isrjournals

International Journal of Advanced Research in Computer Networking,Wireless and Mobile Communications

Volume: 5 Issue: 3 Apr,2025,ISSN_NO: 2320-7248

CUSTOMER CHURNPREDICTION USING RECURRENT NEURAL NETWORKS AND LONG SHORT-TERM MEMORY NETWORKS

Asmitha.KC[1],Stella Marry. S[2], Sandhiya .S[3],Shanmugavalli .H[*4]

[*123]UG Scholar–Dept of CSE, GRT Institute of Engineering and Technology Tiruttani.

[*4]Assistant Professor-Dept of CSE, GRT Institute of Engineering and Technology, Tiruttani.

Asmithakc16@gmail,.com , stellamarry2002@gmail.com , sandhiya230619@gmail.com

[4]*Corresponding Author : shanmugavalli.h@grt.edu.in

***Abstract*:** Customer churn prediction is a critical aspect of customer retention strategies in industries such as telecommunications, banking, and e-commerce. The goal of this project is to predict the likelihood of customers discontinuing services by analyzing historical customer data using advanced deep learning techniques, specifically Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks. These models are well-suited for time-series data and sequential patterns, making them ideal for detecting churn trends in customer behavior over time. In this project, a dataset containing historical customer information such as demographics, transaction history, usage patterns, and service interactions is used to train the RNN and LSTM models. By learning patterns from sequential data, the models can capture dependencies in customer behavior and predict the probability of churn. The RNN helps in processing sequences of customer interactions, while the LSTM addresses the issue of long-term dependencies in the data, enhancing prediction accuracy.

Keywords: Customer Churn, Churn Prediction, Customer Retention, Predictive Modelling, Machine Learning.

1.INTRODUCTION

Customer churn is a significant challenge for businesses in competitive industries like telecommunications, banking, and e-commerce. Churn occurs when customers discontinue a service, leading to revenue loss and increased customer acquisition costs. Predicting churn in advance allows businesses to implement proactive retention strategies, improving customer satisfaction and loyalty. Traditional machine learning models often struggle with sequential data dependencies, making deep learning techniques like Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks more effective for churn prediction. These models customer interactions over time, identifying subtle behavioural patterns that indicate the likelihood of churn. This project utilizes a historical customer dataset, including demographics, transaction history, and service usage, to train and evaluate RNN and LSTM .[2]

2.RELATEDWORK

Customer churn prediction has been an active area of research due to its significant impact on business profitability and customer relationship management. Several studies have explored various machine learning and deep learning techniques to predict churn across different industries such as telecommunications, banking, and e-commerce.

2.1 REVIEW ON MACHINE LEARNING METHODS FOR CUSTOMER CHURN PREDICTION AND RECOMMENDATIONS FOR BUSINESS PRACTITIONER

Customer churn directly impacts business revenue and increases operational costs due to the need for acquiring new customers. Machine learning provides a data-driven approach to predict potential churners and enables businesses to take preventive actions. Various ML models)**[12]** — from traditional statistical methods to advanced deep learning models — have been proposed to address this challenge.
Limitation-Despite significant advancements in machine learning techniques for customer churn prediction, several limitations remain. One major challenge is the poor quality and imbalance of real-world customer datasets, where churners often represent a small portion of the data, leading to biased predictions.

2.2 CHURN-NET: DEEP LEARNING ENHANCED CCP IN TELE INDUSTRY

Churn-Net[11] is a deep learning-based framework designed to enhance customer churn prediction (CCP) in the telecommunications industry. It leverages advanced neural network architectures, particularly Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) models, to sequential customer data such as call records, service usage, and transaction history. **Limitation** Although Churn-Net demonstrates high prediction accuracy, it faces several limitations. The model requires large volumes of high-quality, time-series customer data, which may not always be available.

2.3 ARITHMETIC OPTIMIZATION WITH ENSEMBLE DL SBLSTM-RNN-IGSA MODEL FOR CCP

The Arithmetic Optimization with Ensemble Deep Learning SBLSTM-RNN-IGSA model [2] is proposed for efficient customer churn prediction (CCP). It integrates stacked bidirectional LSTM (SBLSTM), RNN, and improved grey wolf optimizer with arithmetic optimization to enhance prediction accuracy and feature selection
Limitation The model involves high computational complexity due to multiple integrated algorithms. Additionally, its performance heavily depends on parameter tuning and large-scale datasets, which may limit practical implementation in resource-constrained environments.

