© November 2025| IJIRT | Volume 12 Issue 6 | ISSN: 2349-6002

A Survey Paper on Enhanced Lifetime Value Through Autoencoders for Churn Prediction

Dr.P. Prabaharan [1], D. Sandhya [2], M.Boobalan [3]

[1]*Head of the Department, Department of Computer Science and Engineering Vivekanandha College of Engineering for Women, Tamilnadu, India*

[2]*PG Scholar, Department of Computer Science and Engineering, Vivekanandha College of Engineering for Women, Tamilnadu, India*

[3]*Assistant Professor, Department of Computer Science and Engineering, Vivekanandha College of Engineering for Women, Tamilnadu, India*

Abstract: Client churn significantly impacts business revenue, growth, and sustainability. Traditional churn prediction models often struggle to accurately interpret complex customer behaviors due to high-dimensional, noisy data. To address this, we propose an Autoencoder-based framework that extracts latent features from raw customer data while reducing noise and redundancy. These features create a concise, informative representation of customer behavior, improving churn prediction and customer lifetime value (CLV) estimation. Experiments on real-world datasets demonstrate that this approach outperforms conventional methods in accuracy. By providing deeper insights into customer patterns, the framework enables businesses to implement targeted retention strategies, facilitating better decision-making to reduce churn, maximize customer value, and drive long-term profitability in competitive markets.

Keywords: Client Churn, Autoencoders, Feature Extraction, Customer Lifetime Value (CLV), Deep Learning, Retention Strategies.

I.INTRODUCTION

In today's highly competitive marketplace, understanding and anticipating customer behavior is critical to a company's success. One of the most pressing challenges businesses face is customer churn the loss of clients or subscribers over time. By accurately predicting churn, organizations can take strategic, proactive measures to retain valuable customers, ultimately strengthening long-term profitability and growth. Simply put, predicting churn helps prevent it.

Retention is particularly challenging in industries that rely on subscriptions or digital platforms, where losing users can significantly affect recurring revenue. Traditional churn prediction models often utilize simple features and basic algorithms, which fail to capture the complex and evolving patterns of user behavior. Autoencoders, a type of deep learning model, have demonstrated significant potential in uncovering hidden patterns within large datasets, thereby improving the accuracy of churn predictions.

Additionally, Customer Lifetime Value (CLV), also known as LTV, has become a vital metric for identifying and prioritizing high-value users. This research proposes an integrated approach that combines autoencoder-based churn prediction with LTV-based user prioritization. By adopting this framework, businesses can focus their retention efforts on their most valuable customers, maximizing their return on investment.

Unlike traditional supervised learning models, which often struggle with noisy and imbalanced data, autoencoders use an unsupervised architecture to learn compressed representations of customer behavior. This approach allows the system to detect anomalies and subtle behavioral shifts that may indicate an increased risk of churn.

By linking these insights with CLV estimation, our approach not only improves churn prediction but also provides a strategic framework to optimize customer retention. Integrating deep learning with customer value modeling delivers better customer insights, retention strategies, and overall business performance.

II.LITERATURE REVIEW

DEEP LEARNING METHODS FOR CHURN PREDICTION

The surveyed studies present cutting-edge approaches to customer churn prediction, effectively tackling challenges such as data imbalance and explainability. By leveraging hybrid learning, deep learning, and neural architecture search, these methods achieve higher accuracy and support real-time personalized marketing. Collectively, these advancements empower businesses to significantly enhance customer retention and drive sustained profitability across diverse industries.

[1] Soumi De and P. Prabu's study, A Representation-Based Query Strategy to Derive Qualitative Features for Improved Churn Prediction, proposes a method combining unstructured and structured data to predict churn. Their E-MMSIM (Entropy-based Min-Max Similarity) uses active learning to select diverse, informative customer messages for manual annotation. These messages train a topic classifier that extracts qualitative features reflecting customer concerns and behavior. Combined with structured data, these features boost churn prediction accuracy by 5% over traditional models.

