ISSN 2277-2685
IJESR/Jan-Mar. 2025/ Vol-15/Issue-1/656-665
T Baba *et. al.*, / International Journal of Engineering & Science Research

Using Big Data And Explainable Ai In A Hybrid Model To Predict Churn And Boost Customer Retention In Streaming Services

[1] T Baba, [2] Ganapa Jayaveera Raghav
[1] Assistant Professor, M. Tech, [2] M. Tech (Student)
[1, 2] Department of Computer Science and Engineering, PVKK institute of technology, Anantapur, Andhra Pradesh, India.

Abstract

For streaming services, where keeping current members is vital for ongoing success, customer churn prediction is a major concern. This paper offers a large data-driven hybrid model predicting customer turnover with excellent accuracy by sophisticated machine learning and deep learning technologies. The version addresses data imbalance using SMOTE oversampling on the Churn data dataset. Predictive performance is improved by means of Chi-square (Chi2) and Sequential feature selection (SFS), optimising feature selection. Although several algorithms were used, the emphasis is on a voting Classifier combining boosted models (LightGBM and XGBoost) and CNN-LSTM to reach best accuracy. These methods together have a prediction accuracy over 95%. Integrated to improve interpretability, explainable artificial intelligence gives insights on churn causes. A Flask-based user interface with authentication also guarantees smooth user interaction, hence facilitating brief testing and deployment. In competitive industries, this hybrid approach shows great promise for increasing accuracy of churn prediction and supporting client retention.

"Index Terms - *Artificial intelligence, classification algorithms, deep learning, decision support systems, explainable AI, model interpretation, semi-supervised learning, big data analysis".*

1. INTRODUCTION

Customer churn [1] refers to the sad state of affairs in which consumers terminate their commercial association with a firm. A consumer moving to a rival, cancelling a subscription or ceasing use of a product or service could all contribute to this. Businesses are greatly worried about customer churn since it can result in lower income, lower customer lifetime value, and a bad effect on brand reputation [2]. Drawing new clients can be expensive and the loss of customers means a depletion of important income streams [3]. High customer attrition rates can imply more marketing and acquisition expenses for companies trying to replace lost customers [3]. According to Harvard business review, "the cost of creating a new customer is 5 to 25 instances the cost of keeping an old customer" before any company uses a new strategy, it is crucial to remember that their customer base should not be diminished. Any company adopting a new strategy should first consider that their client base should not be diminished. Reducing churn is extra important than ever.

Analysts say customer fall has caused the company notable income loss. The process of examining those consumers' behaviour and taking the suitable actions to address its [1] gains tremendous weight and relevance as a result. Reducing customer churn calls for good data analysis, customised marketing strategies and improved customer service [5]. Using machine learning and other techniques to forecast consumer behaviour helps businesses enhance customer satisfaction, lower churn rates, and promote corporate expansion in several ways. Groups can take proactive steps to keep consumers by finding those most likely to leave and exposing the causes of their departure. These steps should include providing focused promotions, fair pricing, or enhancing customer service. Many statistics exist for certain sectors like communications to forecast consumer churn or customer behaviour. But, there are hardly ever no big datasets on line when one considers the study of consumer behaviour for streaming sites or e-commerce sectors. Collected from the Udacity program, a customer log dataset [6] was developed in partnership with Appen, IBM Watson, Bertelsmann, insight, Kaggle, and Starbucks. This dataset is ideal for this research since it closely suits a real-life dataset or log file of any streaming service including Spotify, Deezer, Soundcloud, Gaana etc.

This paper aims to build a machine-learning pipeline capable of predicting consumer attrition in the streaming service sector. This project aims to assist businesses increase their profitability. While the modern literature has often used both deep neural network models and classic machine learning models, our approach's innovation is the blending of new deep learning techniques with conventional machine learning approaches. Our deep learning algorithms as well as traditional models like Catboost, XGBoost and logistic regression are used. Furthermore, we suggest a novel hybrid machine-learning model that combines the benefits of deep learning with conventional machine-learning techniques. Our research also provides a primer on feature selection using the Chi-square test and "Sequential feature selection (SFS), followed via the Shap and Explainable Boosting model (EBM)" for model interpretation. Furthermore, we have included a semi-supervised learning approach to evaluate our model against unlabelled facts, that is more appropriate in situations while labelled data is limited but unlabelled data is ample such as deployments in actual-world settings. The performance of the churn prediction model can be improved via maximising using the huge unlabelled data using this combined method.

