# Predicting B2B Customer Churn and Measuring the Impact of Machine Learning-Based Retention Strategies

Victória Emanuela Alves Oliveira[2,3][a], Amanda Cristina da Costa Guimarães[3][b],
Arthur Rodrigues Soares de Quadros[1,3][c], Reynold Navarro Mazo[3,4][d],
Rickson Livio de Souza Gaspar[3], Alessandro Vieira[3][e] and Wladmir Cardoso Brandão[1,3][f]

[1]*Department of Computer Science, Pontifical Catholic University of Minas Gerais (PUC Minas), Brazil*

[2]*Center for Exact and Technological Sciences, Federal University of Recôncavo da Bahia (UFRB), Cruz das Almas, BA, Brazil*

[3]*Data Science Laboratory (SOLAB), Sólides S.A. - Belo Horizonte, MG, Brazil*

[4]*Institute of Mathematical and Computer Sciences (ICMC), University of São Paulo (USP), São Carlos, SP, Brazil*

Keywords: Customer Churn, Churn Prediction, B2B Churn, Support Vector Machines (SVM), Machine Learning, Feature Selection.

Abstract: Acquiring new customers often costs five times more than retaining existing ones. Customer churn significantly threatens B2B companies, causing revenue loss and reduced market share. Analyzing historical customer data, including frequency on product usage, allow us to predict churn and implement timely retention strategies to prevent this loss. We propose using Support Vector Machines (SVMs) to predict at-risk customers while retraining it, if necessary. By monitoring its recall over 15-day periods, we retrain the model if its recall on new data falls below 60%. Our research focuses on feature selection to identify key churn factors. Our experiments show that when constantly retraining the model, we avoid accuracy loss by updating the customer's context, providing valuable insights on how to reduce churn rates and increase revenue.

## 1 INTRODUCTION

In today's dynamic business environment, companies often prioritize switching partners over building strong relationships (Tamaddoni Jahromi et al., 2014). To avoid revenue loss from customer churn, B2B companies need predictive methods. Analyzing past churn data and product usage can help predict and prevent future churn, protecting revenue.

Churn prediction allows targeted customer strategies. By analyzing recent product usage and comparing it to similar companies that experienced churn, models can identify warning signs and proactively prevent attrition.

Existing research has explored various automated customer retention methods. However, a crucial gap remains in understanding the temporal dynamics of churn prediction models. As customer needs and product demands evolve (Li et al., 2023), the underlying patterns driving churn may also shift. Consequently, static churn prediction models, trained on historical data, may lose their effectiveness over time. Our hypothesis is that, to maintain predictive accuracy, churn models must be dynamically updated to reflect these evolving customer behaviors and product requirements.

This paper uses Support Vector Machines (SVMs) on historical churn data to predict customer churn in 15-day intervals, using strategic feature selection. Every two weeks, the predictions are re-evaluated based and classified as correct or incorrect. This reduces churn and measures the revenue from strategies for at-risk customers. The dataset used in this study comprises over 8,000 Small and Medium Enterprises (SMEs) in the context of HR services. Although reactive methods allow for some level of customer retention, the dynamic updates on churn predictors provide valuable insights to B2B companies, helping businesses take proactive retention strategies. Namely, we contribute by:

- Applying SVMs to predict customer churn with dynamic data and updated models.

[a] https://orcid.org/0009-0000-2777-4581
[b] https://orcid.org/0009-0007-0321-3334
[c] https://orcid.org/0009-0004-9593-7601
[d] https://orcid.org/0009-0003-2011-3715
[e] https://orcid.org/0000-0002-9921-3588
[f] https://orcid.org/0000-0002-1523-1616



Oliveira, V. E. A., Guimarães, A. C. C., Soares de Quadros, A. R., Mazo, R. N., Gaspar, R. L. S., Vieira, A. and Brandão, W. C.
Predicting B2B Customer Churn and Measuring the Impact of Machine Learning-Based Retention Strategies.
DOI: 10.5220/0013436300003929
In *Proceedings of the 27th International Conference on Enterprise Information Systems (ICEIS 2025) - Volume 1*, pages 572-581
ISBN: 978-989-758-749-8; ISSN: 2184-4992

- Demonstrating the performance degradation of static churn prediction models over time, highlighting the need for dynamic updates.
- Using feature selection to identify key features for churn prediction, improving customer retention strategies.

This paper has five sections. Section 2 introduces customer churn, supervised machine learning, and the models used: SVM, Logistic Regression, and Sequential Feature Selection. Section 3 discusses related work. Section 4 details the methodology: data preprocessing, feature selection, model architecture, and evaluation metrics. Section 5 presents and analyzes the results. Section 6 concludes and suggests future research.

