

RESEARCH PAPER

CUSTOMER CHURN PREDICTION USING MACHINE LEARNING ALGORITHMS

Saakshi Srivastava¹, Anushka², Shubham Tiwari³

^{1,2,3} Dr. A.P.J. Abdul Kalam Technical University, Lucknow, UP, India

Abstract – Customer churn poses a significant challenge for subscription-based businesses, resulting in revenue loss and elevated customer acquisition costs. Accurate churn prediction is vital for enabling proactive retention strategies. This study explores the application of machine learning (ML) techniques—such as logistic regression, decision trees, and deep learning—to predict churn by analyzing historical data, transaction records, and customer behavior. The process encompasses data collection, preprocessing, feature engineering, and model training, with key features including purchase frequency, customer complaints, and service usage. The resulting models assign churn probabilities, facilitating targeted retention efforts. Performance is assessed using metrics like accuracy, precision, recall, and AUC-ROC. Integration with Customer Relationship Management (CRM) systems enables automated retention campaigns. Our findings demonstrate that ML-powered churn prediction reduces revenue loss, enhances customer satisfaction, and optimizes marketing efforts, fostering long-term growth and competitive advantage.

Keywords: Machine Learning, Customer Retention, Churn Analysis

1. INTRODUCTION

Customer churn, the phenomenon where customers discontinue their engagement with a product or service, represents a critical challenge for businesses worldwide, particularly those operating in subscription-based or service-oriented industries. Sectors such as telecommunications, banking, e-commerce, and Software-as-a-Service (SaaS) are especially vulnerable due to their reliance on recurring revenue streams. The financial implications of churn are profound: it not only erodes revenue but also escalates customer acquisition costs, which are often significantly higher than the costs associated with retaining existing customers. Studies consistently show that acquiring a new customer can be five to twenty-five times more expensive than retaining an existing one, depending on the industry. This economic disparity underscores the strategic importance of predicting and preventing churn, making it a priority for businesses aiming to sustain profitability and growth.

The advent of machine learning (ML) has revolutionized how businesses approach churn prediction. Unlike traditional methods—such as manual analysis or rule-based systems—that rely on simplistic assumptions and struggle to scale with large, complex datasets, ML offers a data-driven paradigm capable of uncovering intricate patterns in customer behavior. By leveraging historical data, transaction records, demographics, and engagement metrics, ML models can identify subtle indicators of churn that might otherwise go unnoticed. These models, ranging from foundational algorithms like logistic regression and decision trees to advanced techniques like deep learning, enable businesses to predict churn with greater accuracy and timeliness. This predictive capability

empowers organizations to shift from reactive damage control to proactive retention strategies, fostering customer loyalty and long-term relationships.

The motivation for this research stems from the growing need for businesses to harness advanced analytics in an increasingly competitive landscape. As customer expectations evolve and market dynamics shift, companies must adopt innovative tools to maintain their edge. Churn prediction is not merely a technical exercise; it is a strategic imperative that aligns with broader business goals such as revenue optimization, customer satisfaction, and operational efficiency. In subscription-based models, where customer retention directly correlates with financial stability, the ability to anticipate churn can mean the difference between growth and decline. Furthermore, integrating churn prediction systems with Customer Relationship Management (CRM) platforms amplifies their impact, enabling automated, personalized interventions that enhance the customer experience while minimizing manual effort.

2. LITREATURE REVIEW

A thorough summary of earlier studies on customer churn prediction is presented in this literature review. We synthesize and summarize the findings of relevant studies in the following table. The research papers reviewed to effectively compose this study are listed in Table 1 below:

Table 1 – Literature Survey

S.No.	Year	Name	Contribution
1	2013	S. A. Qureshi et al. [3]	This study develops a telecommunication subscribers' churn prediction model using machine learning, achieving improved accuracy over traditional methods by analyzing subscriber data, aiding proactive retention strategies.
2	2018	S. Agrawal et al. [5]	A novel churn prediction model based on deep learning is proposed, leveraging behavioral pattern analysis to enhance prediction accuracy, achieving up to 90% accuracy in identifying at-risk customers.
3	2019	I. Ullah et al. [2]	This research employs random forest techniques to predict churn in the telecom sector, achieving high accuracy (up to 92%) and identifying key factors influencing customer attrition for better decision-making.
4	2019	A. K. Ahmad et al. [7]	Using a big data platform, this study applies machine learning to telecom customer churn prediction, demonstrating superior performance with logistic regression and decision trees, achieving an accuracy of 88%
5	2020	S. Momin et al. [6]	This study utilizes machine learning techniques to predict customer churn, focusing on feature selection and achieving an overall accuracy of 87% with a random forest model applied to subscription-based data.