2.4 A REPRESENTATION- BASED QUERY STRATEGY TO DERIVE QUALITATIVE FEATURES FOR IMPROVED CCP

The Representation-Based Query Strategyis introduced to extract qualitative[4] features from customer data for improved customer churn prediction (CCP). This approach focuses on generating meaningful feature representations that enhance model learning and prediction accuracy.
Limitation The method relies on complex feature extraction processes, which may increase the overall computational cost. Additionally, its effectiveness may decrease when applied to datasets with limited or noisy customer information.

2.5 A SAMPLING-BASED STACK FRAMEWORK FOR IMBALANCED LEARNING IN CCP

The Sampling-Based Stack Framework addresses the class imbalance problem in customer churn prediction (CCP) by integrating various sampling techniques with stacked ensemble models to improve prediction accuracy. However, the framework may suffer from over fitting and loss of important data patterns due to excessive sampling or improper selection of sampling methods

2.6 INTEGRATED CCP AND CSF FOR TELCO BUSINESS

The Integrated Customer Churn Prediction (CCP) and Critical Success Factors (CSF) framework for the Telco business focuses on combining churn prediction models with key business factors like customer satisfaction, service quality, and pricing strategies to improve customer retention. However, the approach may face challenges in accurately identifying and quantifying critical success factors, and its effectiveness may vary across different telecom environments due to dynamic customer behaviour.

3.PROPOSED SYSTEM

The proposed system for Customer Churn Prediction using RNN and LSTM Networks aims to address the limitations of existing systems by leveraging advanced deep learning models to analyze sequential customer behaviour data more effectively. Unlike traditional models utilizes Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, which are specifically designed to handle time-series data and capture long-term dependencies in customer behaviour. The system will be capable of processing large volumes of customer interaction data, including transaction history, usage patterns, customer service interactions, and demographic details, to generate accurate churn predictions.
The key feature of the proposed system is its ability to learn from sequential patterns in customer behavior, which allows it to predict churn based on both short-term and long-term customer interactions. This dynamic and adaptive approach enhances the predictive power of the model, making it suitable for industries with frequent and evolving customer engagements, such as telecommunications, e-commerce, and subscription-based services. Additionally, the system can be trained to detect subtle behavioural changes—such as reduced login frequency, changes in purchase habits, or negative support interactions—that may indicate churn risk before it becomes apparent.
To further improve prediction accuracy and real-time responsiveness, the proposed system can be integrated with streaming data platforms that support continuous model updates and live monitoring of customer activity. This enables the generation of timely alerts and actionable insights, empowering businesses to intervene proactively with personalized retention strategies. Moreover, the use of LSTM's memory cell architecture allows the model to retain crucial historical context over extended periods, providing deeper insights into long-term trends and lifecycle changes.
Incorporating attention mechanisms or hybrid models with Convolutional Neural Networks (CNNs) for feature extraction could further enhance performance by emphasizing critical time steps or patterns in the data. The system can also benefit from explainable AI techniques that provide transparency into the prediction process, helping stakeholders understand which behavioural features most influence churn probability.
Overall, the proposed system represents a significant advancement traditional churn prediction methods by introducing a more intelligent, context-aware, and time-sensitive approach to understanding .

4. CASE MODULE

In general, case can be defined as a course or principle of action adopted or proposed by an organization or individual.

4.1 DATA PREPROCESSING AND FEATURE ENGINEERING:

This module focuses on preparing the dataset for training the deep learning models. It involves collecting customer data, handling missing values, and normalizing features to ensure consistency. Feature engineering techniques such as encoding categorical variables, scaling numerical values, and creating time-based features are applied to enhance model performance. The dataset includes demographic details, transaction history, service usage, and customer interaction logs. Outlier detection methods help eliminate noise in the data, ensuring accurate predictions. Time-series transformations are implemented to structure data in a sequential format for RNN and LSTM models. The processed data is then split into training, validation, and test sets. Additionally, feature selection techniques are used to identify the most influential attributes for churn prediction. The final dataset serves as the foundation for training deep learning models effectively.

