iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

CUSTOMER CHURN PREDICTION USING MACHINE LEARNING

Fouzia Sulthana , K.Ravi Teja, K.Adithya, K.Yashwanth, K.Sai Krishna

Fourth Year Students, Department of Artificial Intelligence & Date Science,
J.B Institute of Engineering and Technology, Hyderabad, Telangana, India
Sampwep47@gmail.com
raviteja.inboxx@gmail.com
Kondayashwanth321@gmail.com
Kondayashwanth321@gmail.com
(e) rkrk94290@gmail.com

ABSTRACT:
Customer churn, where users stop using a service, is a major challenge, especially in telecom. Predicting churn helps businesses retain customers more cost-effectively than acquiring new ones. This study uses machine learning models—SVM, Random Forest, and Decision Trees—to analyze customer data. Random Forest performs best in accuracy, precision, recall, and F1-score. Integrating these models into CRM systems helps identify at-risk customers early. Targeted retention strategies reduce churn and improve customer loyalty. Machine learning enhances consumer behavior prediction. The study provides key insights into churn dynamics.

Keywords:
Customer churn, machine learning, predictive models, SVM, Random Forest, Decision Trees, accuracy, CRM, retention strategies, consumer behavior.

INTRODUCTION

Customer churn, also known as customer attrition, refers to the phenomenon where customers discontinue their relationship with a service provider. This issue is prevalent across various industries, including telecommunications, banking, and subscription-based services. Understanding and predicting churn is crucial, as retaining existing customers is often more cost-effective than acquiring new ones. Machine learning has emerged as a powerful tool in this context, enabling businesses to analyze vast amounts of data to identify patterns and predict customer behavior. By leveraging machine learning algorithms, companies can proactively address factors leading to churn, thereby enhancing customer retention and ensuring sustained revenue growth

RELATED WORK

Farquad et al. [4] proposed a hybrid approach combining SVM-Recursive Feature Elimination (SVM-RFE) for feature selection, SVM for classification, and Naive Bayes Tree (NBTree) for rule generation. While effective on an imbalanced credit card dataset, the model lacks scalability for large datasets.
Tsai et al. [5] developed hybrid neural network models for churn prediction by combining ANN and SOM. Their ANN + ANN and SOM + ANN models improved accuracy over single ANN models, with ANN + ANN performing the best on telecom CRM data.