# 2 BACKGROUND

This section defines key concepts used in this paper, including Customer Churn, Machine Learning Algorithms, Support Vector Machines, and Feature Selection Methods.

## 2.1 Customer Churn

Customer churn is the probability that a customer will end their relationship with a company (Kamakura et al., 2005). Since acquiring new customers is often more expensive than retaining existing ones (Athanassopoulos, 2000; Tamaddoni Jahromi et al., 2014), accurately measuring and predicting churn is essential for business health.

## 2.2 Supervised Machine Learning

Supervised machine learning trains a model to predict a target variable ($Y$) from input features ($X$) (Nasteski, 2017). The model can be a regressor (for continuous values) or a classifier (for categorical values). Typically, the data is split into training (e.g., 80%) and validation (e.g., 20%) sets.

## 2.3 Support Vector Machine

A Support Vector Machine (SVM) is a supervised learning algorithm that separates data points from different classes using a linear or non-linear function (e.g., a line or hyperplane). See (Mammone et al., 2009) for more details.

## 2.4 Feature Selection Methods

Feature selection is crucial in AI modeling, especially with real-world data (Kumar and Minz, 2014). Irrelevant or redundant features can hurt model performance. This study uses Sequential Feature Selection and Logistic Regression for feature selection.

### 2.4.1 Sequential Feature Selection

Sequential Feature Selection (SFS) selects the $n$ features that maximize accuracy, measured across cross-validated models. The best performing set is kept.

Sequential Forward Floating Selection (SFFS) is a variation of SFS. SFFS iteratively adds or removes features based on their contribution to accuracy. The "floating" search allows the algorithm to remove previously selected features if they become redundant.

### 2.4.2 Logistic Regression

Logistic Regression is a supervised learning model often used for feature selection. It can effectively measure feature importance and reduce data dimensionality without significant loss of accuracy (Cheng et al., 2006). In this study, Logistic Regression is used for feature selection, not classification.

# 3 RELATED WORKS

Customer churn prediction has gained significant attention in recent years, particularly in industries such as telecommunications, banking, e-commerce, and B2B services (Manzoor et al., 2024). Various machine learning models have been applied to identify patterns and predict churn, with particular focus on models such as Support Vector Machines (SVM), Random Forest, K-Nearest Neighbors, Gradient Boosting, Logistic Regression, Naive Bayes, Decision Trees and Neural Networks (A. and D., 2016).

In the telecommunications sector, multiple studies emphasize the effectiveness of ensemble methods and Support Vector Machines (SVMs) for churn prediction. According to (Poudel et al., 2024), which utilizes a telecommunications dataset from Kaggle, Gradient Boosting Machine (GBM) achieved the highest accuracy at 81%, outperforming models like SVM, Logistic Regression, Random Forest, and Neural Networks. In contrast, (Ullah et al., 2019), analyzing a South Asian mobile communications service provider dataset, showed that Random Forest outperformed Decision Trees, Naive Bayes, Bagging, and Boosting, with an accuracy of 88.63%. However, (Rodan et al., 2014), which examined customer churn in a Jordanian telecommunications company, demonstrated that SVM achieved an outstanding accuracy of 98.7%, surpassing other models like Neural Networks, Decision Trees, and Naive Bayes, highlighting its robustness in handling telecommunications datasets.

In the banking domain, studies such as (Tran et al., 2023) and (Xiahou and Harada, 2022) highlight the effectiveness of advanced algorithms, demonstrating that integrating k-means clustering with SVM can significantly enhance prediction accuracy. The first study, based on a Kaggle dataset containing information on banking customers, found that Random Forest consistently outperformed other models, achieving an accuracy of up to 97.4%. Similarly, in the B2C e-commerce sector, research by (Xiahou and Harada, 2022), using a dataset from Alibaba Cloud Tianchi, showed that SVM, after performing customer segmentation, outperformed Logistic Regression with an accuracy of 91.56%, further highlighting its effectiveness in predicting customer churn.

In the B2B sector, churn prediction often involves analyzing complex customer behavioral and transactional patterns. (Tamaddoni Jahromi et al., 2014) explored models such as Decision Trees and AdaBoost, highlighting the latter's effectiveness in addressing class imbalances. Similarly,(Gordini and Veglio, 2017), using data from an Italian online retail company, demonstrated the superiority of AUC-optimized SVM over traditional SVM, Neural Networks, and Logistic Regression, achieving an accuracy of 89.67%. In subscription-based services, (Coussement and Van den Poel, 2008), analyzing data from a Belgian newspaper publisher, highlighted the utility of SVM, particularly with AUC-based parameter tuning. However, Random Forest ultimately delivered the highest predictive accuracy. Collectively, these findings highlight SVM's adaptability and effectiveness, positioning it as a strong candidate for B2B churn prediction tasks where precise customer retention strategies are critical.