4.2 RECURRENT NEURAL NETWORK (RNN) MODEL

This module implements a basic Recurrent Neural Network (RNN) to analyze sequential customer behavior. The RNN model processes time-dependent features, learning patterns in customer interactions to predict churn probabilities. It consists of multiple recurrent layers that maintain past information while making predictions for future customer behavior. The model is trained using historical data, allowing it to recognize behavioral trends that lead to churn. To optimize performance, activation functions such as ReLU and tanh are used in different layers. Regularization techniques like dropout are applied to prevent overfitting and improve generalization. The model is evaluated using accuracy, precision, recall, and F1-score metrics. Hyperparameter tuning, such as adjusting the learning rate and batch size, is performed to enhance model efficiency. The results from the RNN model provide insights into customer engagement and early churn detection.

4.3 LONG SHORT-TERM MEMORY (LSTM) MODEL

The LSTM model builds upon RNN's capabilities by addressing the problem of long-term dependencies in sequential data. This module integrates LSTM layers, which consist of memory cells capable of retaining information over extended periods. Unlike traditional RNNs, LSTM networks mitigate the vanishing gradient problem, making them more effective for capturing complex customer behavior trends. The model is trained with past customer interactions, learning patterns that indicate potential churn. Techniques such as batch normalization and dropout are implemented to enhance model stability and prevent overfitting. The architecture includes multiple LSTM layers stacked with dense output layers for prediction. Performance is measured using AUC-ROC, accuracy, precision, and recall metrics. Hyperparameter tuning, including adjustments to the number of LSTM units and learning rate, helps refine the model. The LSTM model delivers improved accuracy in predicting customer churn by capturing both short-term and long-term dependencies.

4.4 MODEL EVALUATION AND OPTIMIZATION

This module focuses on assessing the performance of both the RNN and LSTM models to determine the most effective approach for churn prediction. Evaluation metrics such as accuracy, precision, recall, F1-score, and AUC-ROC are used to measure model effectiveness. A confusion matrix is generated to analyze false positives and false negatives in predictions. Cross-validation techniques ensure the model generalizes well to new customer data. Fine-tuning of hyperparameters, such as optimizer selection and learning rate adjustments, is performed to enhance model performance. Ensemble learning techniques may be explored to combine predictions from multiple models for improved accuracy. Additionally, explainability methods such as SHAP values and feature importance analysis provide insights into which factors contribute most to churn. The final optimized model is integrated into a business framework for real-time churn prediction and customer retention strategies.

Input Sequence Processing:
ht=f(Wh·ht−1+Wx·xt+b)h_t = f(W_h \cdot h_{t-1} + W_x \cdot x_t + b)

Recurrent Neural Networks (RNNs) are a class of artificial neural networks specifically designed to process sequential data, making them highly effective for applications that involve temporal or time-dependent patterns—such as natural language processing, speech recognition, and, notably, customer churn prediction. Unlike traditional feedforward neural networks, which treat all input data as independent and identically distributed, RNNs are capable of capturing temporal dependencies and contextual relationships between

inputs by maintaining a memory of previous inputs through their internal state.
The key architectural innovation of RNNs lies in their recursive structure, where connections between nodes form directed cycles. This looped architecture allows the network to retain information from prior time steps by passing the hidden state from one time step to the next. As a result, RNNs can maintain a form of short-term memory that is essential for tasks where the current input is influenced by past events. This makes them especially suitable for customer churn prediction, where customer behaviour is inherently sequential and evolves over time.
In the context of churn prediction, customer data is rarely static. Behavioural patterns, such as product usage frequency, login activity, transaction history, payment delays, and interactions with customer service, change over time and can provide early indicators of dissatisfaction or disengagement. RNNs can learn from these evolving sequences to recognize behavioural trends and recurring patterns that correlate with churn. For example, a gradual decrease in usage, increased complaints, or changes in purchasing habits can be sequentially analyzed by RNNs to estimate the probability that a customer is likely to churn in the near future.