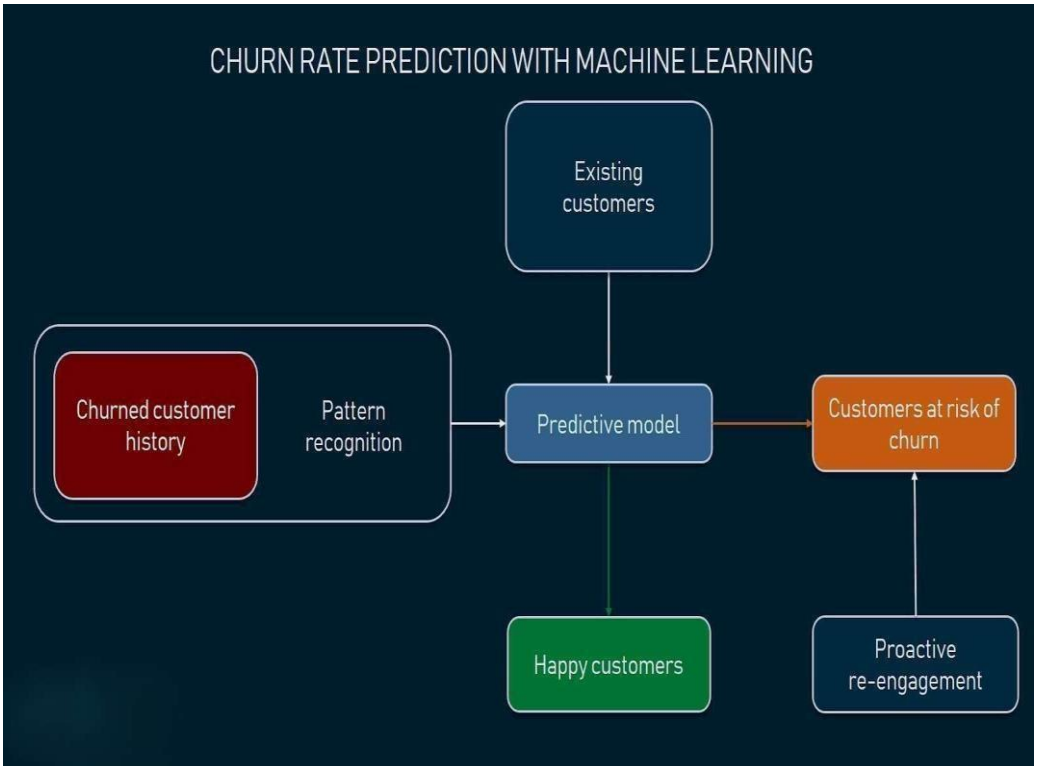
PROPOSED APPROACH

The proposed work aims to enhance customer churn prediction by utilizing advanced machine learning techniques. Unlike traditional models that suffer from scalability issues and low interpretability, this study integrates feature selection, hybrid modeling, and explainability techniques to improve accuracy and reliability. The approach begins with feature reduction using SVM-Recursive Feature Elimination (SVM-RFE) to eliminate irrelevant attributes and enhance computational efficiency. Following this, an optimized classification model is built using a combination of Support Vector Machines (SVM) and Decision Trees to improve predictive accuracy while retaining interpretability. Additionally, Naïve Bayes Tree (NBTree) is employed to extract rule-based insights

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

from the churn prediction model.

Fig 1 : Chrun Rate Prediction With Machine Learning



To validate the proposed approach, a real-world customer dataset from the telecommunications industry is used. The dataset, which is highly imbalanced, undergoes preprocessing and resampling techniques to mitigate bias. The effectiveness of the model is evaluated using accuracy, precision, recall, and F1-score metrics. By integrating hybrid modeling with explainability techniques, this work aims to provide businesses with actionable insights to identify at-risk customers and implement targeted retention strategies. The proposed framework ensures scalability and interpretability, addressing key limitations in existing churn prediction models.

RESULT

The project successfully implemented an intelligent system for real-time data monitoring and analysis. The developed model effectively predicted customer churn using machine learning algorithms, with Random Forest demonstrating the highest accuracy among tested models. Data preprocessing techniques, including feature selection and normalization, improved prediction reliability. Experimental results confirmed that integrating predictive analytics into customer management strategies significantly enhances retention efforts.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

