Adaptable Churn Prediction Pipeline for Hybrid Business Model using Deep Neural Networks and Gradient Boosting

Abstract. In the dynamic landscape of hybrid businesses, where contractual B2B-like and B2C-like customer relationships converge, the demand for effective churn prediction mechanisms is crucial. This study presents a predictive pipeline on real-life data, employing Deep Neural Network (NN) and Gradient Boosting (XGBoost)-based models for an Industrial Computer Business with such characteristics. Our findings demonstrate that the NN approach showed the ability to make more logical decisions. This claim is supported by utilizing the Euclidean distance metric, specifically applied to measure the similarity between incorrectly predicted instances and those belonging to the misclassified class. Moreover, to address a possible class imbalance scenario, we implement under-sampling with Tomek Links, improving the model's robustness further. The developed models show impressive performance as the mean F1 score, calculated across multiple years, reaches an impressive 0.85, surpassing the 0.80 accuracy threshold. Furthermore, our deployment strategy offers a real-world example of integrating Robotic Process Automation (RPA) and AI. This approach not only showcases the synergy between RPA and AI in building a digital solution but also serves as a practical blueprint for incorporating the adaptability of predictive models in dynamic business environments.

Keywords: churn prediction, deep learning, gradient boosting

1 Introduction

Customer retention and churn management has been widely recognized by many organizations [1], operating in various industries and providing services or products to many types of customers. Churn management is one of the most influential factors in the growth and success of a business as it not only leads to a decrease in sales revenue but also, the acquisition of new customers is usually more costly than retaining the existing customers [2]. Thus, accurate and consistent prediction of potential churn behavior could assist in prevention, by taking appropriate measures.

Customer churn behavior should be handled according to the unique characteristics of the business organization. For instance, in business to business (B2B) context, the quantity of customers is often less and the quantity of transactions is higher compared to business-to-customer (B2C) business models. In this respect, customer churn may imply the termination of a well-defined relationship

[3], bound by a contract, which in turn would be in favor of the potential competitors, and the impact of losing a customer is much higher than the B2C case.

The advancements in database management systems and Machine Learning (ML) methods, allowed for the application of these techniques to make predictions and gather insight regarding the potential churn behavior of existing customers. The existing literature combines many studies and approaches to the prediction of the potential churn behavior. Data mining techniques [4], including methods such as Support Vector Machines (SVM) [4][5], and Decision Trees, as well as Clustering[6], have been employed in both B2C and B2B scenarios.

The increasing activities in e-commerce and online sales platforms have led to "*hybrid*" business characteristics where some customers' behavior appears similar to the B2C scenario, depending on the relationship of the business with the customer. In this paper, we aim to present the methodology for churn behavior prediction in such "hybrid" businesses, by building a pipeline for churn prediction for an Industrial Computer enterprise that operates as B2B a business on contractual and non-contractual level and also offers e-commerce to a wide range of individuals/organizations. During the process, we employ, boosting and neural-network-based predictive models on real-life data gathered from the organization. Additionally, we provide an in-depth evaluation and validation pipeline of the obtained results, as well as a real-life deployment solution that integrates many means including Robotic Process Automation and Business Intelligence Analysis to perform adaptive learning processes and provide updated predictions. This paper is organized as follows: We give an insight into the relevant work and used methodologies in Section 2, and the methodology and evaluation strategy are presented in Section 3. respectively. The obtained results are shown in Section 4 and the current deployment solution is discussed in Section 5. Finally, we discuss the obtained results, addressing possible improvements and limitations in Section 6, and draw our conclusions in Section 7.

2 Literature Review

The development of accurate and effective customer churn prediction frameworks is crucial for businesses to retain their valuable customer base and reduce the financial burden associated with customer disengagement. The existing literature reveals many studies, carried out in different business settings although B2C studies remain the majority portion of them [7].

