Andrew Ferreira

**Church Member Churn Prediction and Growth Analysis: A Framework for Churches to Analyze Their Church Management Data**

**Abstract**

Churches, similar to other organizations, encounter challenges in member retention. Unlike businesses, many churches lack advanced tools for this purpose. This study examined the potential of business-based churn prediction methods to address member churn in churches. Utilizing data from Rock RMS, an open-source church management tool, a Decision Tree algorithm was applied to pinpoint member churn predictors. Key insights show that profiles organically created, with lower online engagement, older than 12 months, and with marital statuses other than single have a higher likelihood of retention.

Additionally, the study incorporated tech-industry metrics like Activation Rate, Monthly Active Users (MAU), and Stickiness to assess church member engagement. Preliminary findings indicate that business-oriented churn prediction models are suited for churches, and tech-industry metrics also offer valuable engagement insights. By adopting these methodologies, churches have the potential to deepen their relationships with members and enhance retention.

Figure 1: Graphical Abstract

A diagram of a church

Description automatically generated

1. **Introduction**

Regardless of industry, products, or services, the customer fuels any organization's existence. Every organization has strategies in place to find their potential customers. Once they do, they must find ways to keep them because campaigns to acquire new customers are usually expensive [2]. Churn is a term many organizations use to describe a customer who stops returning at some point [3]. Churn analysis helps organizations find their customers' churning points so they can act proactively by removing potential barriers that cause the churn, extending the time their customers will remain loyal to them. In a church, people are not called "customers" but "members." While members switching from church A to church B is not considered a customer loss because of the business's nature, churches do not know whether the switch is happening or if the member is giving up on their faith, which is a critical loss for churches. So many new church management tools were created to help churches track the engagement of their members. However, few churches have enough resources to analyze the data from these tools, and very few studies have implemented some type of churn analysis; in most cases, churches will only have demographic data from the Census or private organizations that measure the population of Christians in a specific area. In this study, we intend to create a framework for predicting member churn using the data from Rock RMS [14], an open-source church management tool, and propose metrics to track church member growth. The T-SQL queries and code used in this study are available on a public GitHub repository [4].

1. **Literature Review**
   1. Customer Churn Prediction

Customer churn models attempt to identify early churn signals and predict what customers might soon leave the organization [2]. Many researchers are trying to apply different algorithms, machine learning techniques, and data processing methods to increase the overall accuracy of churn prediction. [5]

Studies have applied tree-based algorithms such as decision trees, random forest, XGBoost, and more to predict churners [6]. Others have tried support vector machines, naïve Bayes, and logistic regression [7] or even more recent and advanced techniques such as neural networks [8]. However, the discussions about these studies are about applying different techniques to improve their overall accuracy.

While some studies focus on algorithms, others concentrate on the data processing phase on the premise that proper data processing yields better predictions [9].

As Sulim Kim et al. mentioned, most churn prediction will keep their focus on comparing different techniques to decrease the prediction error as much as possible, and most of them are either in the telecommunication or finance industry instead of emerging businesses and churches. [5] This study aims to apply the customer or member churn prediction for churches to predict the churning point of their members.

* 1. Customer Engagement Metrics

Customer engagement metrics or product metrics are helpful ways to measure how customers interact with products and services organizations offer [10]. Many big-tech companies such as Meta [11] or Spotify [12] will include in their earnings reports activation metrics that show how long it takes for a customer to move from acquisition to consumption and engagement metrics to measure how often customers engage with the product and service [10]. Even articles targeted for churches recommends a few metrics such as regular and holiday attendance, donations, and volunteer numbers. [13] Even though some free church management tools, such as Rock RMS [14] or Church Metrics [15], help church leaders track and visualize their data, articles rarely show them how to use their raw data to measure engagement or conduct churn analysis. This study aims to apply product metrics to measure church engagement, and all the T-SQL queries and code used will be available in a public GitHub repository for churches to use. [4]

Different studies have different definitions of churn customers and engaged users [5][10]. Each organization is free to define churn and engagement depending on the service each organization offers [5]. For this study, an engaged member is defined as a member who has at least one trackable interaction (attendance, volunteering or financial contribution) with the church in a month, and a churned member is a member who goes two months without any interaction.