# 4 METHODOLOGY

This section presents the proposed customer churn prediction model, as illustrated in Figure 1.

This diagram outlines the methodology for a customer churn prediction project within an HR company. The process begins with gathering datasets related to churn prediction and customer usability. A sampling process focuses on active customers relevant to churn prediction. Feature engineering follows, involving calculating differences in months and days, creating interval features, remapping product categories, applying one-hot encoding, and handling missing values. Feature selection utilizes logistic regression, Sequential Forward Floating Selection (SFFS), and 5-fold cross-validation. A Support Vector Machine (SVM) model is then trained on the selected features. The model is evaluated using metrics such as accuracy, precision, recall, and F1-score. Finally, explainer models, including, Fast API, and Streamlit, are used to interpret the model's predictions and understand the key drivers of churn.

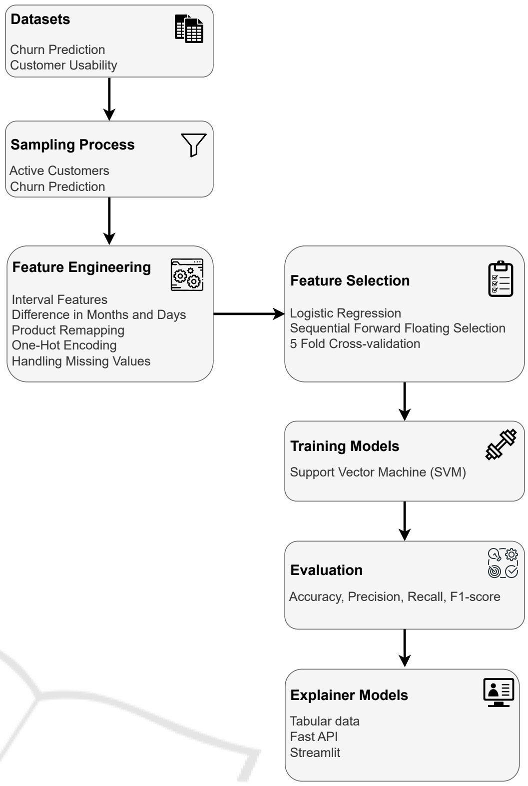


Figure 1: An illustration of the data processing, model training, evaluation and explainer models on the customer churn data.

## 4.1 Data Preprocessing

### 4.1.1 Dataset

The dataset used in this study comes from an Human Resources Technologies (HR Tech) company and includes 8,878 business customers, all categorized as Small and Medium Enterprises (SMEs). These customers span various sectors, including commerce, retail, technology, consultancy, marketing, and services. They are classified within the company's portfolio based on the number of employees registered on the platform. The Customer Relation Manager (CRM) database was used during the period from March 2024 to November 2024.

Additionally, all available data from canceled customers were used, while for active customers, a stratified sample was used. The combination of these two datasets forms the final dataset used to train the SVM model. Including canceled customers is crucial for the model to learn to identify the factors that lead to churn and, consequently, predict the probability of churn for new customers.

#### 4.1.2 Sampling Process

Active user data is divided into groups based on the unique products. For each product category, a sample size is calculated proportionally to the representation of that stratum in the original dataset. Equation 1 displays how we calculate this.

$$\bar{S} = c\frac{S}{N} \tag{1}$$

where $c$ is the multiplication factor defined as 1.5 times the size of the canceled users dataset. This ensures that the final dataset has a higher proportion of users who canceled. $S$ is the number of active users belonging to the specific product category, and $N$ is the total number of active users. If $\bar{S} \leq 0$, $\bar{S}$ is set to 1, ensuring that all products are present in the training dataset, regardless of length. For each group, we randomly select $\bar{S}$ samples, and combine each random sample, generating a final stratified sample containing all groups.

#### 4.1.3 Feature Engineering

Feature engineering is performed in several stages to prepare raw data for training the churn prediction model. The primary transformations include:

- Difference in Months: Calculates the number of months between a customer's subscription date and the current date.
- Difference in Days: Calculates the number of days between various event dates (last login, last job post created, etc.) and the current date. The generated features are named as "days without the event".
- Interval Features: Creates interval features (bins) for numerical features calculated in the previous steps. The values of each feature are divided into intervals based on quartiles, generating features such as "interval months subscriber" (from 0 to 12), "interval days without logging" (from 0 to 7), and so on.
- Product Remapping: Groups similar product categories into more general categories.
- One-Hot Encoding: Converts categorical features into dummy variables (0 or 1).
- Missing Value Treatment: Replaces missing values with 0.