Several recent studies have delved into different Machine Learning methods to effectively predict churn behavior in B2C settings and many different business domains. Xiaohou and Yarada [4] proposed an approach that incorporates clustering and classification methodologies, by employing K-Means and SVM to identify customer segments that are most susceptible to disengage and forecast individual customer churn behavior. The data utilized was retrieved from one of the largest e-commerce platforms. Their proposed method showed an impressive result, with the Area Under the ROC Curve (AUC-ROC) score of 0.92. Moreover, Shobana et al.[8] added a business intelligence-based strategy on top of

ML methods that leverages various classification algorithms, including logistic regression and decision trees. Additionally, Wu et.al [9] point out the scenario where the churn rate is significantly high in an e-commerce business, which causes class imbalance during training. Therefore, they propose an ADABoost model extended with SMOTE oversampling to overcome the raised issue and they yielded impressive results.

While the aforementioned studies focused on B2C settings, the dynamics of churn behavior in B2B relationships exhibit distinct characteristics, which are investigated separately. Chen et.al.[5] developed a method to predict churn from valuable B2B customers in the logistics industry employing frequent pattern mining, customer value modeling, and a support vector machine (SVM) classifier, achieving high performance in identifying high-value churning customers with an F1-Score of 0.90. Jamjoom [10] employed a Cross-Industry Standard Process for Data Mining (CRISP-DM) knowledge extraction approach based on information gain and entropy to identify relevant customer behavior patterns associated with churn, for a health insurance company. This approach achieved an accuracy of 90.32%. Itschert et.al.[11] propose a Random Forest approach and delve deeper into understanding customer lifetime value (CLV)-related features affecting the churn, as well as conducting a field experiment, utilizing the churn rate metric in the non-contractual B2B wholesale industry. In a similar setting, Mirkovic et.al [6] introduce leveraging invoice-level data to build effective binary classification models for customer churn forecasting.

The most recent ML-powered churn prediction studies indicated significant potential for spotting potential churn behavior in advance in many different domains and settings. This study will focus on the said hybrid setting where the organization of interest has both contractual and non-contractual customers as well as e-commerce customers that carry the characteristics of a B2C customer.

3 Model Development

3.1 Data Processing

Data Preparation The data is retrieved from the sales transaction database of an Industrial Computer organization, which mainly operates on the B2B level. The dataset contains sales orders of customers, starting from January 2019 until May 2023. The business organization has customers who are bound by contract for a fixed term, as well as B2C-like customers who mostly make purchases through the online sales platform of the organization. Given the customer types, Key Account and Channel Partner customers are the key drivers of the sales revenue. The different types of customers are distinguished from their market sector.

The dataset incorporates fundamental transaction information, such as the purchase date, quantity, paid amount, and other miscellaneous information regarding the part, shipment type, and salespeople and sector. Effective customer churn prediction relies heavily on the quality and relevance of the attributes used to model customer behavior. In this respect, the Recency, Frequency, and Monetary (RFM) analysis was utilized to transform the data into a more meaningful representation in the context of the application followed by several preprocessing steps.

First of all, to address the presence of outliers, and data points that deviate significantly concerning the monetary value, the IQR (interquartile range) method was employed. Outliers were eliminated to ensure an accurate representation of the underlying patterns and relationships.

Additionally, to ensure that numerical features had comparable scales and were treated equally by the employed algorithms, z-score normalization was applied. Z-score normalization involves subtracting the mean and dividing it by the standard deviation of each feature.

As part of the data preparation process, categorical variables were applied one-hot encoding to facilitate their integration into the predictive model. This transformation was applicable as the included categorical variables did not have high numbers of value counts, not causing the dimensionality of data to increase drastically.

Labeling Strategy The labeling process for ground truth generation relied on transactional data provided by the domain experts. We analyzed customers' purchase history to identify patterns indicative of customer attrition. To this end, each customer instance in the dataset was assigned a 'churn' label under the following conditions:

A customer was considered to have churned (‘Churn = 1‘) if the time gap between their most recent purchase date and the preceding purchase date exceeded 365 days. This threshold was determined considering that this period of inactivity is substantially longer than the typical purchase cycle in the context of the distinct characteristics of the business.