1. **Data & Method**
   1. Data
      1. Data Source

Rock is an open-source solution used by hundreds of churches worldwide. It helps track and manage the relationship between the church and its members by tracking common life events and demographics such as marriage, graduation, birthdays, and addresses. It also helps to track events such as attendance, financial contributions, volunteering, baptism, joining a small group, and more. The system stores the normalized data on SQL Server, a relational database management system, allowing data analysts to query the database using T-SQL and provide insights for the organization. [14]

3.1.2 Data Summary

We used the Rock RMS from a church that started using the application in 2018. All the T-SQL queries used to gather the following variables are available in the public GitHub repository [4]:

Table 1: Descriptive statistics

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Variable | Obs. | Min | 1st Quad. | Median | Mean | 3rd Quad. | Max |
| Members | 188,400 |  |  |  |  |  |  |
| Married | 412 |  |  |  |  |  |  |
| Single | 150,981 |  |  |  |  |  |  |
| Unknown marital status | 37,006 |  |  |  |  |  |  |
| Older than 18 years old | 79,531 |  |  |  |  |  |  |
| Younger than 18 years old | 108,851 |  |  |  |  |  |  |
| Female | 79,652 |  |  |  |  |  |  |
| Male | 77,264 |  |  |  |  |  |  |
| Unknown gender | 31,483 |  |  |  |  |  |  |
| Imported profile | 152,749 |  |  |  |  |  |  |
| Profile organically created | 35,650 |  |  |  |  |  |  |
| Total attendance |  | 1 | 2 | 3 | 8.6 | 8 | 73 |
| Days to first attendance |  | 0 | 0 | 0 | 0.1 | 0 | 7 |
| Months between first and last attendance |  | 0 | 0 | 1 | 17.2 | 26 | 125 |
| People in the household |  | 1 | 1 | 2 | 2.4 | 4 | 8 |
| Financial contributions |  | 0 | 0 | 0 | 0.1 | 0 | 182 |
| Serving Interest |  | 0 | 0 | 0 | 0 | 0 | 5 |
| Engagement Streak |  | 1 | 1 | 1 | 1.9 | 2 | 91 |
| Percentage of online engagement |  | 0 | 0 | 0 | 12.54 | 0 | 10 |

Table 2: Training and test dataset distributions

|  |  |  |  |
| --- | --- | --- | --- |
| Training/Testing datasets | Type | Number of observations | Percentage |
| Training dataset | Total | 150,719 | 100% |
| Churners | 138,466 | 92% |
| Non-churners | 12,253 | 8% |
| Test dataset | Total | 37,680 | 100% |
| Churners | 34,541 | 92% |
| Non-churners | 3,139 | 8% |

3.2 Method

3.2.1 Decision Trees

This study uses the customer or member churn prediction algorithm to identify patterns or churning points for a church member based on previous studies of customer churn prediction. The Decision Tree (DT) is frequently used for classification because it is "the most interpretable model by a human" with great accuracy. [5][16] The output of a DT is a tree displaying how variables are connected and what variables lead to a final classification, so the interpretation is simpler. [16] We decided on the churning point based on the number of trackable activities recorded in Rock RMS. If a member has not been active for two months by attending, serving, or contributing financially to the church, it will be considered a churn (the member left the church); otherwise, the person stayed at church. In a business setting, we usually define churn as a customer buying a product or service the organization offers for the second or third time. Each organization has the freedom to define what action and how often that action needs to happen to distinguish between a loyal or a churn customer [10]. Churches are not trying to sell any product or service; their objective is to keep members engaged and respecting their own pace on what actions each member is willing to take. Therefore, any attendance, volunteering, or financial contribution records are enough for this paper to activate a member. Frequency is the second variable used to classify loyalty or churn; how often each member performs one of these events defines a churn or not, and according to Sulim Kim et al., the frequency varies from organization to organization. Churches expect members to come to their services once a week, except when members are traveling on vacation or ill, making many church leaders consider four weeks of absence as usual. For the case study of this project, the average weeks between the first and last trackable events for each member are eight weeks, so anyone who goes longer than two months without any activity will be considered a churn.