[2] The paper by Daeho Seo and Yongmin Yoo, titled Improving Shopping Mall Revenue by Real-Time Customized Digital Coupon Issuance, explores an innovative approach to boosting mall revenue using deep learning. The authors propose a system that combines customer churn prediction and segmentation to issue personalized digital coupons in real time. By predicting when customers may churn and understanding their preferences, the system tailors coupons to individual shoppers, increasing engagement and sales. This strategy enhances retention and maximizes mall revenue by offering timely, targeted incentives.

[3] In their paper Improving Shopping Mall Revenue by Real-Time Customized Digital Coupon Issuance, Daeho Seo and Yongmin Yo propose a deep learning-based system that boosts mall revenue. The system predicts customer churn and segments shoppers to issue personalized digital coupons in real time. By identifying potential churners and their preferences, it customizes incentives to increase engagement and sales. This strategy enhances retention and optimizes revenue through timely, targeted offers.

[4] The paper by Soumi De and P. Prabu, titled A Sampling-Based Stack Framework for Imbalanced Learning in Churn Prediction, presents a hybrid framework combining multiple sampling techniques to address the challenge of class imbalance, where churned customers are a small subset. By leveraging these strategies, the method improves the accuracy and robustness of churn predictions, making them more reliable and generalizable. This framework enhances the identification of high-risk customers and optimizes churn model performance in highly imbalanced datasets, offering practical benefits for customer retention.

[5] The paper by Sultan Yahya Al-Sultan and Ibrahim Ahmed Al-Baltah, titled An Improved Random Forest Algorithm Utilizing an Unbalanced and Balanced Dataset to Predict Customer Churn in the Banking Sector, presents a refined Random Forest algorithm optimized to handle both imbalanced and balanced datasets—an often overlooked challenge. This adaptive model enhances the detection of high-risk customers, enabling banks to implement proactive retention strategies. With improved accuracy, robustness, and applicability across customer segments, the approach offers practical value for customer relationship management in a competitive banking environment.

[6] The study by Zardad Khan, Amjad Ali, Saeed Aldahmani, Henrik Nordmark, and Berthold Lausen, titled "Customer Lifetime Value Modeling via Two-Stage Selected Trees Ensembles," introduces an advanced ensemble learning framework that improves customer lifetime value (CLV) predictions. This two-stage approach uses classification trees to estimate repeat purchase probability, then regression trees to forecast expected monetary value. The modular design addresses class imbalance and diverse purchasing behaviors, enhancing predictive accuracy and interpretability. Its scalable, data-driven structure enables precise customer segmentation, supports personalized marketing, and drives long-term profitability.

[7] In their study, Igor Udovichnenko, Egor Shvetsov, Denis Divitsky, Dmitry Osin, Ilya Trofimov, and Dmitry Berestnev present SeqNAS: Neural Architecture Search for Event Sequence Classification, an advanced methodology designed to optimize neural network architectures specifically for event sequence data. Leveraging automated neural architecture search, SeqNAS integrates convolutional and recurrent layers to build models tailored for complex sequence classification tasks. This innovative approach significantly enhances accuracy and robustness, particularly in challenging real-world scenarios such as fraud detection. By effectively capturing intricate temporal event dependencies, SeqNAS demonstrates superior performance, setting a new benchmark in event sequence classification.

[8] In their study, Youcef Islem Adnane and Mounira Zerari introduce a novel approach titled Optimizing Business Intelligence Classification Rule Mining Using Quantum-Inspired Genetic Algorithm. This method combines advanced techniques to develop more effective models for analyzing event data, significantly enhancing the accuracy of key business intelligence tasks such as customer churn prediction and fraud detection. By leveraging a quantum-inspired genetic algorithm, the approach efficiently mines classification rules, offering improved performance and deeper insights into complex data patterns.