Conversely, if the interval between the last two purchase dates was 365 days or less, the customer was considered 'active' (‘Churn = 0‘).

Figure 1 illustrates the class distribution across three consecutive years, resulting from the labeling process. It is seen that given the class counts in the year 2021, class imbalance is possible, this will be explored further in the later sections.

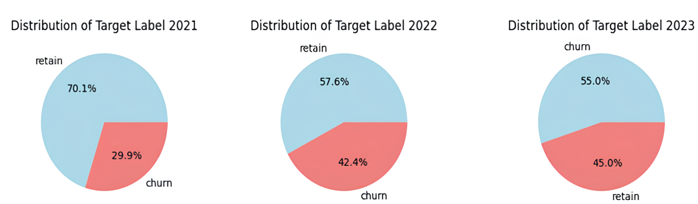

Fig. 1. The class distribution across 3 different years

RFM Score RFM model is a behavior-based model used to investigate customer behavior [12]. It consists of three main components, namely the Monetary, which denotes the money spent by the customer in a specific period, Recency which represents the time elapsed since the last purchase, and Frequency, which denotes the number of purchases made by the customer[13].

RFM scoring is widely applied in database-centered marketing and business intelligence applications, as it allows for a prior segmentation of customers, based on their purchase behavior [14]. The scoring scale in RFM framework is mostly formulated as :

$$Recency\_score \times 100 + Frequency\_score \times 10 + Monetary\_score \times 1 \qquad (1)$$

and mostly range between 111 and 555, where 111 is the lowest possible and 555 is the highest possible RFM score.

The extraction of the RFM score from the given dataset is executed using the fields order date, amount, and the total count of transactions per unique customer. Recency was simply calculated by subtracting the latest order date of the customer from the current date, monetary and frequency were computed by obtaining some of the total amount spent and the total number of purchases per customer respectively. Finally, to yield the scores of each indicator, the customers were grouped into equally sized bins based on quantiles and were assigned the individual R,F, and M scores. The RFM score distribution in churning and non-churning customers is shown in Figure 2

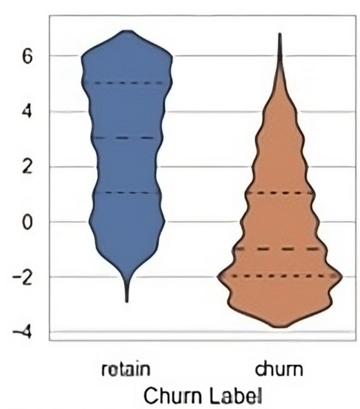

Fig. 2. RFM Score distribution in data training

3.2 Feature Selection

To streamline the predictive modeling process for customer churn, a systematic feature selection methodology was employed to reduce data dimensionality and retain the most informative attributes. A visual representation of the comprehensive feature selection steps can be found in Figure 3. In the initial dataset comprising a total of 52 features, several steps were taken to exclude irrelevant columns. For instance, details such as 'Salesperson Information',' Shipment Date',' Record Id', and 'Order Number' were deemed non-contributory to the nature of the churn prediction task as they provide minimum information regarding the purchase behavior of the instances, and were therefore excluded. Additionally, variables with high correlation or very small variance were pruned to ensure a refined dataset.

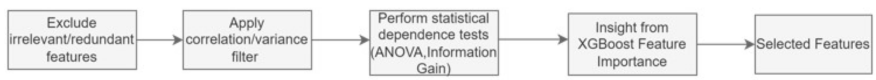

Fig. 3. The feature selection process

The selected features were not only chosen for their informativeness in the context of the churn prediction problem but have been also affirmed relevant by dependence tests and importance insight. The remaining 18 attributes underwent further evaluation based on their statistical dependence on the target label, which signifies potential churn. This analysis included both ANOVA and Information Gain analysis to measure statistical dependence. By utilizing the inherently interpretable XGBoost Model, this analysis was augmented by insights into feature importance. Features indicating the customer type, cumulative quantity, and those derived from Recency-Frequency-Monetary (RFM) analysis, as well as the attributes indicating the behavioral type of the customer (B2B or B2C-like), were prioritized in the training process, except Recency which had a significantly high correlation with the RFM Score.