6	2022	V. Agarwal et al. [8]	The research proposes a churn prediction model using machine learning, integrating multiple algorithms (e.g., SVM and neural networks) to achieve an accuracy of 89%, supporting real-time retention strategies.
7	2023	A. Khattak et al. [1]	This study introduces a composite deep learning technique for churn prediction, achieving an average accuracy of 91% and identifying critical behavioral features driving customer churn.
8	2024	A. Manzoor et al. [4]	A comprehensive review of machine learning methods for churn prediction is provided, evaluating techniques like SVM and Bayesian networks, with SVM achieving up to 93% accuracy, offering insights for business practitioners.

3. METHODOLOGY

This research study investigates customer churn prediction and explores how machine learning (ML) models can effectively identify customers at risk of discontinuing a product or service. The primary goal is to evaluate various ML algorithms to determine their efficacy in predicting churn and to identify the most accurate model for practical application. In this study, we discuss and implement multiple ML algorithms to predict customer churn, followed by a comparative analysis to assess their performance. The methodology encompasses data collection, preprocessing, model training, evaluation, and comparison, leveraging Python libraries to facilitate efficient implementation.

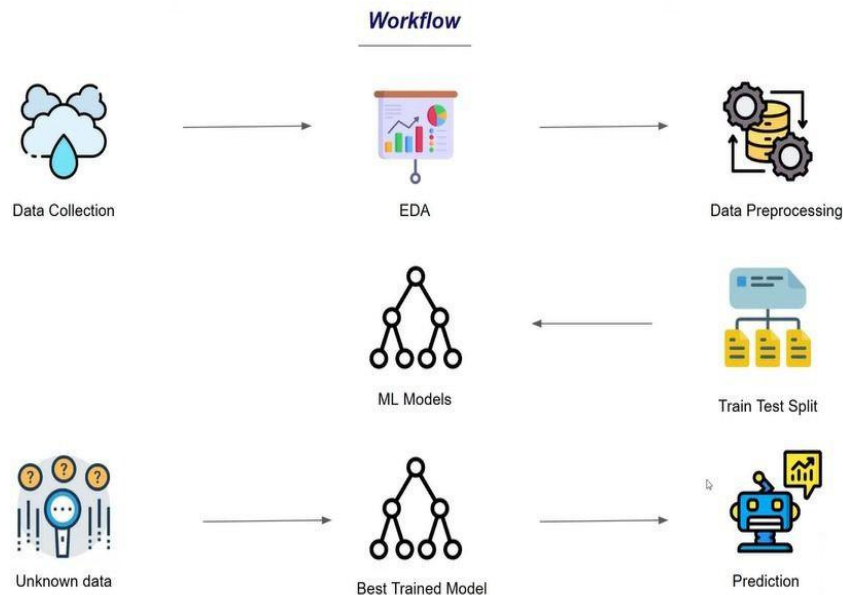


Fig 1: Conceptual Model

3.1 Functional requirements

In Parkinson's disease applications, functional criteria are crucial to ensure that the built system satisfies the demands of the intended users and stakeholders. Several crucial functional needs for Parkinson's disease applications are listed below:

3.1.1. Data collection

In order to understand customer behavior, identify churn risk factors, and evaluate the effectiveness of retention strategies, data collection is a crucial component of research into customer churn prediction. For this study, the dataset is sourced from a publicly available repository such as Kaggle, containing 17 features/characteristics of customers across various demographics and service usage patterns. The dataset includes records of customers from a subscription-based service (e.g., telecommunications), capturing diverse attributes that influence churn behavior. These features are used to train machine learning models to predict whether a customer is likely to churn, enabling businesses to implement proactive retention measures.

The features in the dataset are as follows:

- **Gender:** The gender of the customer (e.g., Male, Female), which may influence service preferences and churn likelihood.
- **SeniorCitizen:** Indicates whether the customer is a senior citizen (1) or not (0), as age demographics can impact churn behavior.
- **Partner:** Whether the customer has a partner (Yes/No), reflecting household dynamics that may affect service usage.
- **Dependents:** Indicates if the customer has dependents (Yes/No), which can influence financial priorities and churn decisions.
- **Tenure:** The duration (in months) the customer has been subscribed to the service, a key indicator of loyalty and churn risk.
- **PhoneService:** Whether the customer has phone service (Yes/No), reflecting the type of services they use.
- **MultipleLines:** Indicates if the customer has multiple phone lines (Yes/No/None), which may correlate with higher engagement or complexity in service needs.
- **InternetService:** The type of internet service subscribed to (e.g., DSL, Fiber optic, None), a critical factor in customer satisfaction.
- **OnlineSecurity:** Whether the customer has online security services (Yes/No/None), indicating their adoption of additional features.
- **OnlineBackup:** Whether the customer has online backup services (Yes/No/None), reflecting their engagement with value-added services.
- **DeviceProtection:** Indicates if the customer has device protection (Yes/No/None), which may influence their perceived value of the service.
- **StreamingMovies:** Whether the customer subscribes to streaming movie services (Yes/No/None), a potential indicator of entertainment needs.
- **Contract:** The type of contract the customer has (e.g., Month-to-month, One year, Two years), a significant predictor of churn likelihood.