5. ASSUMPTION FOR RNN:

Temporal Dependencies: RNNs work under the assumption that the order of data points matters, meaning that historical information plays a key role in predicting future outcomes. In churn prediction, this assumption is crucial as the model must learn how past behaviours or interactions influence future churn decisions.
Sequential Data: RNNs assume that the data points are sequential in nature. This means that the model relies on the sequence of inputs (e.g., customer interactions or service usage) to make predictions.
Short-term and Long-term Memory: RNNs assume that both short-term and long-term +dependencies in the data are important for accurate predictions. In churn prediction, it's vital to capture both immediate behaviour changes and more gradual shifts in customer engagement.

6.LONG SHORT – TERM MEMORY (LSTM)ALGORITHM

Long Short-Term Memory (LSTM) networks are a specialized type of Recurrent Neural Networks (RNNs) that are specifically designed to address the challenges faced by stan dard RNNs in learning and remembering long-term dependencies in sequential data. LSTM models have revolutionized the way we handle time-series and sequential data, which are prevalent in many domains, including customer churn prediction.

LSTM's primary advantage over traditional RNNs lies in its ability to selectively remember or forget information over long sequences. This is achieved by using a set of gates (forget gate, input gate, and output gate) that control the flow of information within the network. This makes LSTM networks highly effective in capturing complex temporal patterns, which is key when predicting customer churn, where the customer behaviour over time plays a significant role in determining whether or not they will leave a service .In the context of customer churn prediction, LSTM networks can analyses historical data such as customer transactions, interaction logs, and behaviour patterns to identify which customers are likely to churn based on the patterns over time.

6.1 ASSIMPTION FOR LSTM IN CCP.

Sequential Nature of Data: LSTM assumes that the data is sequential, meaning the order of events matters. For churn prediction, a customer's past behaviour (e.g., past interactions, purchases, or complaints) can influence future behaviour. Since churn prediction requires analyzing a sequence of customer actions over time, LSTM is an ideal model to capture these temporal dependencies.
Memory Requirement: LSTM assumes that the network needs to remember important information over long periods. This is crucial for churn prediction because certain events, even if they occurred months ago, might still influence a customer's decision to stay or leave.
Selectively Forgetting and Remembering Information:
Forget Gate:
The forget gate decides what information from the previous time step should be discarded from the cell's memory. It takes two inputs: the current input (x_t) and the previous hidden state (h_{t-1}).
Formula:
ft=σ(Wf·[ht−1,xt]+bf)f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)

Input Gate:
The input gate controls how much of the new information should be stored in the memory cell.
Formula:
it=σ(Wi·[ht−1,xt]+bi)i_t = \sigma(W_i \cdot [h_{t-1}, x_t] +b_i) Ct′=tanh(WC·[ht−1,xt]+bC)C'_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)

Cell State Update:
The cell state (C_t) is updated by combining the previous memory (C_{t-1}), the forget gate, and the input gate
Formula:
Ct=ft·Ct−1+it·Ct′C_t = f_t \cdot C_{t-1} + i_t \cdot C'_t

isrjournals

International Journal of Advanced Research in Computer Networking,Wireless and Mobile Communications

Volume: 5 Issue: 3 Apr,2025,ISSN_NO: 2320-7248

Adaptability to Evolving Customer Behaviour : Another important assumption when using LSTM models in churn prediction is that customer behaviour is the for dynamic and continuously evolving. Traditional models often assume stationarity in behavior patterns, but LSTM is capable of adapting to shifting trends over time due to its flexible learning mechanisms.This adaptability allows the model to recalibrate itself as new data becomes available,making it suitable for industries where customer preferences,market condition and service usage patterns change frequently.

7.EXPERIMENTAL RESULTS

This result discusses about the implementation of the policy-based security for various cases are identified and the below Fig 7.1 , Fig 7.2 , Fig 7.3 , Fig 7.4 , Fig 7.5 , Fig 7.6 , Fig 7.7 and Fig 7.8Shows the implementation of admin policy based on the proposed methodology.