Based on the transformations performed, the features were classified into two main types: numerical and categorical. These types were organized into functional groups to facilitate analysis, including Usability, Engagement, Performance, Interval Usability, and Interval Performance, as described in Table 1. The Engagement group includes numerical features that capture users' direct interaction with the platform, while the Usability group contains both numerical and interval features that monitor behaviors and usage patterns. The Performance group evaluates the outcomes generated by user actions and also includes both numerical and interval features. The Product group includes information about the service terms contracted by the customer. This classification was designed to preserve the confidentiality of the organization's operational details while ensuring that the data provides relevant insights for churn analysis.

#### 4.1.4 Feature Selection

To select features, we employed a Logistic Regression model and selected the most significant features by removing those that did not contribute to the model's accuracy (Ververidis and Kotropoulos, 2005). After that, we applied SFFS for feature selection. The selected features were chosen based on the 5-fold cross-validation results of the trained model.

The accuracy metric, reflecting the proportion of correct classifications, was used to evaluate the model's performance. To reduce dimensionality without sacrificing crucial information, only the top 34 most relevant features were selected, as shown in Table 1.

### 4.2 Model Architecture

After selecting the most significant features, the dataset was divided into subsets of training and testing, with 75% and 25% of the data being allocated for training and testing, respectively.

To address the class imbalance in the dataset, weights were assigned to the classes, with the no-churn class given a higher weight (2) and the churn class a lower weight (1). This exploratory approach aimed to balance the influence of both classes during model training, ensuring the no-churn class had

Table 1: Features analyzed in the model.

| Type | Group | Feature |
|---|---|---|
| Numerical | Engagement | F1 |
| | | F2 |
| | Usability | F3 |
| | | F4 |
| | | F5 |
| | | F6 |
| | Performance | F7 |
| Categorical | Product | F8 |
| | | F9 |
| | | F10 |
| | Interval Usability | F11 |
| | | F12 |
| | | F13 |
| | | F14 |
| | | F15 |
| | | F16 |
| | | F17 |
| | | F18 |
| | | F19 |
| | | F20 |
| | | F21 |
| | | F22 |
| | | F23 |
| | | F24 |
| | | F25 |
| | | F26 |
| | | F27 |
| | Interval Performance | F28 |
| | | F29 |
| | | F30 |
| | | F31 |
| | | F32 |
| | | F33 |
| | | F34 |

greater significance in shaping the model's learning process.

However, this balancing approach is not equivalent to the more well-known techniques in the literature, such as oversampling or undersampling (Drummond and Holte, 2003). Unlike these techniques, which alter the quantity of data in each class, this approach adjusts the importance of each class, allowing the model to account for the disproportionate impact of imbalanced classes without altering the original data structure. This helps improve the model's accuracy by mitigating potential biases toward the majority class and balancing learning and prediction between the two classes.

The model used to predict customer non-churn was the Support Vector Machine (SVM), which proved to be suitable for binary classification problems with high-dimensional data (Mammone et al., 2009). In the model construction process, the Linear kernel was chosen, which tries to separate the classes through a linear hyperplane in the original data space. The choice of this kernel is due to its computational simplicity and suitability for problems where the classes are approximately linearly separable (Cortes and Vapnik, 1995).

The performance of the SVM with the Linear kernel strongly depends on the regularization parameter $C$, which controls the trade-off between maximizing the margin separating the classes and minimizing classification errors. When the value of $C$ is high, the model tries to minimize classification errors during training at all costs, which can lead to overfitting. On the other hand, when the value of $C$ is low, the model accepts more errors, prioritizing the separation between classes, which can result in underfitting (Coussement and Van den Poel, 2008). Based on this, the chosen value was 1.0, as it provided a good balance between maximizing the margin of separation and minimizing classification errors.

## 4.3 Evaluation Metrics

For the model's quality assessment, the metrics of accuracy (Equation 3), recall (Equation 4), and f1-score (Equation 5) were applied.

$$\text{precision} = \frac{TP}{TP+FP} \tag{2}$$

$$\text{accuracy} = \frac{TP+TN}{TP+FN+TN+FP} \tag{3}$$

$$\text{recall} = \frac{TP}{TP+FN} \tag{4}$$

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{5}$$

In equations 2, 3, 4, and 5, $TP$, $FP$, $TN$, and $FN$ represent true positive, false positive, true negative, and false negative, respectively. In our context, for a given observation, the model:

- Correctly predicts a non-churned customer, for $TP$.
- Incorrectly predicts a non-churned customer as churned, for $FP$.
- Correctly predicts a churned customer, for $TN$.
- Incorrectly predicts a churned customer as non-churned, for $FN$.

