



“Churn Prediction in Telecom Sector Using NLP & ML”

[1] Akash Raut [2] Shivanjali Kadam [3] Mohit Sakalkale [4] Ritesh Sambhus

Guide-Prof. A. S. Thorat

Department of Computer Engineering

Amrutvahini College of Engineering, Sangamner

AMRUT EXPO 2024

Comp
B-16



1. Introduction:

• What is Churn in telecom sector?

Churn in the telecom sector refers to the rate at which customers terminate or discontinue their services with a particular telecommunications service provider. High churn rates can impact the financial health of telecom companies, making customer retention strategies crucial for their success.

• Problem Definition

-The project aims to address the challenge of customer churn in the telecom sector. The Telecom Company is identified as a primary user for particular project. The problem is to develop a churn prediction solution using NLP (Natural Language Processing) and ML (Machine Learning) to accurately identify customers at risk of churning.

- Integrating structured customer data with unstructured textual data presents challenges in data pre-processing, feature engineering, and model building. We have used Random Forest Algorithm. The goal is to create an effective model that offers interpretability and actionable insights, enabling telecom companies to proactively retain at-risk customers and improve overall customer satisfaction.

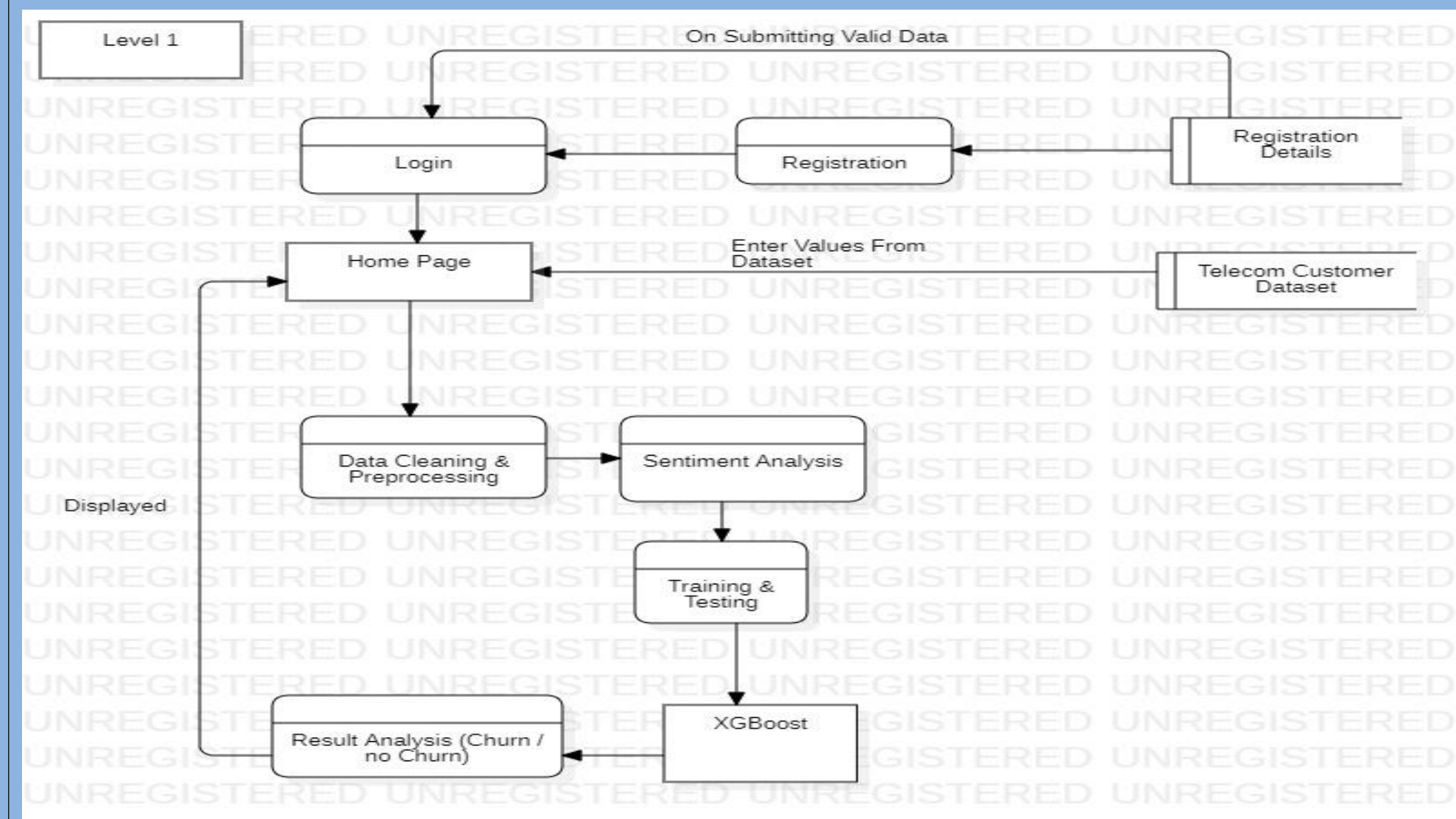
2. Need of Project:

The project is essential to tackle the prevalent issue of customer churn in the telecom sector. By employing NLP and machine learning, the goal is to develop a predictive model that accurately identifies at-risk customers. Integrating structured and unstructured data poses challenges, necessitating adept data handling and model creation. The project's significance lies in its potential to provide telecom companies with actionable insights, enabling proactive customer retention and overall satisfaction improvement. Strict adherence to ethical data practices ensures privacy and consent considerations are prioritized throughout the project.

3. Methodology:

In the methodology section of the churn prediction project, provide a detailed overview of the approach taken to achieve the project objectives. Outline the key steps involved in data collection, preprocessing, and feature engineering. Explain the utilization of NLP techniques for sentiment analysis on customer reviews. Clearly state the machine learning algorithms employed, with a focus on XGboost, and describe the model training and evaluation process. Emphasize any innovative techniques or tools used to enhance model performance.

4. Flowchart:



5. Screenshot:

Churn Report	
customerID	1231
gender	Male
SeniorCitizen	1
Dependents	Yes
Partner	Yes
PaymentMethod	Electronic check
MonthlyCharges	118
TotalCharges	306
Feedback	high maintenance
ChurnCategory	Price
Prediction Result	
Predicted Churn	
View Report	

6. Results and Discussion

The project has met expectations, demonstrating robust performance as the model yields anticipated results. Leveraging a diverse dataset comprising structured information and customer reviews, our approach proves valuable in predicting customers at risk of churn. This predictive capability empowers companies to proactively implement strategic interventions. The results effectively indicate the likelihood of customer churn, aiding companies in making informed decisions to optimize customer retention strategies.

7. Advantages:

1. Earlier identification of a customer at a risk of Churn.
2. Improved Customer Experience.
3. Resource Optimization.

8. Limitations:

1. **Data Quality:** Churn prediction models heavily rely on data quality. Inaccurate or incomplete data can lead to inaccurate predictions.
2. **Model Overfitting.**
3. **Data Privacy and Ethics:** Analyzing customer data, especially in the context of NLP, raises privacy and ethical concerns.

9. Objective:

1. To develop a Churn Prediction Model that identifies customers at risk of churning.
2. To help telecom companies for customer retention.
3. To improve the customer satisfaction by focusing on areas where improvement is required.

10. Application:

1. **Customer Retention:** Identifying potential churners allows telecom companies to take proactive measures to retain customers.
2. **Targeted Marketing:** By understanding customer preferences and behavior through ML models, telecom providers can create highly targeted marketing campaigns.
3. **Pricing Strategies:** ML models can help in determining optimal pricing strategies, such as identifying which customers are more price-sensitive.
4. **Revenue Growth:** By Applying different marketing & pricing strategies, Companies can achieve good revenue growth.

11. Conclusion:

In summary, our study highlights the vital role of churn prediction in the competitive telecom sector. Customer clusters are identified based on churn risk offer actionable insights for telecom decision-makers to refine customer retention strategies. The integration of NLP and ML not only enhances customer satisfaction but also optimizes overall business performance. Future research will explore innovative approaches, delving into the application of advanced analytics for deeper trend analysis and gaining insights into the evolving patterns of customer behavior.