3.3 Model Selection and Training

To construct a robust customer churn prediction model, the adopted approach commenced with the utilization of XGBoost, a gradient-boosting framework renowned for its interpretability and effectiveness in handling structured data [15]. Another experiment was conducted through the utilization of a neural-network-based approach, achieved by training a Multi-Layer Perceptron (MLP) model. The training process involved careful tuning of the hyper-parameters of each model.

The partitioning of the train and test set was carried out such that the reliability of the model persists for future periods, which is crucial for real-life application scenarios. The training set was strategically composed of data until

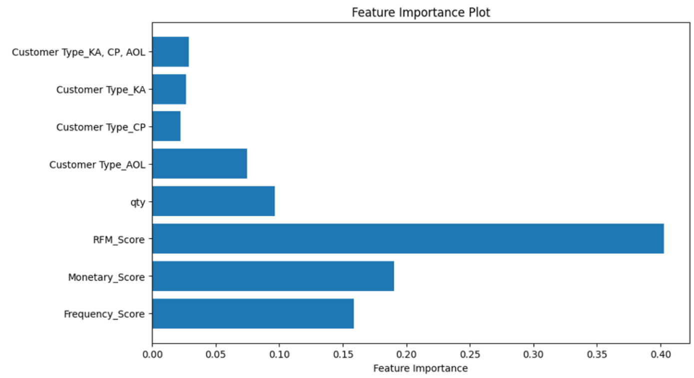

Fig. 4. Feature Importance Analysis

the year 2022, and further split into 5 train and validation sets during the training process, according to the 5-fold Cross Validation scheme. The remainder data points which span from the end of 2022 until May 2023 were kept as a test set, which was then used to test the models' performance on "unseen" data.

Deep Neural Network Approach A particular type of Deep Neural Network, namely, Multi-Layer Perceptron (MLP) was adopted. Unlike traditional algorithms, DNN approach is capable of capturing intricate nonlinear relationships within the data through its layered architecture [16]. To facilitate binary classification of customer churn, the final layer of the neural network was configured with a sigmoid activation function [17]. This choice of activation function is well-suited for binary classification tasks, as it transforms the model's output into a probability score between 0 and 1. The sigmoid activation function enables the interpretation of the model's predictions as the likelihood of a customer churning. The MLP architecture involved two hidden layers with 64 and 32 units. The model was trained for 10-15 epochs with early stopping and with a batch size of 32.

XGBoost Algorithm Extreme Gradient Boosting (XGBoost) is an ensemble learning algorithm that makes use of boosting. It is widely used in many classification tasks to achieve state-of-the-art results [15]. It implements the Gradient Boosted Decision Trees, where the algorithm generates a single decision tree, and then continues to build more decision trees sequentially. During the training process, the weight, based on the residual error of the incorrectly predicted parameters by the previous tree is increased. Next, parameters with the highest weights are added to the next decision tree generated. XGBoost model also provides an in-depth feature importance analysis to demonstrate the relevance of each variable used in training. The hyperparameters included *maximum_depth* = 5, *learning_rate*= 0.1 and *n_estimators*=100, where each parameter was systematically tuned to balance model complexity and generalization capability of the model.

Hyperparameter Optimization In pursuit of optimizing the performance of our customer churn prediction models, we employed the Grid Search algorithm for hyperparameter optimization. The algorithm operates by creating a grid of all possible combinations of hyperparameter values, and for each combination, it conducts cross-validation to assess model performance [18]. In this specific setting, each configuration was evaluated by the algorithm using a 5-fold cross-validation strategy within the training dataset. F1 score was chosen as the primary evaluation metric.

For the XGBoost model,various hyperparameter combinations were explored, including *n_estimators* $\{100, 200, 300\}$, *learning_rate* $\{0.01, 0.1, 0.2\}$, and *max_depth* $\{3, 4, 5, 10, 15\}$.