More specifically, Equation 3 measures the overall proportion of correct predictions. Equation 4 quantifies the proportion of true positive instances among all actual positive instances. Equation 5 calculates the harmonic mean of precision and recall, providing a balanced measure of the model's performance. Ideally, all values should approach 1, indicating optimal performance.

Recall was specifically used for determining if the model needed to be retrained or not. On a threshold of 60% recall, we defined the need for retraining if a trained model achieved a recall of less than 60%. The use for recall instead of any other metric is that there is a need to update the true model's capacity for avoiding churns, and the recall measures exactly this, while others do not.

# 5 RESULTS

## 5.1 Experiments

The analysis of predictions for all customers over a 6-month period, using a single train/test split, revealed distinct results for the SVM model. In this scenario, the model achieved an accuracy of 0.79, precision of 0.44, recall of 1.00, and an F1-Score of 0.62, as shown in Table 2.

Table 2: Overall confusion matrix.

| – | Non-churners | Churners |
|---|---|---|
| Non-churners | 5459 | 1901 |
| Churners | 582 | 936 |

Comparing these results with the biweekly analysis, we identified significant differences. These variations are detailed in tables 5 to 8, highlighting potential implications for retention strategies and necessary adjustments to the model for different time windows.

Table 3: Confusion matrix for the first biweekly period.

| – | Non-churners | Churners |
|---|---|---|
| Non-churners | 5514 | 1665 |
| Churners | 56 | 112 |

Table 4: Confusion matrix for the second biweekly period.

| – | Non-churners | Churners |
|---|---|---|
| Non-churners | 5735 | 1695 |
| Churners | 58 | 152 |

Table 5: Confusion matrix for the third biweekly period.

| – | Non-churners | Churners |
|---|---|---|
| Non-churners | 4886 | 1531 |
| Churners | 37 | 75 |

Table 6: Confusion matrix for the fourth biweekly period.

| – | Non-churners | Churners |
|---|---|---|
| Non-churners | 5171 | 1546 |
| Churners | 34 | 49 |

Table 7: Confusion matrix for the fifth biweekly period.

| – | Non-churners | Churners |
|---|---|---|
| Non-churners | 5453 | 1900 |
| Churners | 40 | 84 |

Table 8: Confusion matrix for the sixth biweekly period.

| – | Non-churners | Churners |
|---|---|---|
| Non-churners | 5514 | 1665 |
| Churners | 56 | 112 |

## 5.2 Metrics Analysis

The biweekly analysis accuracy (0.74 to 0.77) was similar to the previous analysis (0.79), showing consistent overall classification performance. However, accuracy can be misleading with imbalanced classes or unequal error costs (false positives/negatives) (Coussement and Van den Poel, 2008). A model predicting all customers as "non-churn" can have high accuracy but be useless (Coussement and Van den Poel, 2008). Precision is important in churn prediction because false positives lead to unnecessary retention costs (Coussement and Van den Poel, 2008).

The biweekly analysis precision (0.03 to 0.08) was much lower than the previous precision (0.44), indicating a significant increase in false positives. While the earlier analysis balanced precision and recall, the biweekly analysis prioritized recall at the cost of precision. This low precision suggests a persistent problem with false positives (Coussement and Van den Poel, 2008). The drop in precision suggests the traditional approach overestimated the model's generalization ability. (Tamaddoni Jahromi et al., 2014) and (Coussement and Van den Poel, 2008) emphasize the importance of considering the time-varying nature of churn data, as customer behavior changes. The biweekly analysis captures this variation for a more realistic performance assessment.

The biweekly analysis recall (0.59 to 0.72) was lower than the previous perfect recall (1.00). This indicates the model missed more churning customers in the biweekly analysis.

The biweekly analysis F1-score (0.06 to 0.15) was also much lower than the previous F1-score (0.62), reflecting the poor precision.

The main difference between the results is the lower precision in the biweekly analysis. This suggests the single train/test split overestimated the model's generalization. The biweekly analysis reveals a tendency for more false positives, negatively impacting precision. Comparing the biweekly results with the single train/test split results (accuracy 0.79, precision 0.44, recall 1.00, F1-score 0.62) is crucial. The precision drop suggests the traditional approach overestimated generalization. (Tamaddoni Jahromi et al., 2014) and (Coussement and Van den Poel, 2008) highlight the importance of considering temporal data variation in churn modeling. The biweekly analysis captures this and provides a more realistic performance view.