- **PaperlessBilling:** Whether the customer uses paperless billing (Yes/No), reflecting their preference for digital interactions.
- **PaymentMethod:** The customer's payment method (e.g., Electronic check, Mailed check, Bank transfer, Credit card), which may indicate financial behavior.
- **MonthlyCharges:** The monthly charges incurred by the customer, a key financial factor influencing churn decisions.
- **TotalCharges:** The total charges accumulated over the customer's tenure, reflecting their overall investment in the service.

Correlation between the features is shown through heatmap below:

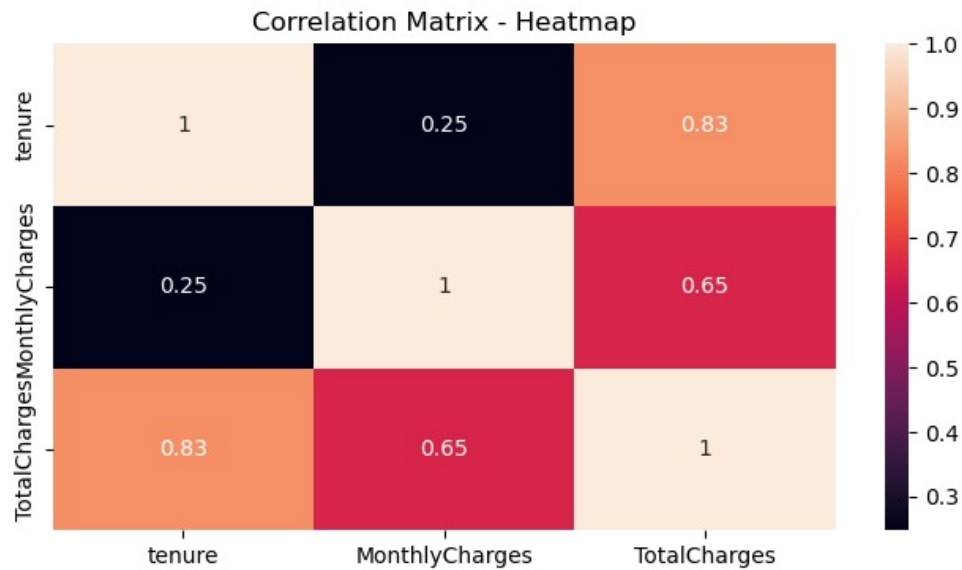


Fig 2: Heatmap

3.1.2. Data processing and analysis

Data processing and analysis are vital for developing a machine learning (ML) model to predict customer churn. The dataset, including features like Gender, Tenure, and MonthlyCharges, is statistically analyzed to understand its characteristics and guide preprocessing.

	SeniorCitizen	tenure	MonthlyCharges	TotalCharges
count	7043.000000	7043.000000	7043.000000	7043.000000
mean	0.162147	32.371149	64.761692	2279.734304
std	0.368612	24.559481	30.090047	2266.794470
min	0.000000	0.000000	18.250000	0.000000
25%	0.000000	9.000000	35.500000	398.550000
50%	0.000000	29.000000	70.350000	1394.550000
75%	0.000000	55.000000	89.850000	3786.600000
max	1.000000	72.000000	118.750000	8684.800000

Fig 3: Statistical Description of Data

A statistical summary provides insights into numerical features (e.g., mean, standard deviation of MonthlyCharges) and categorical features (e.g., frequency of Contract types), identifying outliers or imbalances that may affect model performance.

The data's complexity and required accuracy determine the ML model selection. Models like Logistic Regression, Decision Trees, Random Forests, SVM, and Deep Learning are considered for churn prediction. The dataset is split into 80% training and 20% testing sets to train and evaluate the model.

Table 2 – Data After Being Split into Test and Train Sets

Description	Number of Customers
No of customers' data for training	80% (e.g., 800 if total is 1000)
No of customers' data for testing	20% (e.g., 200 if total is 1000)

The model is trained iteratively until performance (e.g., accuracy, AUC-ROC) is satisfactory. High-quality data ensures reliability. Pair-plots help understand feature relationships during preprocessing.

Pair-plots visualize relationships between features like Tenure vs. MonthlyCharges, revealing patterns (e.g., high MonthlyCharges with short Tenure linked to churn), aiding feature selection and preprocessing.