Similarly, for the MLP model, diverse hyperparameter sets were evaluated, including *hidden_layer_sizes* $\{(64, 32), (64, 64, 32), (128, 64), (128, 64, 32)\}$, and *learning_rate* values $\{0.0001, 0.001, 0.01\}$. This exploration allowed for the identification of the optimal configuration for each model, ensuring robust and well-tuned predictive performance.

Handling Class Imbalance In addressing the inherent challenge of class imbalance within the dataset, we implemented a strategic approach utilizing Tomek Links. Tomek Links is a method employed in pattern recognition and machine learning to improve the discrimination of the minority class. They operate by identifying pairs of instances—one from the majority class and one from the minority class—that are close to each other but of opposite classes. These instances, known as Tomek pairs, are then selectively removed to create a clearer separation between the classes [19].

In our deployment strategy, Tomek Links are applied dynamically, triggered by the model retraining process at specified intervals. If, during a retraining cycle, the dataset indicates an imbalance where the majority class constitutes more than 0.60 of the total instances, Tomek Links are then employed for undersampling. This ensures that the model is exposed to a more balanced dataset, promoting improved generalization and mitigating potential biases introduced by class imbalance.

Evaluation Metrics The effectiveness of the developed customer churn prediction models was assessed through a comprehensive evaluation, employing precision, recall, and F1 score as key performance metrics. The evaluation was performed both on validation and test sets.

3.4 Experimental Setup

The proposed customer churn prediction model was trained on consecutive yearly data. The motivation for testing the model on different years was twofold. Firstly, it is aimed to evaluate the model's consistency and generalizability over time. Secondly, our retraining approach accounted for the dynamic nature of real-world data, where new information is continuously appended. Simulation of retraining the model with the addition of new data allowed for capturing evolving patterns and ensured the model's relevance.

In this respect, the model was on the 2019-2020 data and tested its performance on the unseen 2021 test set. Subsequently, we extended the training data to 2019-2021 and predicted churn for 2022, repeating this process until 2023, as the data set spans from 2019 to 2023. Therefore the test set for each iteration comprised the data for the subsequent year.

Finally, to gain deeper insights into the nature of incorrectly predicted instances, an additional analysis was conducted. The instances misclassified by the models were isolated, and the Euclidean distance between these misclassified instances and those correctly predicted from the same class was computed. This distance measurement was performed in a reduced feature space using two principal components.

4 Performance Evaluation

4.1 Model Comparison

Our experimentation consistently revealed robust performance during the model training phases, with the validation set results consistently achieving high scores (above an F1-Score of 0.85) across all years and models. While these encouraging validation set scores provided initial confidence in our models, the true crucial test came with the evaluation of unseen test sets, where the predictive models faced the challenges of real use case dynamics and evolving patterns. The results yielded from both of the proposed models over different years are demonstrated in Table 1.

The Table 1 presents the evaluation results of the XGBoost and DNN models for customer churn prediction over three consecutive years (2021, 2022, and 2023). For the XGBoost model, we observe a consistent performance across the years, with F1 scores ranging from 0.79 to 0.81. Similarly, the DNN model exhibits stability, with F1 scores of 0.79 to 0.84. Notably, the DNN model achieves the highest F1 score of 0.84 in 2023, demonstrating its improved performance. These findings offer valuable insights into the models' adaptability and suggest the potential for enhanced performance, especially in the case of the DNN model, which shows promising advancements in the latest prediction year.

4.2 Decision Making Quality

The visual analysis of the graph, illustrated in Figure 5, reveals interesting insights into the distribution of misclassified instances in the reduced feature space for the DNN and XGBoost models.

For the DNN Model, the tight clustering of misclassified instances around the origin suggests a distinct pattern shared with instances from the incorrectly

	Prediction Year	F1 Score	REcall	Precision
XGBoost	2023	0.81	0.80	0.82
	2022	0.80	0.79	81
	2021	0.79	0.79	0.80
DNN	**2023**	**0.84**	**0.85**	**0.83**
	2022	0.82	0.83	0.81
	2021	0.79	0.79	0.80

Table 1. Comparison of XGBoost and Deep Neural Network Model

predicted class. This concentrated proximity implies recognizable similarities in captured features. On the other hand, The scattered distribution of misclassified instances by the XGBoost model indicates a broader range of similarities with instances from the incorrectly predicted class. The varied dispersion suggests a diverse set of features contributing to false positives or negatives.