2. RELATED WORK

Its great influence on company profitability and client retention tactics has drawn much interest in accurately forecasting consumer churn. Many studies have investigated different machine learning and deep learning techniques to address this issue. But as consumer behaviour and contact with a service often show temporal linkages that can suggest approaching churn, properly capturing sequential patterns and temporal dependencies in client data has become a vital component.

Incorporating four separate "algorithms—decision Tree, Random forest, Gradient Boosted machine Tree (GBM) and extreme Gradient Boosting (XGBOOST)—Ahmad et al". [9] Created a thorough strategy for telecom companies. With an AUC of 93%, the XGBOOST algorithm showed better performance. A vital 10-fold go-validation was used to fine-tune hyperparameter values, thus splitting the dataset into 70% training and 30% testing. The research, meantime, did not investigate model interpretability or the influence on performance of various feature selection methods.

Using knowledge extraction, Jamjoom [10] investigated data mining methods to forecast client turnover inside insurance companies. While neural networks flourished with a 70:30 distribution, their study showed the efficacy of a 50:50 teach-test distribution for logistic regression. The decision tree models had the fine performance with a 50:50 training/check split, getting an AUC of 0.73. The study investigated clustering methods including k-way for finding homogeneous consumer segments however more sophisticated algorithms like DBSCAN have demonstrated superior performance in recent customer segmentation research. The study also lacked clear discussion on issues or constraints connected to data availability, data quality, or facts preparation, which could affect the accuracy and dependability of the findings.

Using important factors such customer viewing intensity, consumption patterns, payment behaviours and preferences, Li et al. [11] undertook a study aimed at forecasting customer attrition in the cable tv area. This work underlined the want of client loyalty and skilfully used the bell-shaped price curve to identify customer patterns.

"Long short-term memory (LSTM) and Convolutional Neural Networks (CNNs)", among other deep learning architectures, have shown their efficacy in modelling sequential data. Kwon et al. [13] used an LSTM model to extract complex time-series patterns from text messages and customer lifelog data, hence forecasting customer attrition in the digital healthcare sector with an outstanding F1-score of 89%. Although their method successfully used topic modelling and recurrent neural networks, it struggled to identify the ideal amount of topics without previous knowledge, which could affect model accuracy. The quality and completeness of the lifelog data could also affect the model's performance as missing or noisy data could cause biases and errors. Moreover, as the characteristics and patterns in the data ought to fluctuate greatly across various settings, their model's generalisability to other kinds of digital healthcare applications or domains still uninvestigated.

Using textual data from a eu financial services provider, de Caigny et al. [18] proposed an approach meant to improve churn prediction models. Thru careful 5*3 cross-validation and AUC and "top Decile list (TDL) criteria, their work showed that Convolutional Neural Networks (CNNs)" outperformed current techniques in extracting sequential patterns from text material. Their work, however, concentrated just on CNNs as the deep learning architecture for handling textual input. Although CNNs are ideal for text categorisation, other architectures like "recurrent neural networks (RNNs)" or attention-based models may want to provide different viewpoints or maybe superior results in some situations. The research also omitted the CNNbased models' interpretability and explainability features, which are vital for practical uses and regulatory compliance in sectors as banking. Moreover, the research lacked a thorough framework or guidance for practitioners on how to properly include and control textual facts sources into their current "customer relationship management (CRM)" systems or data pipelines.

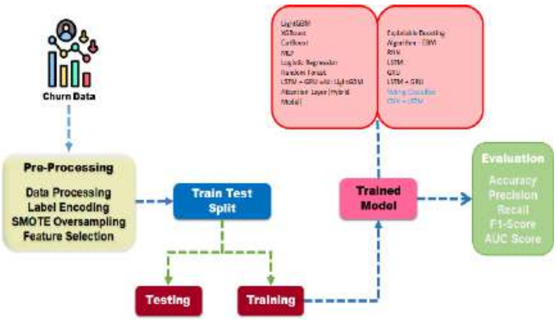
Focussing on forecasting consumer churn in the banking sector, Zorić [19] offered an enlightening case have a look at. Using a neural network in the Alyuda NeuroIntelligence software package, the researchers carefully assessed model performance using measures including CCR (correlation coefficient) and MSE, hence offering insightful analysis of the area of churn prediction. Although neural networks were shown to be successful for customer churn prediction purposes, adding extra machine learning algorithms like decision trees, random forests or ensemble techniques could offer a more thorough view. Comparative studies across several methods would allow a better knowledge of their relative advantages and disadvantages, hence guiding the choice of the most appropriate method for a sure problem setting.