The general idea of DT is to find the most popular class in the final output model [17]. To predict churning members, we implemented the "rpart" library using R, a programming language widely used for data science.

We evaluate churn prediction with the following criteria [18]:

- TP (True positives): "The number of customers that are in the churner category, and the prediction algorithm has determined their category correctly as

churner."

- TN (True negatives): "The number of customers that are in the non-churner category and the prediction algorithm has determined their category correctly as non-churner."

- FP (False positives): "The number of customers who are non-churners but the

algorithm incorrectly categorized them as churners."

- FN (False negatives): "The number of customers who are churners, but the algorithm incorrectly categorized them as non-churners."

"Recall is the ratio of real churners which are correctly identiﬁed," and we use the following formula:

"Precision is the ratio of predicted churners which are correct," and we use the following formula:

"Accuracy is the number of all the correct predictions," and we use the following formula:

"F-measure is the harmonic average of precision and recall," and we use the following formula:

Recall, Precision, Accuracy, and F-measure are metrics used to evaluate the model's ability to predict church member churn. Recall assesses how many actual members who churned were correctly identified by the model, capturing its completeness in identifying churn. Precision gauges the accuracy of the churn predictions by determining how many of the predicted churners truly did churn. Accuracy offers an overall measure of the model's correctness in identifying churners and non-churners relative to all predictions. "Generally, precision and recall are in tension. So, improving one metric causes reducing the other." F-measure, or F1-score, harmonizes Precision and Recall, providing a balanced single metric, especially useful in scenarios where improving one might diminish the other, often observed in imbalanced datasets. [16]

3.2.2 Engagement Metrics

Reports to investors released by major tech companies, such as Spotify, Facebook, Instagram, Messenger, WhatsApp, and Amplitude, have many metrics to measure the growth and engagement of their users. Organizations commonly use a few metrics: Monthly active users (MAU), weekly active users (WAU), and daily active users (DAU), which are ways they measure growth and how engaged their users are. [10] Amplitude is a company that specializes in helping these big tech companies define and evaluate their services by using the following metrics:

“Activation Rate indicates that the customer has oriented themselves with the product and what it can offer them,” and we use the following formula:

For this study, we defined the trigger (events that activate members) as attending a service, volunteering, or contributing financially.

“Monthly or Weekly Active Users (MAU and WAU respectively) refers to the number of active users per month or week.” According to Amplitude, the organization can define what event or action is meaningful enough to be considered an “active” user or member in our instance.

“Stickiness shows how often your users are coming back each month,” The formula to calculate stickiness for a church that provides weekly services is as follows:

If you have a stickiness rate of 50 percent, it shows that your church members come back only 2 out of 4 weeks in a month.

1. **Results**
   1. Member Churn Prediction

Table 3 compares the predicted results with the original data. The number of True Positive (TP) is 34,288, meaning that the algorithm identified 34,288 members who later would be classified as churners by our definition. The number of True Negative (TN) is 655, meaning that the algorithm identified 655 loyal members who would later decide to stay connected to the church for longer. One might say, in this instance, that the False Negative (FN) is the most concerning error since the algorithm did not identify 342 who were about to leave the church; therefore, FN must remain as low as possible, in our case, less than 1% of the dataset.