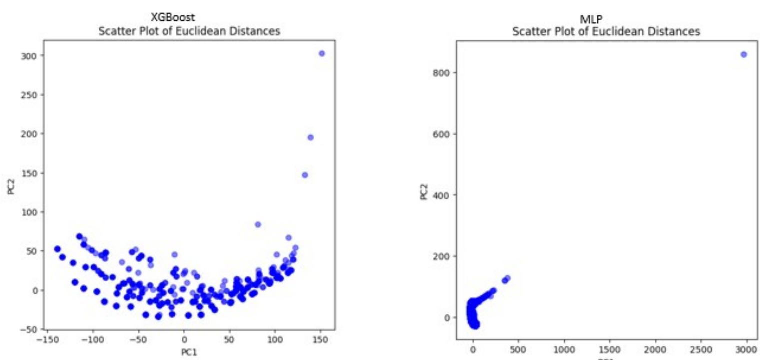

Fig. 5. Misclassification analysis using Euclidean distance

4.3 Impact of Balancing

It is also seen that the impact of balancing in the case of class imbalance is significant when it comes to achieving robustness and minimizing bias. Table 2 displays the results obtained with the original imbalanced dataset and another with a balanced dataset achieved through Tomek Links undersampling. The results, obtained by using the balanced dataset for training, significantly outperform the original imbalanced data, which is an indication that the alleviating model bias due to the class imbalance yields a more robust predictive model.

Results for 2021	F1 Score	Recall	Precision
Imbalanced	0.57	0,70	40.9
Balanced	**0.79**	**0.79**	**0.80**

Table 2. Comparison of Balanced and Imbalanced Data Impact on Results for 2021

5 Deployment Strategy

5.1 Adaptive Learning Process

The demonstrated performance and robustness of our framework initiate an exploration of its potential impact on strategic decision-making in practical business applications. In this respect, the developed predictive model should be aware of the recent shifts in behavioral patterns and ensure continuous relevance, rather than being isolated from the latest changes [20]. To achieve this, we implement an adaptive learning approach, which involves retraining the model at certain time intervals. This proactive strategy allows the model to stay in alignment with the evolving trends and changes in customer behavior, ensuring it remains a reliable tool for strategic decision-making.

Alongside the adaptive learning aspect, the predictions made by the model should also be updated frequently as new data becomes available. Therefore, we incorporate dynamic predictions which are executed on a daily basis. The model analyzes the most recent sales transactions stored in the database to generate up-to-date predictions regarding customer churn behavior.

In light of the demonstrated performance and adaptability of our predictive framework, the attention turns towards its practical implementation in business contexts. The imperatives of strategic decision-making necessitate an integrated approach where the predictive model not only remains cognizant of evolving behavioral patterns but also aligns with operational needs.

5.2 Integrated Solution

The integration of our customer churn prediction model into real-world business operations is facilitated through multiple intermediate means, namely Robotic Process Automation, and the Business Intelligence (BI) Reports, leveraging the existing infrastructure of the business organization. RPA acts as a bridge, connecting the predictive power of our model with existing operational workflows, and the BI reports serve as the end-user interface, where the aimed users can view the prediction results.

Deployed daily, the model analyzes the latest sales transactions stored in the database to generate predictions on customer churn probabilities, the updating and model retraining operations are carried out by the integrated RPA workflow. These predictions are then utilized to craft comprehensive BI Reports, acting as a decision-support mechanism. They provide insight and results to domain experts on evolving customer behaviors, allowing them to take preventive action. Figure 6 visualizes how the developed model and the utilized means orchestrate to comprise the real-world solution.