3. MATERIALS AND METHODS

Using a big data-driven hybrid approach, the suggested method seeks to solve problems with user attrition prediction in streaming services. It optimises feature selection by "Chi-square (Chi2)

ISSN 2277-2685
IJESR/Jan-Mar. 2025/ Vol-15/Issue-1/656-665
T Baba *et. al.,* / International Journal of Engineering & Science Research

and Sequential feature selection (SFS)" techniques [8] and SMOTE oversampling to solve data imbalance using the Churn data dataset. among the many machine learning and deep learning methods the system includes are "LightGBM, XGBoost, CatBoost, Random forest, MLP, Logistic Regression, RNN, LSTM, GRU, CNN-LSTM, a hybrid LSTM-GRU model with a LightGBM attention layer", and a voting Classifier integrating boosted models. The solution offers interpretable forecasts by combining these algorithms with Explainable AI (XAI), hence delivering actionable insights on churn drivers. The adoption of a Flask-based user interface with user authentication guarantees accessibility and simplified interaction as well. By using exact and actionable attrition forecasts, this strategy offers a viable way to improve client retention, hence satisfying the needs of the competitive streaming industry [8].



"Fig.1 Proposed Architecture"

The suggested churn prediction solution increases customer retention in streaming services by means of a methodical pipeline [8]. Beginning with churn data gathering, necessary pre-processing procedures including data processing, label encoding, SMOTE oversampling, and feature selection guarantee data quality and balance. SMOTE (synthetic Minority Over-sampling technique) is used particularly to generate synthetic instances for the minority class, hence correcting class imbalance and enhancing version performance. The dataset is divided into training and testing sets once pre-processing is finished to permit efficient model evaluation. Churn prediction is done by the system using both deep learning and machine learning models. While "deep learning models are RNN, LSTM, GRU, CNN-LSTM, and a hybrid LSTM-GRU model with a LightGBM attention layer", improving feature learning and predictive "accuracy, machine learning methods include LightGBM, XGBoost, CatBoost, MLP, Logistic Regression, and Random forest". A voting Classifier is also included to combine enhanced models for better general performance. Techniques of "explainable artificial intelligence (XAI), including the Explainable Boosting machine (EBM)", are included to improve model interpretability and offer analysis of churn drivers. Ultimately, the trained models are assessed using important performance criteria—Accuracy, Precision, consider, F1-rating, and AUC score—thereby guaranteeing the dependability of churn forecasts. This design lets in streaming services to use data-driven consumer retention tactics by prediction accuracy and interpretability balance [8].

i) Dataset Collection:

Attrition data, a repository with consumer behavioural and subscription-associated characteristics, gives the dataset wished for attrition prediction in streaming services [3]. Critical characteristics in this dataset that significantly have an effect on churn prediction are watch time, subscription duration, fee history, consumer engagement statistics, and platform interaction frequency. Demographic facts and historic churn behaviour are also included to give a complete picture of client retention trends. The dataset goes through pre-processing procedures including missing value management, data normalisation, and feature encoding to guarantee data for model training. "SMOTE (synthetic Minority Over-technique sampling)" is used to provide synthetic samples for the minority (churned) class given that

ISSN 2277-2685
IJESR/Jan-Mar. 2025/ Vol-15/Issue-1/656-665
T Baba *et. al.*, / International Journal of Engineering & Science Research

churn datasets are usually unbalanced, thereby generating a more balanced dataset for higher model performance. The processed dataset is subsequently divided into training and testing sets, hence allowing efficient assessment of deep learning and machine learning models [4]. Using this dataset, the system may create prediction models to precisely spot possible churners, hence enabling streaming services to use proactive client retention tactics.

