

Tasks for Learners

1. Data Cleaning & Preprocessing

- Handle missing values.
- Detect and remove duplicate records.
- Convert categorical variables into numerical form.
- Scale numerical features (only) where necessary.

2. Exploratory Data Analysis (EDA)

- Explore the distribution of Exited (churn vs non-churn).
- Find key patterns:
 - Does age affect churn?
 - Are certain geographies more likely to churn?
 - Does credit score or balance play a role?
- Visualize churn rates across different customer groups.

3. Predictive Modeling

- Split the dataset into training and test sets.
- Train at least two machine learning models (e.g., Logistic Regression, Random Forest).
- Evaluate using accuracy, precision, recall, and F1-score.

4. Insights & Recommendations

- Identify the most important factors influencing churn.
- Suggest business actions (e.g., targeted offers, loyalty programs) to improve retention.

Deliverables

- Jupyter Notebook (with data cleaning, EDA, and model building).
- Summary report highlighting findings, model performance, and recommendations.