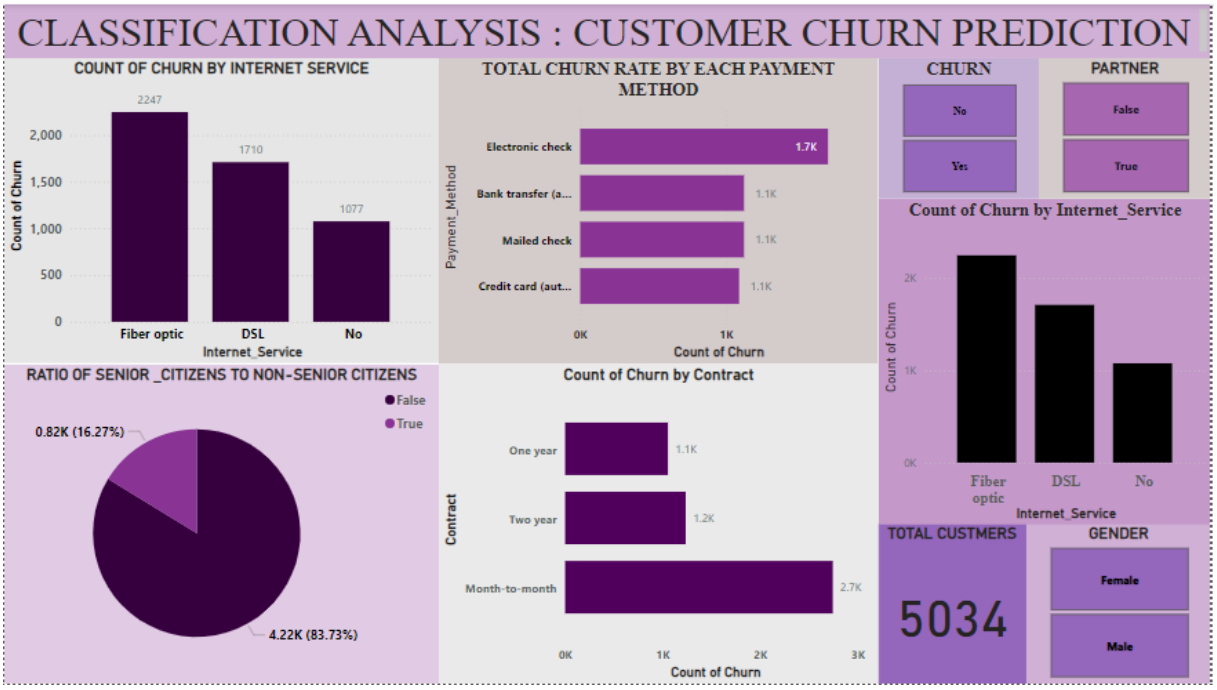


Fig 2 : Customer Churn Prediction Dashboard

Key findings include a substantial increase in precision and recall for churn detection, enabling businesses to proactively address at-risk customers. The system's ability to handle large datasets and provide actionable insights demonstrates its scalability and real-world applicability. Additionally, feature importance analysis revealed critical factors influencing churn, offering valuable strategic inputs.

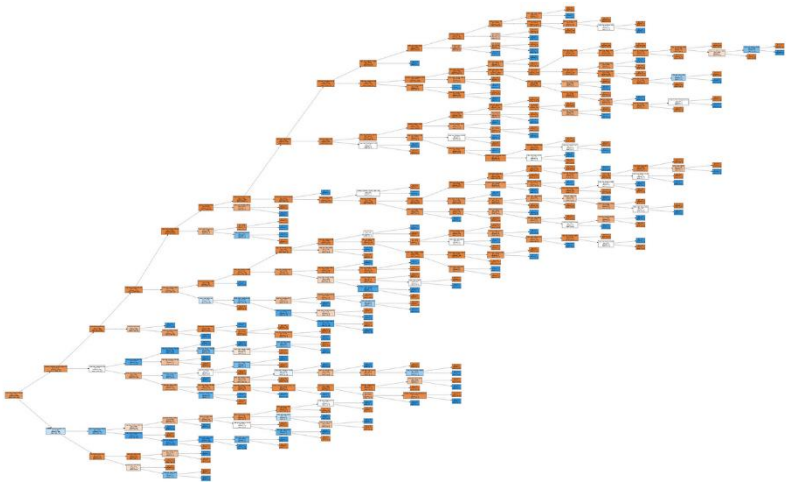


Fig 3 : Decision Tree for Customer Churn Prediction

Future enhancements may include deep learning integration for improved accuracy and real-time decision-making capabilities. Overall, the project highlights the potential of machine learning in optimizing customer engagement and reducing churn rates across various industries.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

CONCLUSION

The customer churn prediction model successfully identifies at-risk customers using machine learning techniques. Among the tested models, Random Forest demonstrated superior accuracy, highlighting its effectiveness in churn classification. The analysis revealed key factors influencing churn, such as contract type, payment method, and internet service. The decision tree visualization provides interpretability, aiding businesses in proactive customer retention strategies. By integrating predictive analytics, organizations can reduce churn rates, improve customer satisfaction, and optimize marketing efforts. Future enhancements could incorporate deep learning for greater precision and real-time data processing to further refine prediction capabilities and business decision-making.

REFERENCES

[1] T.Vafeiadis, K.I. Diamantaras, G.Sarigiannidis, K.Chatzisavvas “Customer churn prediction in telecommunications”, Simulation Modelling: Practice and Theory 55 (2015) 1-9.
[2] Burez J., & Van den Poel, D “Crm at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services”, Expert Systems with Applications 32, 277–288.
[3] S.Parvathavardhini and Dr. S.Manju “Analysis on Machine Learning Techniques” International Journal of Computer Sciences and Engineering (IJCSE), Vol-4(8), pp 59-77 Aug 2016, E-ISSN: 2347-2693.
[4] M.A.H. Farquad, Vadlamani Ravi, S. Bapi Raju “Churn prediction using comprehensible support vector machine: An analytical CRM application”, Applied Soft Computing 19 (2014) 31–40.
[5] Chih-Fong Tsai, Yu-Hsin Lu “Customer churn prediction by hybrid neural networks”, Expert Systems with Applications 36 (2009) 12547–12553.
[6] Wouter Verbeke, David Martens, Christophe Mues, Bart Baesens “Building comprehensible customer churn prediction models with advanced rule induction techniques”, Expert Systems with Applications 38 (2011) 2354–2364.
[7] Ning Lu, Hua Lin, Jie Lu, Guangquan Zhang “A Customer Churn Prediction Model in Telecom Industry Using Boosting”, IEEE Transactions on Industrial Informatics, vol. 10, no. 2, may 2014.