Table 3: Member Churn Prediction Results

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Data | Type | Classification | Predicted |  |  |
| Churner | Non-Churner | Total |
| Original | Count | Churner | 34,288 | 342 | 34,630 |
| Non-Churner | 2,395 | 655 | 3,050 |
| Proportion | Churner | 99.01% | 0.99% | 100% |
| Non-Churner | 78.52% | 21.48% | 100% |

* 1. Member Churn Prediction Evaluation

As we discussed before, there are different metrics to evaluate the results of the DT algorithm. Table 4 shows Recall, Precision, Accuracy, and F-measure, but according to Sulim Kim et al., F-measure is the real accuracy [5]; it “states the equilibrium between the precision and the recall.” [19] Thus, the real accuracy of the Decision Trees (DT) model for our case study using only Rock RMS data is 96.16% regarding f-measure.

*Table 4: Summary of Evaluation*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | Recall | Precision | Accuracy | F-Measure |
| Decision Tress | 99.01% | 93.47% | 92.74% | 96.16% |

4.3 Decision Tree Insights

Still, accuracy is not the only result of DT. Organizations use the algorithm to analyze the nodes of the trees because they might help organizations prevent churn, which is the purpose of running the algorithm [2].

The tree will differ from organization to organization. The source code used to collect and transform the data from Rock RMS for the “rpart” library in R and the R script itself are available in the public GitHub repository [4], so other churches can reuse the code and discover their own trees.

Figure 2: Decision Tree (simplified for readability)



For this case study, the algorithm provided the following insights:

* + Profiles organically created (not imported) have more chances to stay active with the church.
  + Profiles with high online engagement have more chances to leave the church.
  + Older profiles (more than 12 months) have more chances to stay active with the church.
  + Marital statuses other than single has more chances to stay active with the church.

4.4 Engagement Metrics

Figure 1 shows how the Activation Rate, a metric commonly used by tech companies to measure user engagement, can also be applied to a church that uses a church management tool such as Rock RMS. By looking at the trend, one can conclude that the church observed is either losing the engagement of its members or growing the number of Rock profiles created faster than working on activating the members already known. This type of insight is a metric many organizations use, but churches in their church management tools are not.

Figure 3: Activation Rate

![A graph showing the number of years

Description automatically generated]()

Figure 2 shows the application of the Monthly Active Users (MAU) applied to the church setting. It shows that the church has approximately 75 to 80 thousand active monthly members. When combined with the trends observed in Figure 1, it helps churches to understand that Figure 1 is trending down, not because the church is losing their engaged member, but because the creation of new accounts is growing faster than the church is engaging with the ones that are already created; otherwise, Figure 2 would also be trending down.

Figure 4: Monthly Active Users (MAU)

![A graph showing the number of people in the market

Description automatically generated with medium confidence]()

Finally, Figure 3 provides an insight that usual church metrics like regular attendance (the number of people manually counted in the building during a service) cannot capture. It reveals that members are active only 1.8 weeks (45%) out of the 4 weeks in a month. This type of insight can assist churches in evaluating the services they provide, much like it does for other organizations [10].

Figure 5: Stickiness

![A graph with lines and numbers

Description automatically generated]()

All the T-SQL queries to collect and transform the data are publicly available in the GitHub repository so other churches using Rock RMS can use them to evaluate their member engagement.

**5. Conclusion**

Predicting customer churn and engagement metrics is crucial for many organizations. This study attempts to apply these techniques in a church setting by using the Decision Trees (DT) algorithm and proposes new metrics for churches. Our analysis underscores the importance of key member engagement and retention indicators within a church. Specifically, we found that profiles created organically (not imported), those with lower online engagement, older profiles (more than 12 months), and those with marital statuses other than single are more likely to remain active with the church. These insights, derived from the DT algorithm, highlight the nuances of member retention within a church and showcase the feasibility of applying business-oriented prediction algorithms in a church setting. The study also emphasizes the utility of engagement metrics such as Activation Rate, Stickiness, MAU, and WAU for churches. It provides the source code for other churches to collect data from Rock RMS and apply it to an R script that outputs prediction results, allowing churches to better understand their churning points and engagement metrics. The T-SQL queries and R script are publicly available on a GitHub repository.

Works Cited

[1?]