Unnamed: 0	user_id	watch_time	subscription_duration	payment_history	engagement_score
0	101	12.5	24	on-time	85
1	102	3.2	6	late-payment	45
2	103	8.0	12	on-time	67
3	104	1.5	3	missed-payment	20
4	105	10.7	18	on-time	78
...	...	...	...	...	...
9999	11100	4.8	9	late-payment	55

"Fig.2 Dataset"

ii) Pre-Processing:

A key stage in the churn prediction system, pre-processing guarantees that raw data is cleansed, balanced, and optimised for model training. Data processing starts the process; it includes dealing with missing values using imputation techniques like mean or median filling for numerical attributes and mode-based filling for categorical features [5]. Then, label encoding turns categorical variables—such as subscription type or payment history—into numerical representations, hence qualifying them for machine learning models. SMOTE (synthetic Minority Over-sampling technique) is used to provide synthetic instances for the minority class—churned customers—thereby guaranteeing balanced model training and therefore addressing the class imbalance in churn datasets [8]. Techniques of feature selection like "Chi-square (Chi2) and Sequential feature selection (SFS)" are also used to find the most pertinent predictors, hence enhancing model efficiency and interpretation [4]. Applying these pre-processing techniques helps the system guarantee that the data is well-structured, representative, and optimised for predictive modelling, hence producing more accurate and dependable churn forecasts.

iii) Training & Testing:

The dataset is divided into training and testing sets following pre-processing to guarantee efficient model evaluation. Usually, an 80-20 or 70-30 split is used, with the training set used for learning patterns and the testing set assessing model performance [5]. Trained on the processed data are several machine learning (LightGBM, XGBoost, CatBoost, Random forest) and deep learning (LSTM, GRU, CNN-LSTM) models. Metrics including accuracy, precision, recall, F1-score, and AUC score are used to validate performance. This approach guarantees that the model generalises well to unobserved data, therefore allowing dependable and accurate churn forecasts for streaming services [4].

iv) Algorithms:

LightGBM (Light Gradient Boosting Machine)

Developed by Microsoft, LightGBM is a gradient boosting framework using tree-based learning techniques. Optimised for speed and efficiency, it handles big datasets with low memory use. Particularly for structured facts issues like fraud detection, recommendation systems, and financial forecasting, LightGBM is often utilised for classification and regression work. It does exceptionally well with missing values and categorical data management. LightGBM's goal is to keep great accuracy while increasing computing efficiency. Making it appropriate for big-scale machine learning projects where performance is crucial, it speeds up training using histogram-based learning and unique feature bundling.

XGBoost (Extreme Gradient Boosting)

An enhanced gradient boosting technique, XGBoost increases the efficiency, scalability, and accuracy of machine learning models. Designed to reduce computing time while preserving great predictive power, it's far based on decision trees. Often utilised in structured data challenges, XGBoost includes customer churn prediction, credit risk assessment, sales forecasting, and competition-winning data science models. Its performance in corporate applications and Kaggle contests is quite good. By parallel processing, regularisation approaches (L1 and L2), and effective management of sparse data, XGBoost improves conventional gradient boosting. Its performance is enhanced and overfitting is decreased, hence it is favoured for predictive modelling.

$$y\hat{}i=\sum k=1Kfk(xi) \quad (1)$$

CatBoost (Categorical Boosting)

Designed by Yandex, CatBoost is a gradient boosting method meant to effectively handle category features. Fast training and excellent accuracy in machine learning tasks are achieved by use of ordered boosting and GPU acceleration. In situations with categorical data—such as recommendation systems, fraud detection, and medical diagnosis—CatBoost is mainly helpful. It removes the requirement for significant data preprocessing, such one-hot encoding. Using ordered boosting and effective feature management, CatBoost hopes to lower overfitting and increase prediction accuracy. Designed to need little parameter adjustment, it must perform well with modest to large data sets.

$$F(x)=F0(x)+\sum m=1M\sum i=1Nfm(xi) \quad (2)$$

MLP (Multi-Layer Perceptron)

Comprising several layers of neurons—including input, hidden, and output layers—MLP is a kind of artificial neural network. It learns complicated patterns using backpropagation and activation functions. Deep learning applications including image identification, natural language processing, and financial forecasting use MLP. For activities needing nonlinear function approximation, it works well. MLP aims to apply deep learning tools to model complex interactions between inputs and outputs. Often used in supervised learning projects for classification and regression, it offers flexibility in learning patterns.

