

CAPSTONE PROJECT

PROJECT TITLE

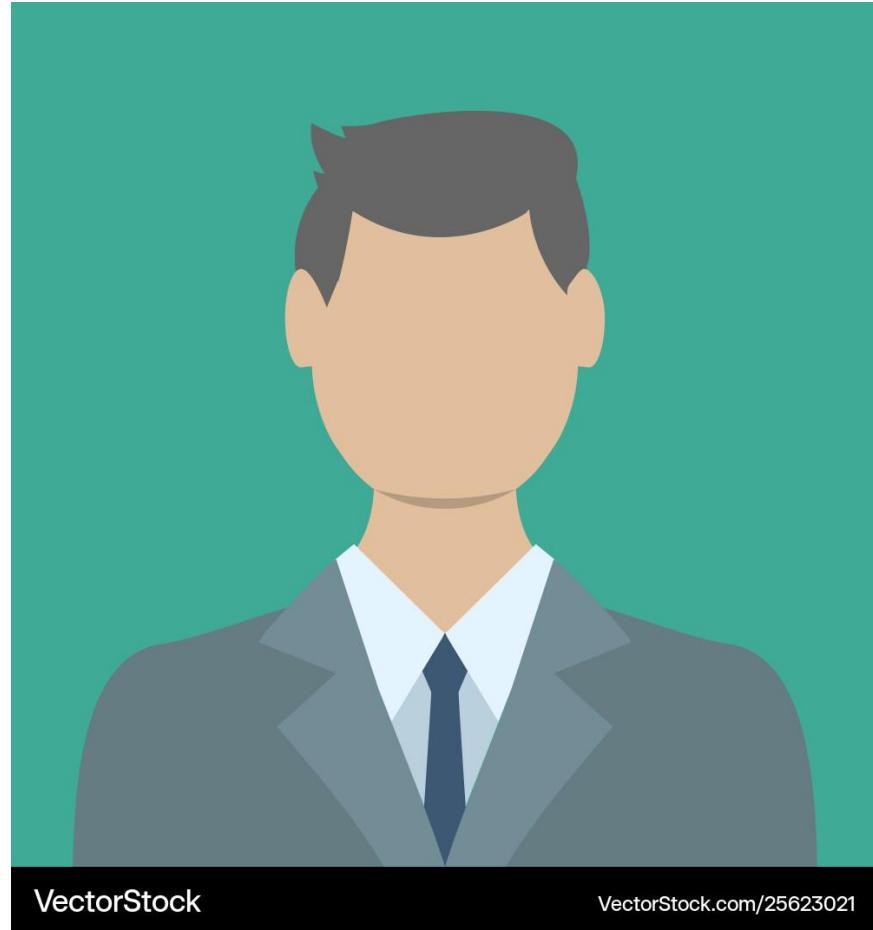
PRESENTED BY

STUDENT NAME: VUPPU ADITHYA KUMAR

COLLEGE NAME: ANURAG UNIVERSITY

DEPARTMENT: DATA SCIENCE

EMAIL ID: vuppuadithyakumar@gmail.com



OUTLINE:

- **Problem Statement** (Customer Churn Analysis)
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result (Output Image)**
- **Conclusion**
- **Future Scope**
- **References**

PROBLEM STATEMENT:

- ❑ In the telecommunications sector , customer "churn" (customers leaving for a competitor) is a primary driver of revenue loss.
- ❑ The current business model lacks a data-driven method to identify "at-risk" customers before they cancel their service.
- ❑ Without historical analysis, the company cannot pinpoint which specific factors (contract types, service issues, or pricing) contribute to high churn rates.
- ❑ The crucial challenge is to analyze 7,043 customer records to identify patterns and predict churn probability to stabilize revenue.

PROPOSED SOLUTION:

- ❑ **Integrated Analytics Framework :** The proposed system shifts the strategy from descriptive to predictive analytics using a multi-layered Power BI environment.
- ❑ **Data Collection & Strategy:** We leverage a comprehensive dataset of 7,043 customers, utilizing historical features such as tenure, contract types, and technical support interactions to build a behavioral profile.
- ❑ **Data Preprocessing:** Cleaning data via Power Query, handling missing values in "TotalCharges," and formatting data types for accurate calculation.
- ❑ **Analytical Engine:** Implementation of a Power BI dashboard using DAX measures (Total Customers, Churn Count, Churn Rate %) to quantify the problem.
- ❑ **Feature Engineering:** Through Power Query, we transform raw data into calculated groups, such as "Tenure Brackets" (0-12 months, 12-24 months), to observe how loyalty changes over time.
- ❑ **AI-Driven Forecasting:** Rather than manual observation, we implement the Key Influencers visual, which utilizes a machine learning backend to rank variables by their impact on the churn percentage.
- ❑ **Prediction:** Leveraging AI-driven visual algorithms to find the "Key Influencers" that lead to customer attrition.