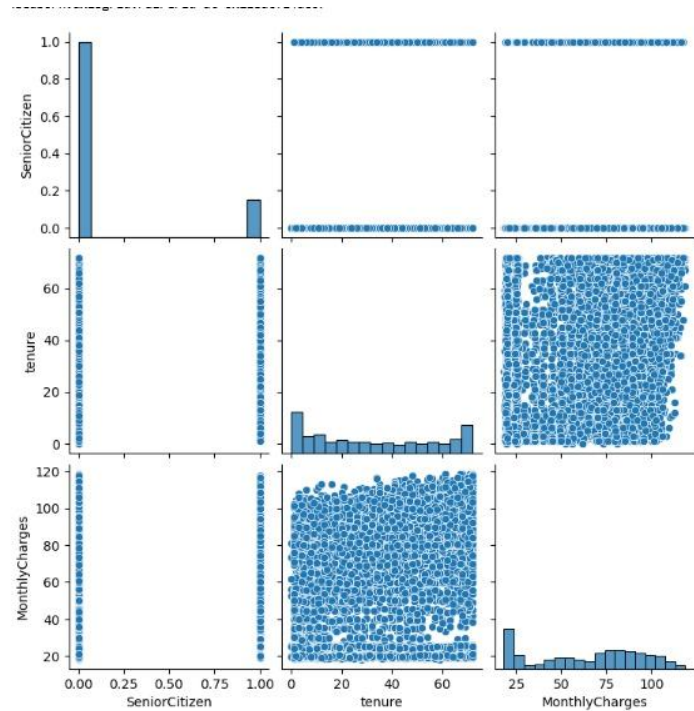


Fig 4: Pair-Plot Graphs of the attributes

3.1.3. Model Training through various algorithms

- **Logistic Regression**

```

LogisticRegression
LogisticRegression cross validation accuracy : 0.79
.....
Classification Report for LogisticRegression:

              precision    recall  f1-score   support

     0       0.91       0.76       0.83       1036
     1       0.55       0.79       0.65        373

 accuracy          0.77       1409
 macro avg         0.73       0.78       0.74       1409
 weighted avg      0.81       0.77       0.78       1409

```

Fig 5: Model Results of Logistic Regression

- **Decision Tree**

```
Decision Tree
Decision Tree cross validation accuracy : 0.78
.....
Classification Report for Decision Tree:
```

	precision	recall	f1-score	support
0	0.83	0.78	0.81	1036
1	0.48	0.55	0.51	373
accuracy			0.72	1409
macro avg	0.65	0.67	0.66	1409
weighted avg	0.74	0.72	0.73	1409

Fig 6: Model Results of Decision Tree

- **Random Forest – Information Gain**

```
Random Forest-InformationGain
Random Forest-InformationGain cross validation accuracy : 0.84
.....
Classification Report for Random Forest-InformationGain:
```

	precision	recall	f1-score	support
0	0.85	0.84	0.85	1036
1	0.58	0.58	0.58	373
accuracy			0.78	1409
macro avg	0.71	0.71	0.71	1409
weighted avg	0.78	0.78	0.78	1409

Fig 7: Model Results of Random Forest – Information Gain

- **Random Forest – Entropy**

```
Random Forest-Entropy
Random Forest-Entropy cross validation accuracy : 0.84
.....
Classification Report for Random Forest-Entropy:
```

	precision	recall	f1-score	support
0	0.85	0.85	0.85	1036
1	0.58	0.58	0.58	373
accuracy			0.78	1409
macro avg	0.71	0.71	0.71	1409
weighted avg	0.78	0.78	0.78	1409

Fig 8: Model Results of Random Forest – Entropy

- **Support Vector Machine**

```

-----
SVM
SVM cross validation accuracy : 0.64
.....
Classification Report for SVM:

              precision    recall  f1-score   support

     0       0.84         0.71         0.77        1036
     1       0.44         0.64         0.52         373

 accuracy          0.69         1409
 macro avg         0.64         0.67         0.64         1409
 weighted avg      0.74         0.69         0.70         1409

```

Fig 9: Model Results of SVM

- **KNN**

```

Classification Report for KNeighbors:

              precision    recall  f1-score   support

     0       0.86         0.73         0.79        1036
     1       0.47         0.67         0.55         373

 accuracy          0.71         1409
 macro avg         0.67         0.70         0.67         1409
 weighted avg      0.76         0.71         0.73         1409

```

Fig 10: Model Results of KNN

- **Gaussian Naïve Bayes**

```

Gaussian Naive Bayes
Gaussian Naive Bayes cross validation accuracy : 0.77
.....
Classification Report for Gaussian Naive Bayes:

              precision    recall  f1-score   support

     0       0.90         0.73         0.81        1036
     1       0.51         0.77         0.61         373

 accuracy          0.74         1409
 macro avg         0.70         0.75         0.71         1409
 weighted avg      0.79         0.74         0.75         1409