Logistic Regression

A statistical model called logistic regression uses a logistic function to simulate binary results. It predicts the likelihood that a certain input fits a certain class. Commonly employed in medical diagnosis, spam detection, credit scoring, and marketing campaigns, it classifies data into two groups: "yes" or "no," "fraud" or "not fraud." Logistic Regression aims to simulate the link between independent variables and a binary-based variable. Its clear findings help to guide decisions.

$$\hat{}y=a+bx \ y \hat{} = a + b x \quad (3)$$

Random Forest

An ensemble learning technique called Random forest increases accuracy and lowers overfitting by combining several decision trees. It runs by building many decision trees and averaging their results. Used in classification and regression projects including medical diagnosis, stock price prediction, and consumer segmentation. It is resilient to missing values and noisy data. “Random forest aims to offer great predicting accuracy and keep interpretability”. Combining several weak learners—decision trees—helps to improve generalisation and stop overfitting.

Explainable Boosting Algorithm (EBM)

Based on "Generalised Additive models (GAMs)", EBM is a clear machine learning technique. It guarantees openness in decision-making and offers great prediction accuracy. It is applied in legal, financial, and medical fields where model interpretability is very important. Valuable for regulatory compliance, it clarifies the basis for a prediction. EBM's primary goal is to provide an understandable machine learning model that moves a compromise between performance and openness.

It lets people believe and grasp decisions made by artificial intelligence.

RNN (Recurrent Neural Network)

Designed for sequential data processing, RNN is a deep learning model. Unlike conventional neural networks, it has links that let information survive over time steps. Applications include time-series forecasting, machine translation, and speech recognition use RNNs. They are efficient in jobs requiring past knowledge to forecast. RNN's goal is to keep context via hidden states, thereby processing sequential data. Natural language processing and dynamic system modelling both make great use of it.

LSTM (Long Short-Term Memory)

A kind of RNN called LSTM uses memory cells and gating systems to keep long-term dependencies, as a result solving the vanishing gradient issue. Applications of LSTM include sentiment analysis, financial time-series forecasting, chatbot creation, and speech recognition. It shines at modelling long-range dependencies. LSTM's goal is to let deep learning models keep pertinent prior information for improved decision-making. it works especially well with long-term dependent sequential data.

“An algorithm calculates classifiers by using the formula”:

$$y = f^* (x). \qquad (4)$$

GRU (Gated Recurrent Unit)

Using just two gates—update and reset gates—GRU is an LSTM variant that streamlines the architecture. Its lower computing costs give comparable performance. Tasks in video captioning, text synthesis, and machine translation employ GRU. In sequential modelling uses, it works well. GRU ambitions to increase computing efficiency while keeping the capacity to learn long-term dependencies. For resource-limited settings, it is a substitute for LSTM.

LSTM + GRU

A hybrid deep learning model using LSTM and GRU components to maximise the benefits of each architectures in managing sequential dependencies. Applied in anomaly identification, financial forecasting, and NLP where many temporal patterns exist in a dataset needing both short-term and long-term memory retention. Combining LSTM and GRU serves to improve sequence learning by using LSTM's long-term memory retention and GRU's computational efficiency.

LSTM + GRU with LightGBM Attention Layer (Hybrid Model)

This hybrid approach uses LightGBM for structured data learning and combines LSTM and GRU with an attention mechanism that gives priority to big features. Applied in healthcare diagnostics, fraud detection, and customer churn forecasting whilst sequential patterns must be paired with structured data insights. By using deep learning for sequential data and boosting techniques for structured feature learning, this model aims to increase prediction accuracy.

Voting Classifier

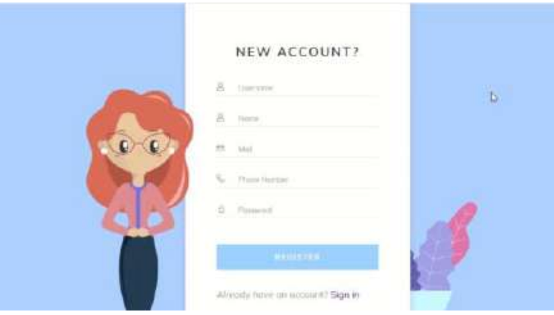
Often employing weighted or majority voting to enhance accuracy, a voting classifier is an ensemble learning approach that aggregates predictions from several models. Applied in advice systems, medical diagnostics, and financial fraud detection where many models have complementing strengths. A voting classifier aims to boom predictive performance by combining outputs from several models, hence lowering individual biases and enhancing generalisation.