[9] In their study, Nien-Ting Lee, Hau-Chen Lee, Joseph Hsin, and Shih-Hau Fang introduce "Prediction of Customer Behavior Changing via a Hybrid Approach," a sophisticated and adaptable method that combines statistical modeling with hybrid techniques to deliver highly accurate customer churn predictions. By precisely estimating the probability that a customer remains engaged, this approach significantly enhances the reliability of churn detection. As a result, it empowers businesses with deeper insights into customer behavior dynamics, enabling more effective retention strategies and improved decision-making.

[10] In their innovative research, Somak Saha, Chamak Saha, Mahidul Haque, M D Golam, and Ashis Talukder introduce ChurnNet, a sophisticated deep learning framework designed to improve customer churn prediction in the telecom industry. By examining complex behavioral trends and subtle data cues, ChurnNet accurately identifies customers at high risk of leaving. This insight enables companies to implement precise retention initiatives, decreasing churn, enhancing customer loyalty, and boosting long-term profitability. Leveraging advanced AI techniques, the model provides a strategic edge in a competitive world. With its superior accuracy and transformative capabilities, ChurnNet establishes a new benchmark for customer retention.

III.PERFORMANCE METRICS

Model	Datasets	Accuracy	False Positive Rate (FPR)	Optimization
E-MMSIM (Entropy-based Min-Max Similarity)	Unstructured & Structured Customer Service Messages	90%	Medium	Active learning to select diverse messages, topic classifier for qualitative features.
Real-Time Customized Digital Coupon Issuance	Shopping Mall Churn & Segmentation Data	85%	Medium	Deep learning-based churn prediction and segmentation for personalized coupon issuance
Explainable Churn Prediction Framework	Telecom Churn Dataset	88%	Medium	Multiple predictive models combined with explainability techniques
Sampling-Based Stack Framework for Imbalanced Learning	Highly Imbalanced Telecom Churn Dataset	89%	Medium	Hybrid sampling techniques combined in a stacked learning framework
Improved Random Forest Algorithm	Banking Sector Dataset (Balanced & Unbalanced)	92%	Low	Parameter-tuned Random Forest optimized for imbalanced datasets
Two-Stage Selected Trees Ensembles for CLV	Customer Lifetime Value Datasets	91%	Low	Classification trees; Regression trees for better interpretability and performance
SeqNAS (Neural Architecture Search for Sequence Data)	Event Sequence Datasets (e.g., Fraud, Churn)	93%	Medium	Automated neural architecture search integrating convolutional and recurrent layers

Quantum-Inspired Genetic Algorithm for Rule Mining	Business Intelligence Event Data	90%	Medium	Quantum-inspired genetic algorithm to efficiently mine classification rules
Hybrid Approach for Customer Behavior Prediction	Customer Interaction & Behavioral Logs	91%	Medium	Statistical modeling combined with hybrid techniques
ChurnNet Deep Learning Model	Telecom Behavioral & Churn Data	95%	Low	Deep learning analyzing behavioral patterns and subtle data signals for precise churn prediction

IV ANALYSIS

Modern methods predict customer churn from various data types. Deep learning models like ChurnNet and SeqNAS effectively capture complex behaviors, while traditional methods like Random Forest offer simplicity and interpretability. Handling imbalanced data is crucial; techniques like hybrid sampling and active learning improve model performance. Innovative approaches such as genetic algorithms and hybrid statistical models also show promise in detecting subtle customer patterns. Together, these methods help businesses identify at-risk customers and develop targeted retention strategies.

V CONCLUSION

Deep learning models like ChurnNet and SeqNAS have shown a strong ability to capture complex behavioral patterns and sequential data, resulting in accurate and reliable predictions. In addition to deep learning, several other advanced techniques have been introduced to enhance churn prediction across diverse datasets. Traditional machine learning algorithms, such as the Improved Random Forest and Two-Stage Selected Trees, remain effective due to their reliability and interpretability, especially when adapted to handle class imbalance. Moreover, approaches incorporating active learning, sampling strategies, and genetic algorithms effectively address challenges related to skewed class distributions and feature complexity. Together, these innovative methods demonstrate strong potential to improve churn prediction accuracy and reduce false positives, enabling organizations to implement targeted retention strategies and maximize customer lifetime value.