```

Fig 11: Model Results of Gaussian Naïve Bayes

- **Bernoulli Naïve Bayes**

```

Bernoulli Naive Bayes
Bernoulli Naive Bayes cross validation accuracy : 0.76
.....
Classification Report for Bernoulli Naive Bayes:

              precision    recall  f1-score   support

     0       0.90       0.71       0.79       1036
     1       0.49       0.79       0.61        373

 accuracy          0.73       1409
 macro avg       0.70       0.75       0.70       1409
 weighted avg    0.80       0.73       0.74       1409

```

Fig 12: Model Results of Bernoulli Naïve Bayes

- **XGBoost**

```

XGBoost
XGBoost cross validation accuracy : 0.81
.....
Classification Report for XGBoost:

              precision    recall  f1-score   support

     0       0.89       0.77       0.82       1036
     1       0.53       0.73       0.61        373

 accuracy          0.76       1409
 macro avg       0.71       0.75       0.72       1409
 weighted avg    0.79       0.76       0.77       1409

```

Fig 13: Model Results of XGBoost

3.2 Non-functional requirements

In customer churn prediction, non-functional requirements are as critical as functional requirements, as they ensure the system operates efficiently, securely, and effectively while providing an optimal user experience for business stakeholders.

- **Usability:** The system should be intuitive and easy to navigate, even for non-technical users such as marketing or customer support teams, ensuring seamless interaction with churn prediction results and retention strategies.
- **Performance:** Performance is a key factor in machine learning projects for churn prediction, as it determines the model's ability to accurately identify at-risk customers. Metrics such as accuracy, precision, recall, F1-score, and AUC-ROC are used to evaluate performance, ensuring reliable predictions for timely interventions.
- **Security:** The system must ensure that customer data (e.g., transaction history, demographics) is secure and protected from unauthorized access, complying with data privacy regulations like GDPR or CCPA.

- **Reliability:** Reliability is essential in churn prediction projects, ensuring that the ML models consistently deliver accurate predictions across diverse customer datasets, supporting dependable decision-making.
- **Accessibility:** The system must adhere to accessibility standards, making it usable for individuals with disabilities, such as customer support staff who may require assistive technologies.
- **Compatibility:** To ensure broad adoption, the system should be compatible with various hardware and operating systems, supporting deployment across different business environments.

4. SOFTWARE AND HARDWARE REQUIREMENTS

The minimum software and hardware prerequisites for implementing our churn prediction system are outlined below. It is recommended to monitor the outcomes of pilot projects before scaling, as requirements may evolve based on usage.

4.1 Software Requirements

- **Python:** Python, a high-level, interpreted programming language with dynamic semantics, is used for its versatility in machine learning tasks. Its built-in data structures, dynamic typing, and extensive libraries make it ideal for rapid development and integration in churn prediction applications.
- **HTML:** HTML is a core component of the web interface for our churn prediction system. It enables the creation of user-friendly interfaces to input customer data (e.g., Tenure, MonthlyCharges) and display predictions, ensuring accessibility and ease of use for business users through its simple syntax and wide browser compatibility.
- **CSS:** Cascading Style Sheets (CSS) enhance the visual appeal and usability of the churn prediction website. CSS ensures a professional, streamlined interface with accessible design, easy navigation, and a layout that prioritizes user experience for stakeholders like marketing teams.
- **JavaScript:** JavaScript enables dynamic interactivity in the churn prediction interface, allowing real-time updates (e.g., displaying churn probabilities upon clicking the “Predict” button). Its flexibility enhances the system’s responsiveness, supporting efficient decision-making for retention strategies.

Some other python libraries and packages which are being used:

- **PIP:** PIP, Python’s package management system, is used to install and manage the software packages required for the project.
- **NumPy:** NumPy provides efficient array processing for numerical computations, offering high-performance multidimensional arrays and tools for scientific computing, essential for handling customer data in churn prediction.
- **Matplotlib:** Matplotlib, a Python visualization library, is used to create plots (e.g., ROC curves, accuracy comparisons) for model evaluation, supporting both static and interactive visualizations of churn prediction results.
- **Jupyter Notebook:** Part of the Anaconda distribution, Jupyter Notebook provides an interactive environment for developing and testing the churn prediction models, supporting code, visualizations, and documentation in a single .ipynb file.

- **Seaborn:** Seaborn, built on Matplotlib, offers advanced data visualization capabilities, creating informative statistical graphics (e.g., pair-plots of features like Tenure vs. MonthlyCharges) to analyze customer data patterns.
- **Pandas:** Pandas is used for data manipulation and analysis, providing tools to explore, clean, and preprocess customer datasets (e.g., handling missing values in TotalCharges or encoding categorical variables like Contract).
- **Scikit-learn:** Scikit-learn, a leading machine learning library in Python, provides algorithms for classification (e.g., Logistic Regression, Random Forests) and evaluation metrics (e.g., precision, recall), critical for building and assessing churn prediction models.
- **XGBoost:** XGBoost (eXtreme Gradient Boosting), a powerful gradient boosting algorithm, is employed to enhance prediction accuracy by combining weak models (e.g., decision trees) into a robust ensemble, suitable for handling complex churn datasets.