The decreased recall in the biweekly analysis is also important, showing a reduced ability to identify at-risk customers. However, the precision drop is the main contributor to the lower F1-score. (Tamaddoni Jahromi et al., 2014) address B2B churn, proposing an approach that considers customer heterogeneity and profit. While not discussing biweekly analysis directly, their emphasis on temporal variation and robust models supports the idea that these metrics are affected by the time frame used (Tamaddoni Jahromi et al., 2014).

Figure 2 shows our six biweekly intervals and model updates when recall is below our threshold.

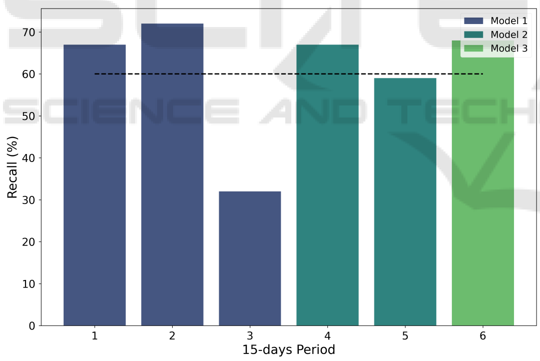


Figure 2: Recall of models leading to retraining.

When recall falls below the threshold, we retrain the model to keep it above 60%. These values indicate that changing customer needs often reduce model capabilities, requiring periodic retraining.

(Coussement and Van den Poel, 2008) compare SVM parameter selection for churn prediction. While focusing on modeling, their discussion of model generalization and performance evaluation across scenarios (including class distributions) supports biweekly analysis. Their use of a time window for predictive variables further supports analyzing temporal data variation (Coussement and Van den Poel, 2008).

## 5.3 Models Analysis

The feature selection resulted in a subset of 10 features deemed most relevant for churn prediction. The accuracy achieved during cross-validation in the feature selection process was 0.72. (Tamaddoni Jahromi et al., 2014) and (Coussement and Van den Poel, 2008) emphasize the importance of considering temporal variation in data when modeling churn, as customer behavior and the factors influencing churn can change over time. Biweekly analysis, by evaluating the model across different periods, captures this variation and provides a more realistic assessment of its performance.

The trained SVM model was then evaluated on a separate test set, similarly, (Rodan et al., 2014) focus on using SVM to predict churn in the telecommunications industry. The authors employ metrics such as accuracy, hit rate, churn rate, and lift coefficient to evaluate the SVM model's performance and compare it with other approaches, such as neural networks and decision trees. Although they do not explicitly mention AUC or Gini, their emphasis is on the model's ability to correctly identify customers likely to churn (Rodan et al., 2014).

The high AUC (0.9029) and Gini coefficient (0.8057) indicate excellent model discrimination, effectively distinguishing churn customers from non-churn customers. According to (Coussement and Van den Poel, 2008), AUC is a crucial metric for evaluating churn models, as it is more robust to varying classification thresholds and class imbalances compared to accuracy. The authors conduct statistical tests to compare the AUC of different models, including SVMs and logistic regression. Additionally, they emphasize the importance of lift and top-decile lift, which assess the model's ability to identify customers with the highest likelihood of churn.

Similarly, (Poudel et al., 2024) compare various models, including GBM and neural networks, using metrics such as accuracy, precision, recall, F1-score, ROC-AUC, and PR-score (area under the Precision-Recall curve). They also employ SHAP values to explain the model's predictions, providing insights into the key features influencing the results.

In summary, the SVM model, trained with an optimized subset of 34 features via SFFS, demonstrated excellent performance in churn prediction, evidenced by high accuracy and AUC. Analysis with coefficient values provided actionable insights into the main drivers of churn, highlighting the importance of customer engagement, especially regarding frequent logins, job creation, and the occurrence of positive interactions on the platform.

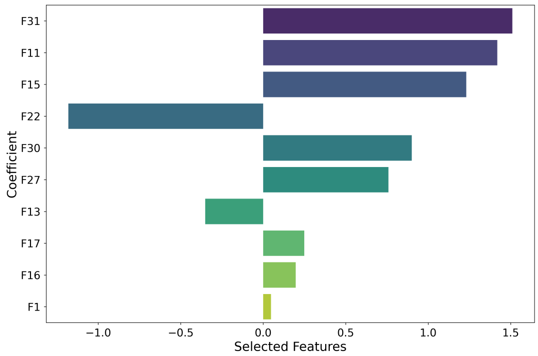


Figure 3: Features importance analyzed.

### 5.4 Features Analysis

Feature F11, with a coefficient of 1.4186, indicates that customers inactive for 8 to 44 days are significantly more likely to churn, highlighting the importance of regular engagement. Strategies such as sending relevant content and personalized notifications could be effective in mitigating churn. The importance of frequent logins is corroborated by various studies on churn in online services. In the paper by (Poudel et al., 2024) the feature “Tenure Months” (subscription duration) has a strong impact on churn, and inactivity for 8-44 days can be an early sign of disinterest. The suggestion to send relevant content and personalized notifications aligns with retention strategies focused on proactive engagement, as discussed by (Tamaddoni Jahromi et al., 2014).