SYSTEM APPROACH:

The "System Approach" section outlines the overall strategy and methodology for developing and implementing the churn analysis. Here's a structure for this section:

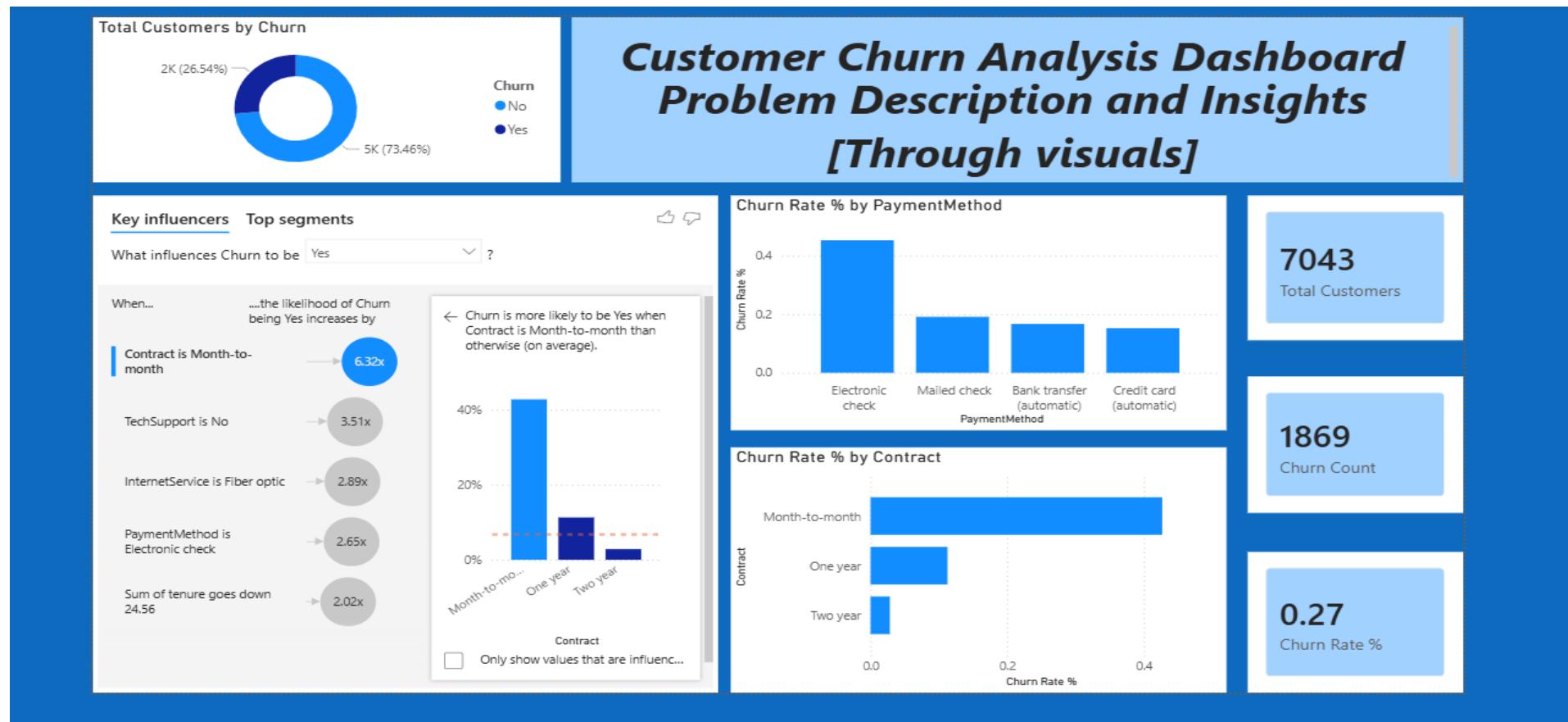
- **System requirements:**
 - **OS:** Windows 10/11
 - **Tool:** Microsoft Power BI Desktop (Latest Version).
 - **Data Source:** CSV-based Customer Churn Dataset.
- **Library required to build the model:**
 - **DAX (Data Analysis Expressions):** For complex metrics calculation.
 - **Power Query:** For ETL (Extract, Transform, Load) processes.
 - **Key Influencers AI:** For automated regression analysis.

ALGORITHM & DEPLOYMENT:

- **Algorithm Selection (Logistic Regression Logic):**
 - We chose the **Key Influencers** AI algorithm because it is specifically designed for binary classification problems (e.g., will a customer leave: Yes or No?).
- **Data Input:**
 - The model analyzes the relationship between the target (Churn) and independent variables like Contract and MonthlyCharges, calculating the "increase in likelihood" for each attribute..
- **Hyperparameter Tuning:**
 - In Power BI, we "tune" the analysis by adding "Expand By" categories, allowing the algorithm to look for patterns within specific sub-segments like Senior Citizens or Fiber Optic users.
 - **Deployment Strategy:** The final solution is deployed as an interactive .PBIX dashboard, allowing stakeholders to filter data by region or service type for localized decision-making.

RESULT:

DASHBOARD:



CONCLUSION:

- **Key Insight:** The project revealed that the "Month-to-Month" contract is the primary risk factor, increasing churn probability by over 2.5x compared to long-term contracts.
- **Effectiveness:** By identifying that "Fiber Optic" users churn more than "DSL" users, we have highlighted a potential service quality issue that was previously hidden in the raw data.
- **Final Impact:** This system provides the business with a stable supply of actionable insights, allowing for the design of targeted "Retention Campaigns" for high-risk customers, thereby securing future revenue streams.

FUTURE SCOPE:

- ❑ **Predictive Evolution:** Integrating **Azure Machine Learning** to deploy more complex models like Random Forest or XGBoost for individual-level churn probability scoring.
- ❑ **Edge Integration:** Exploring the use of **Real-Time Streaming** data to trigger automatic discount emails the moment a customer views the "Cancel Subscription" page.
- ❑ **Global Expansion:** Scaling the dashboard to compare churn patterns across different international markets or service regions.

REFERENCES:

GitHub Link: <https://github.com/adithya4001/CUSTOMER-CHUNK-ANALYSIS>

KAGGLE DATASET: <https://www.kaggle.com/datasets/blastchar/telco-customer-churn>

Thank You