Customer Churn Analysis Dashboard Using Machine Learning

[1]Mekala Devi S [2]Amirtha G [3]Mohana Sangari N [4]Yuvaharshini T
[123]UG Scholar, Department of Computer Science, Mohamed Sathak Engineering College, Ramanathapuram, Tamil Nadu
[4]Assistant Professor, Department of Computer Science, Mohamed Sathak Engineering College, Ramanathapuram, Tamil Nadu

***Abstract*—**Customer churn means losing customers when they stop using a company's products or services. A high churn rate can affect a company's growth and profit. Traditional methods of analyzing churn mostly use manual work and basic statistics, which are not very effective for large and complex data. In this project, a machine learning-based system is used to predict whether a customer is likely to leave. By studying past customer data, the models identify important factors that lead to churn. The results help organizations take actions like giving personalized offers and improving services. Overall, the system helps in better customer retention and supports data-driven decision-making.

***Keywords*—**Churn Analytics, Predictive Modeling, Risk Analysis, Model Optimization, Decision Support Systems, Streamlit Dashboard, Real-Time Insights, Data-Driven Strategy.

I. INTRODUCTION

Customer churn is a critical issue faced by businesses as it has a direct impact on their revenue and long-term growth. Traditional approaches to churn analysis often struggle to handle large volumes of complex and dynamic data effectively. To address this limitation, this project presents a machine learning-based customer churn prediction system that aims to identify customers who are likely to discontinue services. The system utilizes historical customer data and incorporates essential preprocessing techniques such as data cleaning, encoding of categorical variables, and feature scaling to enhance data quality and model performance. Multiple machine learning algorithms, including Logistic Regression, Decision Tree, and Random Forest, are implemented to develop predictive models. These models are evaluated using standard performance metrics such as accuracy, precision, recall, and F1-score to ensure reliability and effectiveness. Furthermore, a user-friendly dashboard is developed using tools like Streamlit or Power BI to visualize predictions and key insights in an intuitive manner. This enables organizations to better understand the factors contributing to customer churn and implement proactive strategies such as personalized offerings and service improvements. Overall, the proposed system enhances customer retention and supports informed, data-driven decision-making.

II. LITERATURE REVIEW

Review of existing research and methods on a topic. It helps identify gaps and improve the proposed system.

A. Traditional Methods of Customer Churn Analysis

Earlier studies on customer churn mainly used basic statistical and rule-based techniques to analyze customer behavior. Methods such as descriptive statistics and simple data analysis were applied to identify patterns in the data. Although these approaches provided some level of understanding, they were not very efficient when dealing with large datasets. In addition, they lacked the ability to accurately predict which customers were likely to leave in the future. Due to these drawbacks, businesses found it challenging to detect potential churn customers early and take effective preventive measures.

B. Machine Learning Approaches for Churn Prediction

Recent research shows that machine learning techniques are more effective in predicting customer churn with better accuracy. Algorithms like Logistic Regression, Decision Tree, and Random Forest are commonly used to study past data and identify patterns that affect customer behavior. These models improve their performance by using methods such as feature selection, cross-validation, and hyperparameter tuning. In addition, the use of visualization tools and dashboards helps to present the results in a clear and easy-to-understand way. This makes it easier for organizations to take data-driven decisions and improve their customer retention strategies.

III. EXISTING SYSTEM

The existing system for customer churn analysis primarily relies on traditional statistical methods and manual data evaluation techniques. Organizations typically use historical reports and basic analytical tools to identify churn trends; however, these approaches are often inadequate for handling large-scale and complex datasets. They lack the capability to uncover hidden patterns and do not support real-time or predictive analysis. Consequently, the existing systems are less efficient in accurately identifying

potential churn customers, limiting the ability of businesses to implement timely and effective retention strategies.

IV. PROPOSED WORK

In this project, we aim to develop a machine learning–based system to predict customer churn and help improve customer retention. We collect and preprocess customer data, then apply algorithms like Logistic Regression, Decision Tree, and Random Forest to build prediction models. The models are trained and tested to find the best one.

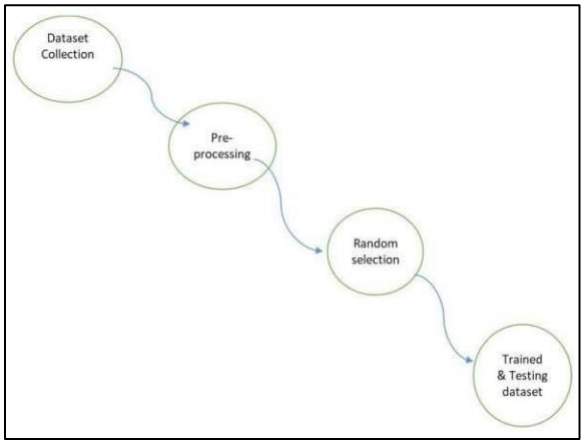
We also create an interactive dashboard using tools like Streamlit or Power BI to display the results. This helps users easily understand customer behavior and identify customers who may leave, so that necessary actions can be taken to reduce churn.