Feature F15, with a coefficient of 1.2309, also points to a high churn risk when customers don’t create new job postings for 69 to 823 days. Actively using this feature suggests they are engaged and less likely to churn. Offering tutorials and support to encourage job postings could be helpful (Coussement and Van den Poel, 2008). Increasing customer engagement with key platform features, like job postings, can effectively reduce churn.

Feature F31, with a coefficient of 1.5098, is the most significant predictor of churn, highlighting the critical impact of positive interactions (e.g., feedback, promotions) on customer retention. A lack of such interactions for 37 to 55 days is an even stronger indicator of churn than inactivity in logins or job postings. This underscores the importance of continuous engagement beyond mere platform usage. To mitigate churn, companies should proactively foster positive interactions through strategies like gamification, rewards, and recognition, ensuring customers feel valued and engaged.

Feature F27, with a coefficient of 0.7588, indicates that not using the engineering feature for 101 to 200 days increases churn risk. Like feature F15, this highlights the importance of regularly using platform features. While “engineering” may be less critical than job creation, promoting its use is still important for retention.

Feature F30, with a coefficient of 0.9003, is similar to F31. It shows that not having positive interactions for 31 to 36 days also increases churn risk, though slightly less. Both features emphasize the importance of positive occurrences, but the 37-55 day period (F31) appears more critical.

Features F16 (coefficient 0.1969) and F17 (coefficient 0.2494) both show that any period of inactivity without creating profiles increases churn risk. This suggests that regularly creating profiles is a key sign of customer engagement, though less impactful than other features.

Feature F1, with a coefficient of 0.0477, indicates that a high number of open job postings slightly increases the likelihood of churn. This might mean users are frustrated with the recruitment process. It’s worth checking if these jobs are hard to fill or if there are platform usability issues.

Feature F13, with a coefficient of -0.3513, shows that short breaks (3 to 7 days) in creating job postings are linked to a lower churn risk. This could be a normal pattern, where users post jobs, wait for results, and then continue.

Feature F22, with a coefficient of -1.1810, reveals that not having evaluations for 3 to 229 days is tied to a lower churn risk. This suggests that how often users evaluate may not directly show their satisfaction. Happy customers may not feel the need to constantly evaluate.

The coefficient analysis of each feature highlights that keeping customers engaged and consistently using platform features is crucial for retention. By understanding these factors, companies can create targeted strategies to reduce churn. Personalized retention strategies should be developed based on the coefficient analysis of each feature. For example, those with high coefficient for days without logging in should receive engagement incentives. Tracking the evolution of each coefficient analysis can also help identify changes in customer behavior and adapt strategies accordingly.

Using coefficient analysis of each feature with other metrics, such as AUC, precision, and recall, provides a more complete picture of the model’s performance and what drives churn.

### 5.5 Result Discussions

The SVM model displays overall good performance regarding recall when it is re-trained periodically. The results show that with time the pattern changes, thus reducing the capabilities of the obsolete model to correctly assess the predictions. This indicates the need for constant updates on model training to optimize recall.

Consistent analysis identifies customer inactivity as a strong predictor of churn. Long periods without logins (F11), job postings (F15), or positive interactions (F31) strongly indicate higher churn risk.

Addressing B2B churn in HR Tech is challenging. To our knowledge, there are few studies specifically on HR Tech churn, highlighting the novelty and importance of our work. Our findings offer valuable insights into HR Tech churn factors, enabling tailored retention strategies. By focusing on proactive engagement and addressing inactivity, HR Tech companies can improve customer retention.

# 6 CONCLUSIONS AND FUTURE WORKS

This study investigated churn prediction in a B2B context using biweekly data and a SVM model with a linear kernel. Feature selection through SFFS, combined with interpretability analysis based on the coefficient values of each feature, enabled the identification of key churn drivers and the development of actionable insights for retention strategies.

The model's accuracy stability across the analyzed biweekly periods demonstrates its robustness to temporal data variations, a crucial aspect in dynamic environments like B2B. Additionally, the consistent minimization of false positives in four out of five biweekly periods reduces the cost of unnecessary interventions, optimizing retention resources.

The coefficient analysis of each feature revealed valuable insights into customer behavior and the factors influencing churn. Platform inactivity, extended periods without creating new job postings, and the absence of positive occurrences emerged as the strongest predictors of churn. These findings align with existing literature emphasizing the importance of customer engagement and positive service experiences for retention. The analysis also highlighted the influence of other features, such as inactivity in the engineering functionality and profile creation, pointing to the need for further investigation into the relationship between evaluation frequency and churn.