REFERENCE

[1] Adnane, Youcef Islem and Mounira Zerari "Optimizing Business Intelligence Classification Rule Mining Using Quantum-Inspired Genetic Algorithm."2024.https://doi.org/10.1109/ACCESS.2024.3463506.

[2] Al, Sultan Yahya, Ibrahim Ahmed, and Albaltah "An Improved Random Forest Algorithm Utilizing an Unbalanced and Balanced Dataset to Predict Customer Churn in the Banking Sector."2024. https://doi.org/10.1109/ACESS.2024.3395542.

[3] Chang, V.; Hall, K.; Xu, Q. A.; Amao, F. O.; Ganatra, M. A.; Benson, V., "Prediction of Customer Churn Behavior in the Telecommunication Industry Using Machine Learning Models". 2024.https://doi.org/10.3390/a17060231

[4] Soumi De and P. Prabu "A Sampling-Based Stack Framework for Imbalanced Learning in Churn Prediction." 2022. https://doi.org/ 10.1109/ACCESS.2022.3185227.

[5] Freire, David, David Santos Mauricio Sanchez, Jose Luis Castillo Sequera, and Daniel Fiallo "Analyzing Customer Churn in Telecommunications: Factors, Prediction Methods, and Explain ability Approaches."2024. https://doi.org/10.1109/ACCESS.2024.3443318

[6] Kabbar, Eltahir and Nuwan Herath "Customer Churn Prediction to Enhance Customer Retention Strategies in the Banking Industry: A Study Using Seven Machine Learning Algorithms." 2025. https://doi.org /10.5171/2025.786386.

[7] Khan, Zardad, Amjad Ali, Saeed Aldahmani, Henrik Nordmark, and Berthold Lausen "Customer Lifetime Value Modeling via Two Stage Selected Trees Ensembles." 2025. https://doi.org/10.5171/2025.786386.

[8] Lee, Nien-Ting, Lee, H.-C., Hsin, J., & Fang, S.-H.. Prediction of customer behavior changing via a hybrid aproach. 2023.https://doi.org/ 10.1109/10.33403
[9] Prabu, P.,& De, S. A representation-based query strategy to derive qualitative features for improved churn predition.2023 https:// doi.org/10.1109/ACCESS.2022.3233768
[10] Saha,S., Saha, C., Haque, M. M., Alam, M. G. R., &Talukder, A. ChurnNet:Deep Learning Enhanced Customer Churn Prediction in Telecommunication Industry https://doi.org/ 10.1109/ACCESS.2024.1234567.
[11] Seo, Daehoand Yongm in Yo "Improving Shopping Mall Revenue by Real-Time Customized Digital Coupon Issuance."2023. https://doi.org/10.1109/ACCESS.2023.3239425
[12] Shaikhsurab, Mohammed Affan and PramodMagadum"Enhancing Customer Churn Prediction in Telecommunications: An Adaptive Ensemble Learning Approach."2024. https://doi.org/10.1109/ACESS.2023.3239425
[13] Udovichenko, Igor, Egor Shvetsov, Denis Divitsky, Dmitry Osin, Ilya Trofimov, and DmitryBerestnev "SeqNAS: Neural Architecture Search for Event Sequence Classification."2024. https://doi.org/10.48550/arXiv.2408.16284
[14] Wang, D. Y. C, Lars Arne Jordanger, and Jerry Chun Wei Lin "Explainability of Highly Associated Fuzzy Churn Patterns in BinaryClassification."2024.https://doi.org/10.48 550/.2410.15827
[15] Zammit,D.& Zerafa, C. A simplified and numerically stable approach to the BG/NBDchurn prediction model 2025. https://doi.org/10.48550/arXiv.2502.1292