Features:

- Predicts churn accurately
- Uses multiple ML algorithms
- Interactive dashboard
- Supports real-time insights

Working:

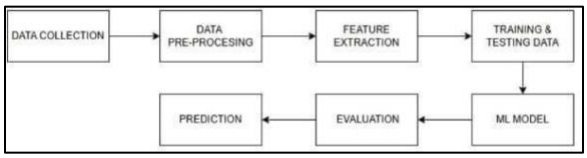
The system preprocesses data, trains and evaluates models, and selects the best one. It then predicts churn for new data and displays results through a dashboard for better decision-making.



[Image: System Architecture / Workflow Diagram — original image corrupted in source file]

V. METHODOLOGY

- Data preprocessing (cleaning, encoding, scaling)
- Feature selection and preparation
- Model building using ML algorithms (LR, DT, RF)
- Train-test split for model training and testing
- Model evaluation using accuracy, precision, recall, F1-score
- Selection of best-performing model
- Deployment using Streamlit/Power BI
- Visualization of results through dashboard

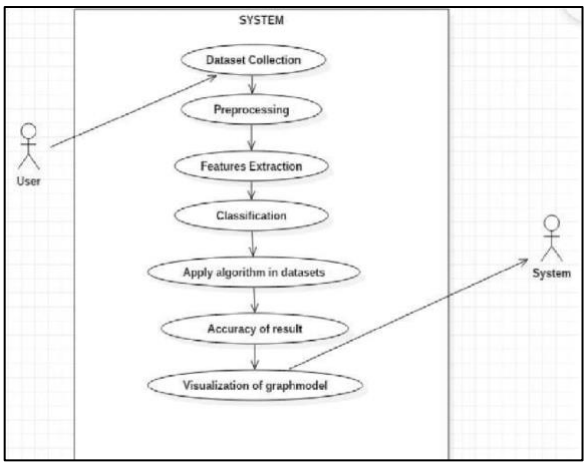


[Image: Methodology Flow / Pipeline Diagram — original image corrupted in source file]

VI. ALGORITHMS/TECHNIQUES USED

- Logistic Regression
- Decision Tree
- Random Forest
- Data Preprocessing (cleaning, encoding, scaling)
- Feature Selection
- Train-Test Split
- Cross-Validation
- Hyperparameter Tuning
- Model Evaluation (Accuracy, Precision, Recall, F1-score)

VII. RESULTS AND DISCUSSION



[Image: Results / Performance Chart — original image corrupted in source file]

The machine learning models implemented in this project successfully predict customer churn with good accuracy. Among the models used, Random Forest showed better performance compared to Logistic Regression and Decision Tree based on evaluation metrics such as accuracy, precision, recall, and F1-score. The results indicate that preprocessing techniques and feature selection significantly improve model performance. The dashboard effectively visualizes churn predictions and key insights, making it easier to understand customer behavior patterns. The system helps identify high-risk customers, allowing

organizations to take timely actions for retention. Overall, the project demonstrates that machine learning provides a reliable and efficient solution for customer churn prediction and supports data-driven decision-making.

VIII. ADVANTAGES

- Early prediction of customer churn
- Improves customer retention
- Supports data-driven decisions
- Provides accurate analysis
- Interactive dashboard visualization
- Saves time and effort

IX. APPLICATIONS

- Telecom companies for customer retention
- Banking and financial services
- E-commerce platforms
- Subscription-based services (OTT, SaaS)
- Insurance companies
- Retail businesses

X. CONCLUSION

The Customer Churn Analysis Dashboard using Machine Learning shows how data-driven methods can be used to predict customer behavior and reduce churn. By applying data preprocessing techniques and machine learning algorithms like Logistic Regression, Decision Tree, and Random Forest, the system can identify customers who are likely to stop using the service. The use of an interactive dashboard through tools like Streamlit or Power BI helps in clearly presenting the results and insights, making them easy to understand. This allows organizations to analyze patterns and take proactive steps to retain customers. Overall, the project improves customer retention strategies, supports better decision-making, and provides a flexible solution for business growth.

XI. FUTURE WORK

- Use advanced models like XGBoost and deep learning to get better accuracy.
- Add real-time prediction to get quick results.
- Deploy the system on the cloud for easy access.
- Make the dashboard more interactive and user-friendly.
- Add customer segmentation and give personalized suggestions.
- Improve the model by using larger datasets.

REFERENCES

[1] Khadka & S. Maharjan, "Customer satisfaction and loyalty," *Centria Univ. Appl. Sci.*, 2017.

[2] J. Ahn et al., "Survey on churn analysis in business domains," *IEEE Access*, 2020.

[3] J. N. Sheth & C. Uslay, "Creating enduring customer value," *J. Creating Value*, 2022.

[4] T. Dierkes et al., "Word of mouth effects on churn," *Decis. Support Syst.*, 2011.

[5] M. Alizadeh et al., "Customer churn model for banking," *IEEE Access*, 2023.

[6] N. Edwine et al., "Churn risk in telecom sector," *Math. Problems Eng.*, 2022.

[7] L. C. Cheng et al., "Customer churn behavior analysis," *J. Global Inf. Manage.*, 2019.