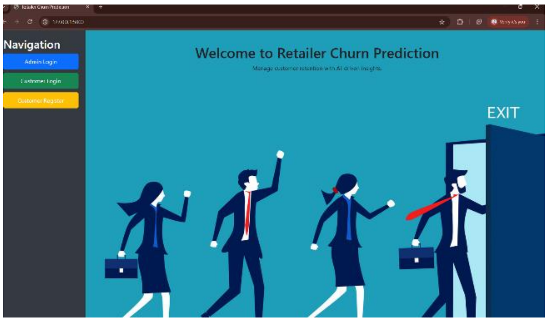

Fig 7.1 Welcome page

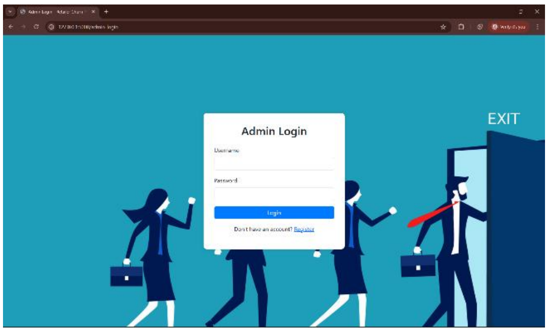

Fig 7.2 AdminLogin

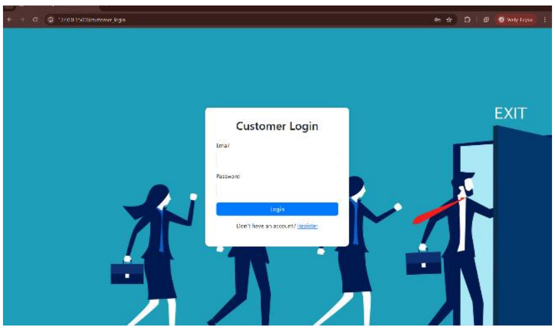

Fig 7.3 Customer Login

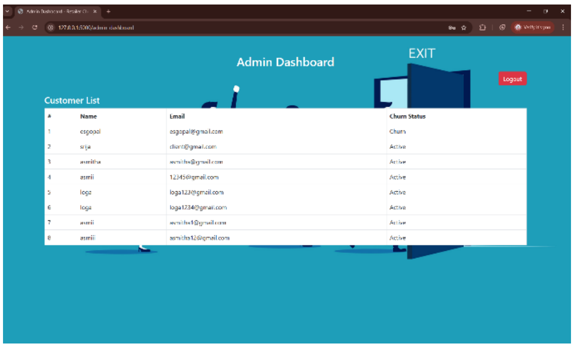

Fig 7.4 Admin Dashboard

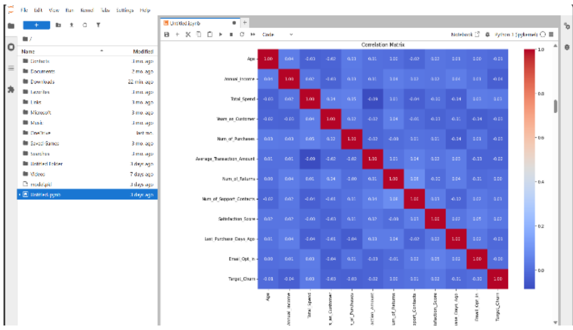

Fig 7.5 Correlation matrix

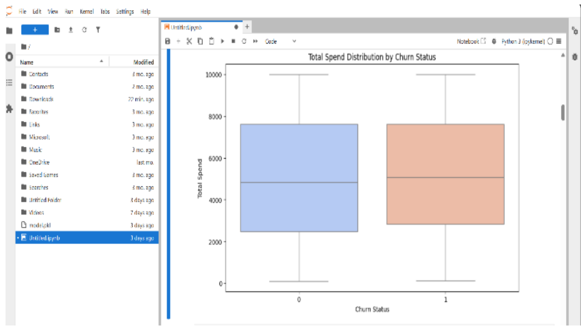

Fig 7.6 Churn Status

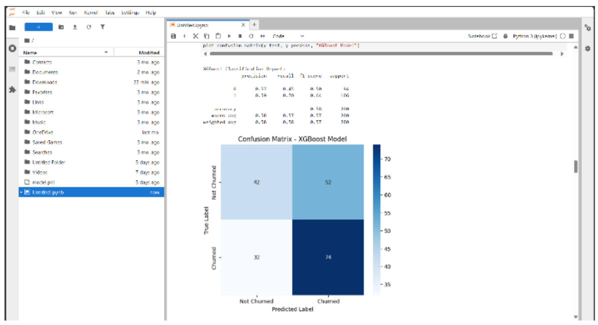

Fig 7.6 Predicted List

8.CONCLUSION & FUTURE WORK

The proposed system effectively leverages deep learning for accurate customer churn prediction, promoting proactive retention and sustainable business growth. Future enhancements can focus on integrating attention mechanisms, external data sources, and real-time learning pipelines. Additionally, incorporating explainable AI and transfer learning will boost model transparency and adaptability across diverse industries.