[2] Lalwani, Praveen, et al. “Customer Churn Prediction System: A Machine Learning Approach.” *Computing*, vol. 104, no. 2, Feb. 2022, pp. 271–94. *EBSCOhost*, https://doi-org.oralroberts.idm.oclc.org/10.1007/s00607-021-00908-y.

[3] Collins Dictionaries, editor. “Churn.” Collins English Dictionary, 12th ed., Collins, 2014. Credo Reference, search.credoreference.com/articles/Qm9va0FydGljbGU6MzU1MzUwMA==. Accessed 22 Aug. 2023.

[4] GitHub

[5] Kim, Sulim, and Heeseok Lee. “Customer Churn Prediction in Influencer Commerce: An Application of Decision Trees.” *Procedia Computer Science*, vol. 199, Elsevier BV, Jan. 2022, pp. 1332–39. https://doi.org/10.1016/j.procs.2022.01.169.

[6] Ahmad AK, Jafar A, Aljoumaa K (2019) Customer churn prediction in telecom using machine learning in big data platform. Journal of Big Data 6(1):28

[7] Asthana P (2018) A comparison of machine learning techniques for customer churn prediction. International Journal of Pure and Applied Mathematics 119(10):1149–1169

[8] Brându¸soiu, I., Toderean, G., Beleiu, H.: Methods for churn prediction in the pre-paid mobile telecommunications industry. In: 2016 International conference on communications (COMM), pp. 97–100. IEEE (2016)

[9] Berger, P., & Kompan, M. (2019). User modeling for churn prediction in E-commerce. IEEE Intelligent Systems, 34(2), 44-52.

[10] Amplitude. (2023). 15 important metrics you should be tracking. <https://info.amplitude.com/product-metrics>

[11] Meta. (2022). Q4 2022 Earnings. In https://investor.fb.com. Retrieved February 3, 2023, from <https://investor.fb.com/investor-events/event-details/2023/Q4-2022-Earnings/default.aspx>

[12] Spotify. (2022b). Q4 2022 Earnings Report. In https://investors.spotify.com/. Retrieved February 3, 2023, from <https://s29.q4cdn.com/175625835/files/doc_financials/2022/q4/Shareholder-Deck-Q4-2022-FINAL.pdf>

[13] <https://outreachmagazine.com/features/leadership/69430-measuring-what-counts.html>

[14] <https://www.rockrms.com/>

[15] <https://churchmetrics.com/>

[16] M, Kiguchi, et al. “Churn Prediction in Digital Game-based Learning Using Data Mining Techniques: Logistic Regression, Decision Tree, and Random Forest.” *Applied Soft Computing*, vol. 118, Elsevier BV, Mar. 2022, p. 108491. <https://doi.org/10.1016/j.asoc.2022.108491>.

[17 Rachid, Ait Daoud, et al. “Clustering Prediction Techniques in Defining and Predicting Customers Defection: The Case of E-Commerce Context.” *International Journal of Electrical and Computer Engineering*, Institute of Advanced Engineering and Science (IAES), Aug. 2018, <https://doi.org/10.11591/ijece.v8i4.pp2367-2383>.

[18] Khodabandehlou, Samira, and Mahmoud Zivari Rahman. “Comparison of Supervised Machine Learning Techniques for Customer Churn Prediction Based on Analysis of Customer Behavior.” *Journal of Systems and Information Technology*, vol. 19, no. 1/2, Emerald Publishing Limited, Mar. 2017, pp. 65–93. https://doi.org/10.1108/jsit-10-2016-0061.

[19] Khodabandehlou, Samira, and Mahmoud Zivari Rahman. “Comparison of Supervised Machine Learning Techniques for Customer Churn Prediction Based on Analysis of Customer Behavior.” *Journal of Systems and Information Technology*, vol. 19, no. 1/2, Emerald Publishing Limited, Mar. 2017, pp. 65–93. <https://doi.org/10.1108/jsit-10-2016-0061>.