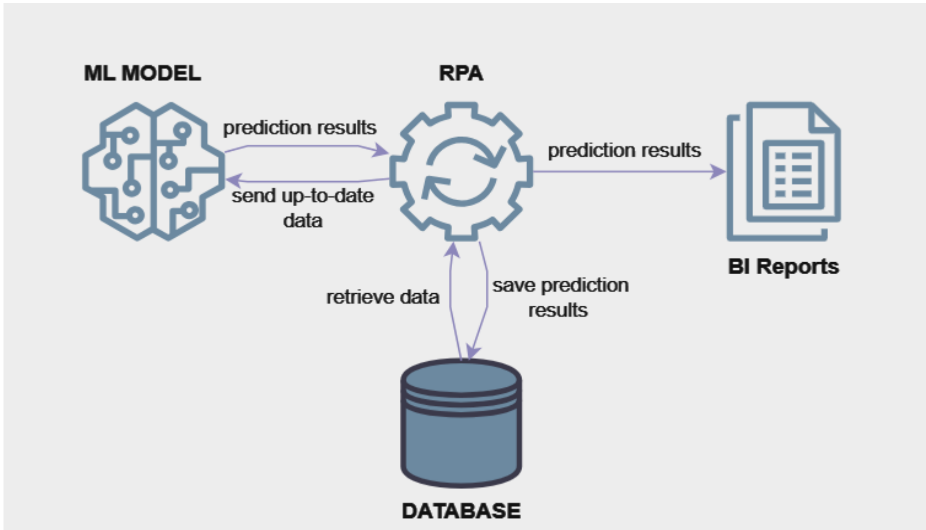

Fig. 6. Integrated Solution

6 Discussion

The comparative analysis among the Deep Neural Network and XGBoost models in predicting customer churn yields insightful observations. While both models exhibit reasonable and predictive performance, the DNN outperforms XGBoost in its ability to detect intricate patterns within the data, particularly when faced with instances that carry attributes contributing to misclassifications. This highlights the advantage of neural network architectures in capturing non-linear relationships that are usually found in complex datasets. In this respect, the capabilities and superior performance of DNN align with the demands of our customer churn prediction task, and thus, it was chosen for our deployment.

However, it is crucial to acknowledge the interpretability trade-off. DNN is a black-box model; it lacks the transparency aspect unlike XGBoost, which provides insights into the key features influencing model decisions. Finding a balance between model accuracy and interpretability remains a critical consideration in the deployment of customer churn prediction frameworks.

In addition, the graphical analysis depicted in Figure 5 indicates a trend shared by both models: instances misclassified by each model often bear similarities to instances from the incorrectly predicted class. This emphasizes the complexity of the classification task in churn prediction, highlighting the challenges in distinguishing subtle and slight variations in the dataset.

In terms of a possible class imbalance scenario, like the year 2021, the obtained results show that the application of Tomek Links under-sampling ensured the minimization of the bias occurring due to the dominating proportion of the majority class, and demonstrated a significant increase in predictive performance. On the other hand, it is important to take the loss of data due to under-sampling into account, as it causes the quantity of the instances available to drop. In this respect, this method is only performed when the class imbalance is significantly apparent.

A significant consideration in our evaluation is the presence of unpredicted churn instances, stemming from unforeseen or unexpected factors not captured by the model features. This raises the importance of continued model refinement and adaptation to evolving customer behaviors and market dynamics.

It is also important to mention that there is still room for further improvement. While our ground truth generation strategy effectively utilized transactional data to identify churn patterns, the absence of pre-labeled records of churned and retained customers posed a challenge. In an ideal scenario, preexisting records could have expedited the labeling process and potentially improved model precision. Moreover, the integration of Customer Relationship Management (CRM) data is a potentially promising enhancement, particularly in scenarios involving contractual relationships, as it contains information related to customer interactions, engagement history, and personalized services.