The integration of real-time data further boosts the system's adaptability to dynamic customer behaviours. The project proves that deep learning models can be effectively applied to churn prediction, offering businesses valuable insights to proactively identify at-risk customers and implement targeted retention strategies. As a result, the proposed system not only contributes to reducing churn but also promotes customer loyalty, providing businesses with a powerful tool for sustainable growth and improved decision- making.

REFERENCE

[1] S. Saha, C. Saha, M. M. Haque, M. G. R. Alam and A. Talukder, "Churn-Net: Deep Learning Enhanced Customer Churn Access, vol. 12, pp. 4471-4484,2024,doi: 10.1109/ACCESS.2024.3349950.

[2]N. Jajam, N. P. Challa, K. S. L. Prasanna and C. H. V. S. Deepthi, "Arithmetic Optimization With Ensemble Deep Learning SBLSTM-RNN-IGSA Model for Customer Churn Prediction," in IEEE Access, vol. 11, pp. 93111-93128, 2023, doi: 10.1109/ACCESS.2023.3304669.

[3]S. Wu, W. -C. Yau, T. -S. Ong and S. -C. Chong, "Integrated Churn Prediction and Customer Segmentation Framework for Telco Business," in IEEE Access, vol. 9, pp. 62118-62136, 2021, doi: 10.1109/ACCESS.2021.3073776.

[4]S. De and P. Prabu, "A Representation-Based Query Strategy to Derive Qualitative Features for Improved Churn Prediction," in IEEE Access, vol. 11, pp. 1213-1223, 2023, doi: 10.1109/ACCESS.2022.3233768.

[5]S. De and P. Prabu, "A Sampling-Based Stack Framework for Imbalanced Learning in Churn Prediction," in IEEE Access, vol. 10, pp. 68017-68028, 2022, doi: 10.1109/ACCESS.2022.3185227.

[6]M. J. Shabankareh, M. A. Shabankareh, A. Nazarian, A. Ranjbaran, and N. Seyyedamiri, "A stacking-based data mining solution to customer churn prediction," J. Relationship Marketing, vol. 21, no. 2, pp. 124–147, Apr. 2022, doi: 10.1080/15332667.2021.1889743.

[7]S. Saleh and S. Saha, "Customer retention and churn prediction in the telecommunication industry: A case study on a Danish university," Social Netw. Appl. Sci., vol. 5, no. 7, p. 173, Jul. 2023

[8]J. K. Sana, M. Z. Abedin, M. S. Rahman, and M. S. Rahman, "A novel customer churn prediction model for the telecommunication industry using data transformation methods and feature selection," PLoS ONE, vol. 17, no. 12, Dec. 2022, Art. no. e0278095.

[9]Y. Liu, M. Shengdong, G. Jijian, and N. Nedjah, "Intelligent prediction of customer churn with a fused attentional deep learning model," Mathematics, vol. 10, no. 24, p. 4733, Dec. 2022.

[10] F. E. Usman-Hamza, A. O. Balogun, L. F. Capretz, H. A. Mojeed, S. Mahamad, SSalihu, A. G. Akintola, S. Basri, R. T. Amosa, and N. K. Salahdeen, "Intelligent decision forest models for customer churn prediction," Appl. Sci., vol. 12, no. 16, p. 8270, Aug. 2022.

[11] Somak Saha , Chamak Saha , Md. Mahidul haque Md. Golam Rabiul Alam and Ashis Talukder ChurnNet: Deep Learning Enhanced Customer Churn Prediction in Telecommunication Industry may 2024.

[12] Awais Manzoor 1,2,3, M. Atif Qureshi 1,2, Etain Kidney2, And Luca Longo 3A Review on Machine Learning Methods for Customer Churn Prediction and Recommendations for Business Practitioners 2024.

[13] Shuli Wu , Wei-Chuen Yau , Thian-Song Ong , And Siew-Chin Chong. Integrated Churn Prediction And Customer Segmentation Framework For Telco Business: 2021