CNN + LSTM

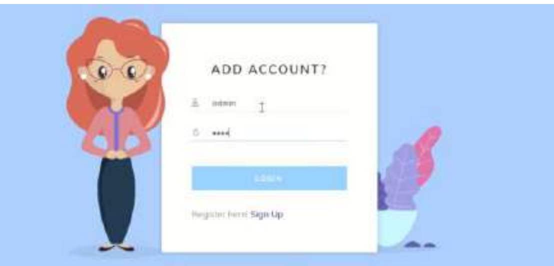
A hybrid deep learning model combining LSTMs for sequential pattern recognition and "Convolutional Neural Networks (CNNs)" for feature extraction. Applied in sentiment classification, voice emotion detection, and video analysis when spatial and temporal aspects must be captured. CNN + LSTM seeks to use CNNs for spatial feature learning and LSTMs for sequential dependencies, therefore enabling strong deep learning models for complicated applications.

4. RESULTS & DISCUSSION

“Fig 3 Welcome dashboard”



“Fig 4 User registration”



“Fig 5 User login”

“Fig 6 Main page”

“Fig 7 User input”

“Fig 8 Prediction result”

5. CONCLUSION

Ultimately, the suggested big data-driven hybrid model for customer attrition prediction in streaming services reveals a strong way to enhance retention tactics. The model produces very correct churn forecasts by using sophisticated machine learning and deep learning technologies to properly handle important issues including data imbalance and feature selection. while the voting Classifier—composed of LightGBM, XGBoost, and CNN-LSTM—greatly improves predictive performance to over 95%, the combination of SMOTE oversampling, Chi-square, and "Sequential feature selection (SFS)" improves the version's capacity to locate important churn-influencing factors. Integration of Explainable AI guarantees openness and aids in delivering actionable insights into churn factors, hence enabling companies to perform focused retention plans. Moreover, the Flask-based user interface with authentication improves accessibility and allows easy testing and deployment. All things considered, this hybrid approach is a useful tool for streaming services trying to lower churn rates and hold a strong, loyal purchaser base within the competitive market since it offers a quick, scalable answer for customer churn forecasting.

Future study will aim at improving the hybrid model by including more sophisticated deep learning architectures, including Transformer-based ones, to raise churn prediction accuracy. Including dynamic customer behaviour tracking and real-time data streams will also assist to improve predictive skills. Methods of explainable artificial intelligence will be enlarged to offer more in-depth interpretation of churn drivers. Moreover, running the model in a scalable cloud environment and testing on various datasets would guarantee compatibility across several streaming services, hence maximising client retention plans.

REFERENCES

[1] Saha, S., Saha, C., Haque, M. M., Alam, M. G. R., & Talukder, A. (2024). ChurnNet: Deep Learning Enhanced Customer Churn Prediction in Telecommunication Industry. *IEEE Access*.

[2] Shaikhsurab, M. A., & Magadum, P. (2024). Enhancing Customer Churn Prediction in Telecommunications: An Adaptive Ensemble Learning Approach. *arXiv preprint arXiv:2408.16284*.

[3] Zhang, L., & Wei, Q. (2024). Personalized and Contextualized Data Analysis for E-Commerce Customer Retention Improvement With Bi-LSTM Churn Prediction. *IEEE Transactions on Consumer Electronics*.

[4] Anudeep, C., Venugopal, R., Aarif, M., Valavan, A. T., Vuyyuru, V. A., & Muthuperumal, S. (2024, June). Predicting Customer Churn in E-commerce Subscription Services using RNN with Attention Mechanisms. In *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-6). IEEE.

[5] Mengash, H. A., Alruwais, N., Kouki, F., Singla, C., Abd Elhameed, E. S., & Mahmud, A. (2023). Archimedes Optimization Algorithm-Based Feature Selection with Hybrid Deep-Learning-Based Churn Prediction in Telecom Industries. *Biomimetics*, *9*(1), 1.

[6] A. A. Jamjoom, ''The use of knowledge extraction in predicting customer churn in B2B,'' J. Big Data, vol. 8, no. 1, pp. 1–14, Dec. 2021.

[7] Y. Li, B. Hou, Y. Wu, D. Zhao, A. Xie, and P. Zou, ''Giant fight: Customer churn prediction in traditional broadcast industry,'' J. Bus. Res., vol. 131, pp. 630–639, Jul. 2021.