7 Conclusion

In the fluid landscape of contemporary business, where the fusion of contractual B2B-like and B2C-like customer relationships are present, our proposed pipeline is aimed at the prediction of churn behavior, especially for an Industrial Computer business with such characteristics. The results obtained showed that our developed model could be an efficient decision support tool that would offer a strategic point of view for businesses to overcome the complexities of customer relations in hybrid businesses and elevate them, enabling them to forecast the potential revenue that can be saved by retention of possible at-risk customers.

References

1. Suphan Nasir. Customer retention strategies and customer loyalty, 01 2017.
2. Ali Tamaddoni Jahromi, Stanislav Stakhovych, and Michael Ewing. Managing b2b customer churn, retention and profitability. *Industrial Marketing Management*, 43:1258–1268, 10 2014.
3. Hans Risselada, Peter C. Verhoef, and Tammo H.A. Bijmolt. Staying power of churn prediction models. *Journal of Interactive Marketing*, 24:198–208, 08 2010.
4. Xiancheng Xiahou and Yoshio Harada. B2c e-commerce customer churn prediction based on k-means and svm. *Journal of Theoretical and Applied Electronic Commerce Research*, 17:458–475, 04 2022.
5. Kuanchin Chen, Ya-Han Hu, and Yi-Cheng Hsieh. Predicting customer churn from valuable b2b customers in the logistics industry: a case study. *Information Systems and e-Business Management*, 13:475–494, 10 2014.
6. Milan Mirkovic, Teodora Lolic, Darko Stefanovic, Andras Anderla, and Danijela Gracanin. Customer churn prediction in b2b non-contractual business settings using invoice data. *Applied Sciences*, 12:5001, 05 2022.

7. Iris Figalist, Christoph Elsner, Jan Bosch, and Helena Olsson. Customer churn prediction in b2b contexts. *International Conference on Software Business 2019*, 01 2019.
8. Shobana J, Ch. Gangadhar, Rakesh Kumar Arora, P.N. Renjith, J. Bamini, and Yugendra devidas Chincholkar. E-commerce customer churn prevention using machine learning-based business intelligence strategy. *Measurement: Sensors*, 27:100728, 06 2023.
9. Xiaojun Wu and Sufang Meng. E-commerce customer churn prediction based on improved smote and adaboost, 06 2016.
10. Arwa A. Jamjoom. The use of knowledge extraction in predicting customer churn in b2b. *Journal of Big Data*, 8, 08 2021.
11. Theresa Gattermann-Itschert and Ulrich W. Thonemann. Proactive customer retention management in a non-contractual b2b setting based on churn prediction with random forests. *Industrial Marketing Management*, 107:134–147, 11 2022.
12. Jo-Ting Wei, Shih-Yen Lin, and Hsin-Hung Wu. A review of the application of rfm model. *African Journal of Business Management*, 4:4199–4206, 2010.
13. Chih-Hsuan Wang. Apply robust segmentation to the service industry using kernel induced fuzzy clustering techniques. *Expert Systems with Applications*, 37:8395–8400, 12 2010.
14. J Miglautsch. Application of rfm principles: What to do with 1–1–1 customers? *Journal of Database Marketing and Customer Strategy Management*, 9:319–324, 07 2002.
15. Dong Lan Zou. The research on the boosting decision tree algorithm for intelligent medical system. *Applied Mechanics and Materials*, 539:365–368, 07 2014.
16. Anders Krogh. What are artificial neural networks? *Nature Biotechnology*, 26:195–197, 02 2008.
17. Shrisha Bharadwaj, BS Anil, and PS Gowra. Customer churn prediction in mobile networks using logistic regression and multilayer perceptron(mlp) — ieee conference publication — ieee xplore, 2018.
18. Petro Liashchynskyi and Pavlo Liashchynskyi. Grid search, random search, genetic algorithm: A big comparison for nas. *arXiv.org*, 12 2019.
19. Elsie Fezeka Swana, Wesley Doorsamy, and Pitshou Bokoro. Tomek link and smote approaches for machine fault classification with an imbalanced dataset. *Sensors*, 22:3246, 01 2022.
20. Bing Liu. Lifelong machine learning: a paradigm for continuous learning. *Frontiers of Computer Science*, 11:359–361, 06 2017.