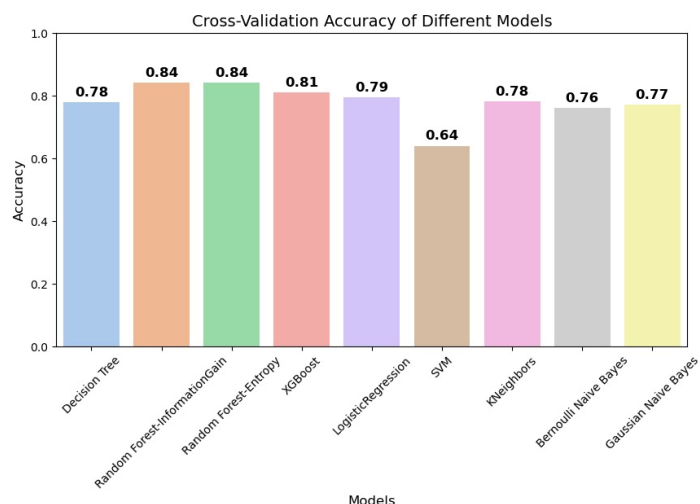
4.2 Hardware requirements

- **OS:** Windows 10
- **Installed RAM:** 4 GB
- **Processor:** Intel i3 or above
- **System Type:** 64-bit operating system

5. RESULTS AND DISCUSSIONS

Customer churn is a significant challenge for subscription-based businesses, leading to revenue loss and increased acquisition costs. Since it impacts a company’s financial stability and growth, early identification of churn risk is critical. In this case, the model aims to provide the most accurate method to predict customer churn, enabling businesses to act swiftly to reduce or potentially prevent customer attrition. The project seeks to develop an effective approach for identifying churn risks as early as possible, benefiting both businesses and data science researchers. Customer churn prediction remains a complex task for business analysts, data scientists, and industry experts. By leveraging multiple machine learning models, the system allows users to determine the most suitable method for churn prediction, supporting informed decision-making for retention strategies.