Despite stable accuracy, our results indicate that the recall falls over customer necessities. Because of this, it may be necessary to periodically retrain the model for more consistent results. The discrepancy suggests potential overfitting in the traditional approach and underscores the importance of temporal validation for a more realistic assessment of model performance.

This study contributes to B2B churn prediction by identifying key churn drivers, demonstrating the value of coefficient-based feature analysis for model interpretability, and proposing targeted retention strategies, such as encouraging frequent logins, promoting key functionalities, and fostering positive platform interactions. For future research, exploring alternative machine learning models, including ensemble methods (Random Forest, Gradient Boosting) and Deep Learning approaches, could provide valuable performance comparisons with SVM. Additionally, incorporating class balancing techniques, such as Class Weight adjustments, SMOTE, and Random Over and Under-Sampling, may improve model effectiveness by addressing data imbalance. Further investigation of new predictive features, such as detailed behavioral metrics, contextual company characteristics, and seasonal factors, could enhance model accuracy and provide deeper insights into customer retention dynamics.

# REFERENCES

A., S. and D., C. (2016). A Survey on Customer Churn Prediction using Machine Learning Techniques. *International Journal of Computer Applications*, 154:13–16.

Athanassopoulos, A. D. (2000). Customer satisfaction cues to support market segmentation and explain switching behavior. *Journal of Business Research*, 47(3):191–207.

Cheng, Q., Varshney, P. K., and Arora, M. K. (2006). Logistic regression for feature selection and soft classification of remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 3(4):491–494.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

Coussement, K. and Van den Poel, D. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. *Expert Systems with Applications*, 34(1):313–327.

Drummond, C. and Holte, R. (2003). C4.5, class imbalance, and cost sensitivity: Why under-sampling beats over-sampling. *Proceedings of the ICML'03 Workshop on Learning from Imbalanced Datasets*.

Gordini, N. and Veglio, V. (2017). Customers churn prediction and marketing retention strategies. An application of support vector machines based on the AUC parameter-selection technique in B2B e-commerce industry. *Industrial Marketing Management*, 62:100–107.

Kamakura, W., Mela, C. F., Ansari, A., Bodapati, A., Fader, P., Iyengar, R., Naik, P., Neslin, S., Sun, B., Verhoef,

P. C., et al. (2005). Choice models and customer relationship management. *Marketing letters*, 16:279–291.

Kumar, V. and Minz, S. (2014). Feature selection. *SmartCR*, 4(3):211–229.

Li, Y., Cheng, Y., Ni, Y., and Zhang, N. (2023). Assessment of product resilience towards the changing customer requirements. *Journal of Engineering Design*, 34(8):561–590.

Mammone, A., Turchi, M., and Cristianini, N. (2009). Support vector machines. *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(3):283–289.

Manzoor, A., Atif Qureshi, M., Kidney, E., and Longo, L. (2024). A Review on Machine Learning Methods for Customer Churn Prediction and Recommendations for Business Practitioners. *IEEE Access*, 12:70434–70463.

Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons. b*, 4(51-62):56.

Poudel, S. S., Pokharel, S., and Timilsina, M. (2024). Explaining customer churn prediction in telecom industry using tabular machine learning models. *Machine Learning with Applications*, 17:100567.

Rodan, A., Faris, H., Al-sakran, J., and Al-Kadi, O. (2014). A Support Vector Machine Approach for Churn Prediction in Telecom Industry. *International journal on information*, 17.

Tamaddoni Jahromi, A., Stakhovych, S., and Ewing, M. (2014). Managing b2b customer churn, retention and profitability. *Industrial Marketing Management*, 43(7):1258–1268.

Tran, H. D., Le, N., and Nguyen, V.-H. (2023). Customer Churn Prediction in the Banking Sector Using Machine Learning-Based Classification Models. *Interdisciplinary Journal of Information, Knowledge, and Management*, 18:087–105.

Ullah, I., Raza, B., Malik, A. K., Imran, M., Islam, S. U., and Kim, S. W. (2019). A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector. *IEEE Access*, 7:60134–60149.

Ververidis, D. and Kotropoulos, C. (2005). Sequential forward feature selection with low computational cost. In *2005 13th European Signal Processing Conference*, pages 1–4.

Xiahou, X. and Harada, Y. (2022). B2C E-Commerce Customer Churn Prediction Based on K-Means and SVM. *Journal of Theoretical and Applied Electronic Commerce Research*, 17(2):458–475.