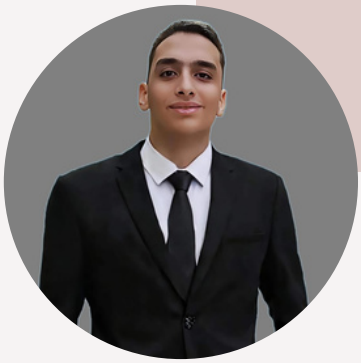
Project

Summary

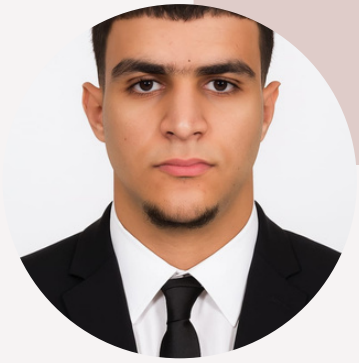
This project will develop an end-to-end data solution to address customer churn for a telecommunications company. We will process and analyze customer data to build a high-accuracy machine learning model capable of predicting which customers are likely to churn. The core deliverable is an interactive web dashboard, built with Python, that allows business stakeholders to get real-time churn predictions for any customer. This tool will empower the company to proactively identify at-risk customers and implement targeted retention strategies, thereby reducing revenue loss and improving customer loyalty.

Our

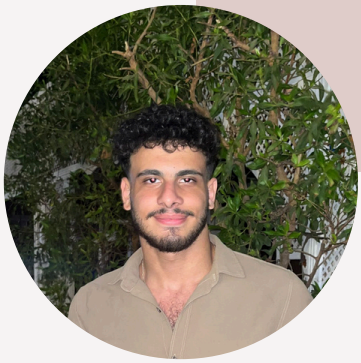
Team



Yousef ELmasry
AI Engineer



Amr Ibrahim
AI Engineer



Mohamed Waleed
Ai Engineer

Amr Elgazar
Team Leader
Ai Engineer

Project

Group Members &Roles

Our team consists of four members, each with a specialized role to ensure comprehensive project coverage and efficient execution·

Data Analyst & Preprocessing Lead**(Milestone_1)**:
[Mohamed Waleed]·
Responsibilities: Data loading, cleaning, exploratory data analysis (EDA), handling missing/duplicate data, and generating the final cleaned CSV file.

Machine Learning **(Milestone_2)**:
[Amr Ibrahim & Amr Mohamed]·
Responsibilities: training and evaluating the XGBoost predictive model, and ensuring model accuracy·

Dashboard Developer (Python/Streamlit)
(Milestone_3) [Yousef ELmasry]·
Responsibilities: Developing the interactive web dashboard using Python (Streamlit), integrating the machine learning model, and ensuring a fast, responsive user interface.·

Project

Objectives

1_To process and clean the raw Telco Churn dataset to create a reliable, analysis-ready data source.

2_To build and train a machine learning model with a target accuracy of over 80% in predicting customer churn.

3_To develop an interactive and user-friendly dashboard that provides real-time predictions and visual insights into key churn drivers.

4_To deliver a final report and presentation that clearly communicates the project's outcomes, business value, and strategic recommendations for customer retention.

Tools &Technologies

- **Programming:** Language: Python•
- **Data Analysis & ML Libraries**: Pandas, NumP y, Scikit-learn, XGBoost•
- **Dashboard & Visualization**: Streamlit, Matplotlib, Seaborn•
- **Development Environment:** VS Code, Google Colab.

Project

Timeline

Milestones 01

Data Preprocessing
DeadLine:
(15 SEP 2025)

Complete data cleaning and generate cleaned_churn_data.csv

Milestones 02

Model Training
DeadLine:
(15 OCT 2025)

Train, evaluate, and save the final predictive model

Milestones 03

Streamlit App
DeadLine:
(3 NOV 2025)

Build interactive web app by using streamlit
Complete the final written report and slide deck.
Prepare for the final project presentation.

KPIs (KeyPerformance Indicators)

We will measure the success of our project against the following customized KPIs, as per the submission guidelines:

1. D**ata Preprocessing (Python script, cleaned CSV)**
    1. 80% of missing/duplicate data correctly handled: Target: 100%. All missing values in TotalCharges will be imputed, and duplicate records will be removed.
    2. **Script efficiency:** The data cleaning script will be optimized to execute in under 15 seconds on the provided dataset.

2. **Visualization (Charts, dashboard)**

1. Dashboard load time: Target: < 3 seconds. The Streamlit dashboard will be optimized for a fast initial load.

2. of required KPIs/metrics visualized: Target: ⩾ 90%. The dashboard will visualize all key churn drivers, including contract type, tenure, and internet service, as identified by the model.

Presentation (Report, slide deck)
Report completeness: Target: 100%. The final report will include all required sections: EDA, Model Performance, Dashboard Guide, and Strategic Recommendations.