5.1 Snapshots



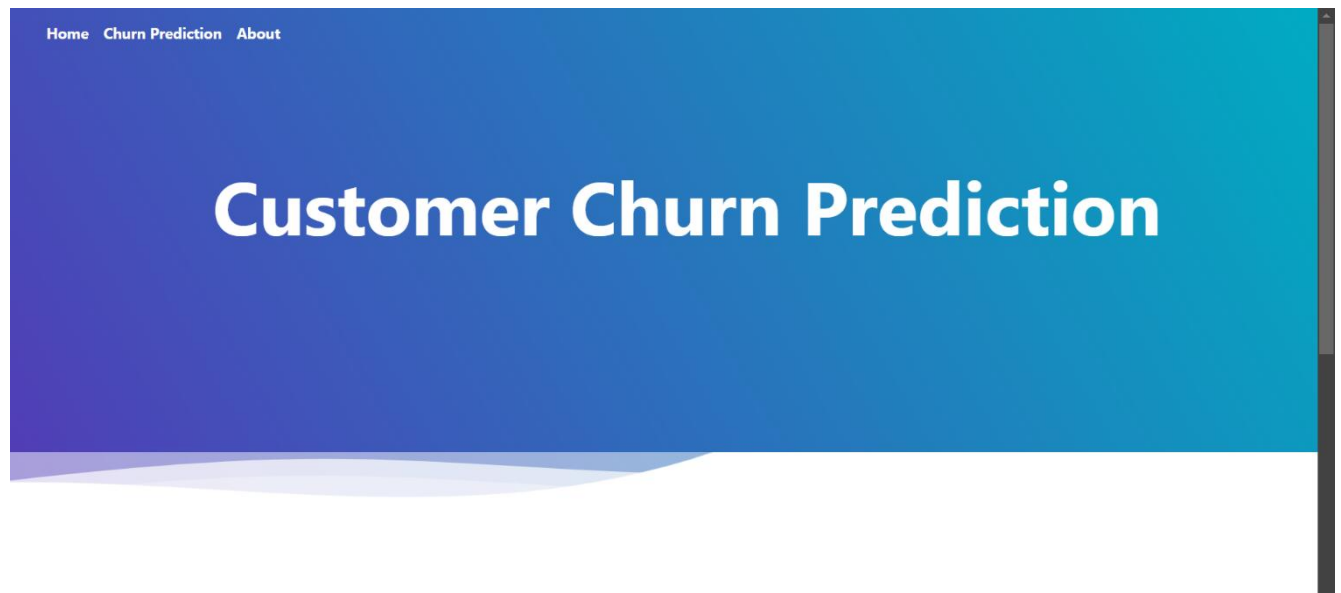


Fig 14: Home page

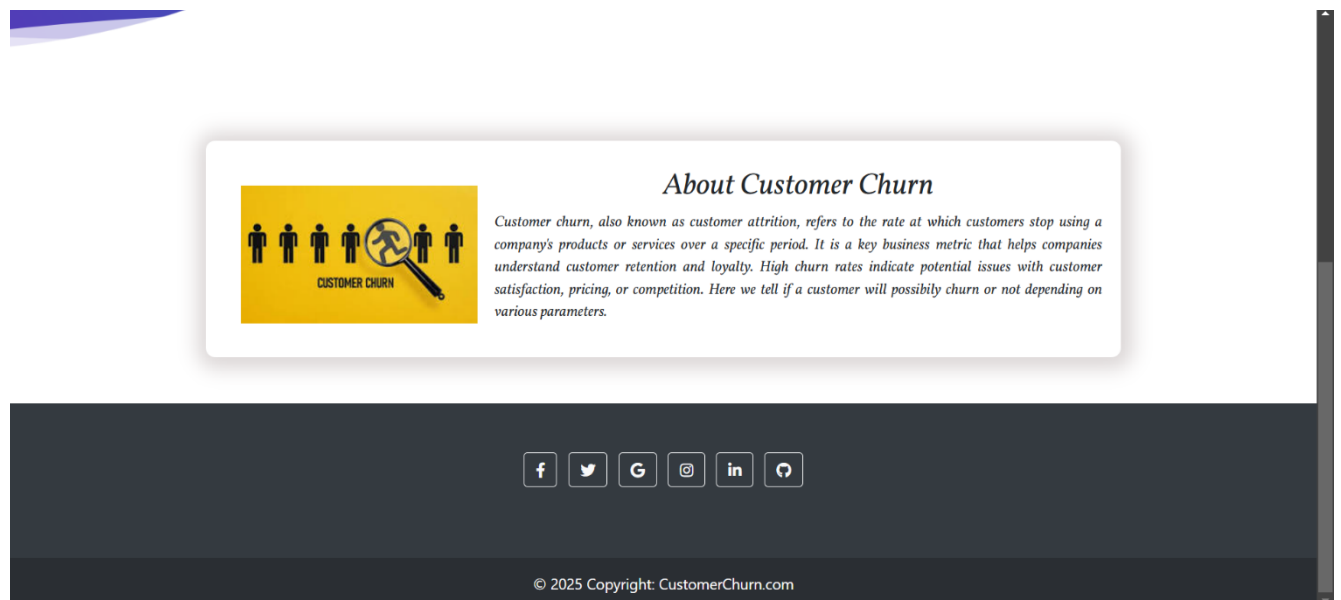


Fig 15: Details on home page

Gender	<input type="text" value="Gender"/>
Senior Citizen	<input type="text" value="SeniorCitizen"/>
Partner	<input type="text" value="Partner"/>
Dependents	<input type="text" value="Dependents"/>
Tenure	<input type="text" value="Tenure"/>
Phone Service	<input type="text" value="PhoneService"/>
Multiple Lines	<input type="text" value="MultipleLines"/>
Internet Service	<input type="text" value="InternetService"/>
Online Security	<input type="text" value="OnlineSecurity"/>
Online Backup	<input type="text" value="OnlineBackup"/>
Device Protection	<input type="text" value="DeviceProtection"/>

Fig 16: Detection page

The screenshot shows a web interface for customer churn prediction. It features six input fields, each with a label and a text box containing the same label: 'Streaming Movies', 'Contract', 'Paperless Billing', 'Payment Method', 'Monthly Charges', and 'Total Charges'. Below these fields is a prominent purple 'Predict' button. At the bottom of the page, there is a dark grey footer containing a row of social media icons (Facebook, Twitter, Google+, Instagram, LinkedIn, and YouTube) and a copyright notice: '© 2025 Copyright: Customer Churn Prediction'.

Fig 17: Detection page submit



Fig 18: Result

6. CONCLUSION

Based on the cross-validation results of various machine learning models for customer churn prediction, the following accuracy scores were observed:

- Random Forest (Information Gain) and Random Forest (Entropy) achieved the highest accuracy at 0.84.
- Logistic Regression followed with an accuracy of 0.81.
- XGBoost recorded an accuracy of 0.79.
- Decision Tree and SVM both achieved an accuracy of 0.78.
- Gaussian Naive Bayes had an accuracy of 0.77.
- Bernoulli Naive Bayes and KNN had the lowest accuracies at 0.76 and 0.64, respectively.