[8] C. Colot, P. Baecke, and I. Linden, ''Leveraging fine-grained mobile data for churn detection through essence random forest,'' J. Big Data, vol. 8, no. 1, p. 63, 2021.

[9] H. Kwon, H. H. Kim, J. An, J.-H. Lee, and Y. R. Park, ''Lifelog data-based prediction model of digital health care app customer churn: Retrospective observational study,'' J. Med. Internet Res., vol. 23, no. 1, Jan. 2021, Art. no. e22184.

[10] J. Kozak, K. Kania, P. Juszczuk, and M. Mitrega, ''Swarm intelligence goal-oriented approach to data-driven innovation in customer churn management,'' Int. J. Inf. Manage., vol. 60, Oct. 2021, Art. no. 102357.

[11] U. Joy. (2023). Customer Churn Dataset. [Online]. Available: https://ieeedataport.org/documents/customer-churn-dataset

[12] Afzal, M., Rahman, S., Singh, D., & Imran, A. (2024). Cross-sector application of machine learning in telecommunications: enhancing customer retention through comparative analysis of ensemble methods. *IEEE Access*.

[13] Chang, V., Hall, K., Xu, Q. A., Amao, F. O., Ganatra, M. A., & Benson, V. (2024). Prediction of Customer Churn Behavior in the Telecommunication Industry Using Machine Learning Models. *Algorithms*, *17*(6), 231.

[14] Madhuranthakam, R. S. (2024). Advanced Algorithms for Enhancing Customer Churn Prediction Accuracy using Real-Time Data Analytics. *International IT Journal of Research, ISSN: 3007-6706*, *2*(2), 205-210.

[15] Patil, S., & Mohammed, A. S. PROACTIVE CRM: PREDICTING CUSTOMER BEHAVIOR

AND CHURN USING MACHINE LEARNING MODELS. *Journal ID*, *2811*, 6201.

[16] Adeniran, I. A., Efunniyi, C. P., Osundare, O. S., Abhulimen, A. O., & OneAdvanced, U. (2024). Implementing machine learning techniques for customer retention and churn prediction in telecommunications. *Computer Science & IT Research Journal*, *5*(8).

[17] Zhang, L., & Wei, Q. (2024). Personalized and Contextualized Data Analysis for E-Commerce Customer Retention Improvement With Bi-LSTM Churn Prediction. *IEEE Transactions on Consumer Electronics*.

[18] Manzoor, A., Qureshi, M. A., Kidney, E., & Longo, L. (2024). A Review on Machine Learning Methods for Customer Churn Prediction and Recommendations for Business Practitioners. *IEEE Access*.

[19] Jahan, I., & Sanam, T. F. (2024). A comprehensive framework for customer retention in E-commerce using machine learning based on churn prediction, customer segmentation, and recommendation. *Electronic Commerce Research*, 1-44.

[20] Feng, Y., Yin, Y., Wang, D., Ignatius, J., Cheng, T. C. E., Marra, M., & Guo, Y. (2024). Enhancing e-commerce customer churn management with a profit-and AUC-focused prescriptive analytics approach. *Journal of Business Research*, *184*, 114872.

[21] Văduva, A. G., Oprea, S. V., Niculae, A. M., Bâra, A., & Andreescu, A. I. (2024). Improving Churn Detection in the Banking Sector: A Machine Learning Approach with Probability Calibration Techniques. *Electronics*, *13*(22), 4527.

[22] Wanda, R. W. K. (2024). *A Customer churn prediction and corrective action suggestion model for the telecommunications industry using predictive analytics* (Doctoral dissertation, Strathmore University).

[23] Behl, A. R., Hadpe, H., Bonde, K., & Patel, H. (2024). Customer Churn Prediction for a Telecommunication Company. *International Journal of Research in Engineering, Science and Management*, *7*(3), 107-113.

[24] Shahabikargar, M., Beheshti, A., Mansoor, W., Zhang, X., Foo, J., Jolfaei, A., ... & Shabani, N. (2024). Generative AI-enabled Knowledge Base Fine-tuning: Enhancing Feature Engineering for Customer Churn.

[25] Hoff, F. (2024). Generative AI in Customer Churn Processes: The integration and role of Generative AI in Customer Churn.