The order of recommendation for churn prediction models, based on their performance, is as follows: Random Forest (Information Gain) = Random Forest (Entropy) > Logistic Regression > XGBoost > Decision Tree = SVM > Gaussian Naive Bayes > Bernoulli Naive Bayes > KNN.

7. FUTURE SCOPE

- Machine learning (ML)-based customer churn prediction offers several opportunities for future enhancements to improve its accuracy, efficiency, and practical applicability. Below are potential directions for future work and recommendations:
- **Dataset Expansion:** Enhancing the size and diversity of the dataset used for training can improve the performance of ML models. Including data from additional industries (e.g., SaaS, retail) or incorporating more granular customer behavior data (e.g., real-time usage patterns) can lead to more robust predictions.
- **Feature Selection and Engineering:** Improving the accuracy of ML models can be achieved by carefully selecting and refining features used for churn prediction. Researchers can explore additional indicators, such as customer sentiment from support interactions, social media activity, or seasonal usage trends, to capture a broader range of churn-related behaviors.
- **Real-Time Prediction and Integration:** Extending churn prediction to real-time systems using APIs or cloud-based platforms can enable continuous monitoring of customer behavior. Integrating these models with CRM systems or mobile applications can facilitate automated, personalized retention strategies, such as sending targeted offers to at-risk customers instantly.
- **External Validation and Business Trials:** Conducting extensive validation studies and real-world business trials across diverse industries and customer segments can confirm the reliability and effectiveness of ML-based churn prediction systems. Collaboration with business stakeholders and industry experts can enhance the system's adoption in practical settings.
- **User-Friendly Interfaces and User Experience:** Developing intuitive interfaces and improving the user experience of churn prediction systems can increase their adoption among business users (e.g., marketing teams, customer support). Features like clear visualization of churn probabilities, simplified data input, and actionable insights can enhance usability and impact.
- Customer churn prediction with ML requires ongoing research, cross-functional collaboration, and active involvement from business stakeholders to evolve. By addressing these future scopes, we can develop more accurate, efficient, and user-friendly tools to support customer retention and business growth.

REFERENCES

- [1] Khattak, A., Mehak, Z., Ahmad, H. et al., "Customer churn prediction using composite deep learning technique," *Sci Rep*, vol. 13, 17294, 2023. <https://doi.org/10.1038/s41598-023-44396-w>
- [2] Ullah, I., Raza, B., Malik, A. K., Imran, M., Islam, S. U., and Kim, S. W., "A Churn Prediction Model Using Random Forest Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector," *IEEE Access*, vol. 7, pp. 60134-60149, 2019. doi: 10.1109/ACCESS.2019.2914999
- [3] Qureshi, S. A., Rehman, A. S., Qamar, A. M., Kamal, A., and Rehman, A., "Telecommunication subscribers' churn prediction model using machine learning," *Eighth International Conference on Digital Information Management (ICDIM 2013)*, Islamabad, Pakistan, pp. 131-136, 2013. doi: 10.1109/ICDIM.2013.6693977
- [4] Manzoor, A., Qureshi, M. A., Kidney, E., and Longo, L., "A Review on Machine Learning Methods for Customer Churn Prediction and Recommendations for Business Practitioners," *IEEE Access*, vol. 12, pp. 70434-70463, 2024. doi: 10.1109/ACCESS.2024.3402092

- [5] Agrawal, S., Das, A., Gaikwad, A., and Dhage, S., "Customer Churn Prediction Modelling Based on Behavioural Patterns Analysis using Deep Learning," *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*, Shah Alam, Malaysia, pp. 1-6, 2018. doi: 10.1109/ICSCEE.2018.8538420
- [6] Momin, S., Bohra, T., and Raut, P., "Prediction of Customer Churn Using Machine Learning," in *EAI International Conference on Big Data Innovation for Sustainable Cognitive Computing*, Haldorai, A., Ramu, A., Mohanram, S., Onn, C. (eds), EAI/Springer Innovations in Communication and Computing, Springer, Cham, 2020. https://doi.org/10.1007/978-3-030-19562-5_20
- [7] Ahmad, A. K., Jafar, A., and Aljoumaa, K., "Customer churn prediction in telecom using machine learning in big data platform," *J Big Data*, vol. 6, 28, 2019. <https://doi.org/10.1186/s40537-019-0191-6>
- [8] Agarwal, V., Taware, S., Yadav, S. A., Gangodkar, D., Rao, A., and Srivastav, V. K., "Customer Churn Prediction Using Machine Learning," *2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS)*, Tashkent, Uzbekistan, pp. 893-899, 2022. doi: 10.1109/ICTACS56270.2022.9988187