Churn Playbook
Churn Playbook

Churn is everywhere and presents itself in different forms for different parts of a business. HR departments use churn to refer to employees leaving the company voluntarily (e.g., staff churn at stores/call centers). Sales departments use churn to refer to a customer terminating a subscription, becoming an inactive user of a product, or ceasing to buy a product. Churn can also be used to refer to equipment failure at manufacturers. In healthcare, one can use churn to model concepts such as when a patient is likely to get a specific disease, be hospitalized, or even die.

This playbook is geared towards helping you use AI to model churn in your business. It goes over the following topics:

- How AI Helps Churn
- Questions to Consider During the Ideation of an AI Churn Model
- AI Solution for Churn
- Single Row per Entity
- Multiple Rows per Entity
- What Data is Best Suited for an AI Churn Solution?
- Tips for Better Churn Predictions in DataRobot
  - Use DataRobot Model Output to Seek Feedback from Domain Experts
  - Assess Model Accuracy
  - Use Group Partitioning
How AI Helps Churn

Consider the example of employee churn depicted in Figure 1, which shows how most organizations handle churn. Companies invest significant resources in hiring, onboarding, and training new employees. In most industries, it may be several months before an employee is fully on-boarded. Employee churn costs companies money and time. HR teams try their best to address this issue, but are usually limited to one-size-fits-all solutions for employee retention, resulting in some employees still leaving. In some departments this churn can be so big that it eats into the bottom line.

Now, consider the business process illustrated in Figure 2 once AI is introduced. The staff responsible uses automated AI in DataRobot to accurately predict individuals who are likely to churn and take preventive action in an effective manner. For instance, if the model suggests that a long commute to work is one of the reasons this employee is likely to churn, the company may consider allowing that employee to work remotely. In the retail business, a company like Hello Fresh may choose to send coupons to a specific customer if they think that customer is about to unsubscribe/churn. The same idea applies to manufacturing if you consider a failing part in a machine as analogous to an employee quitting the organization - in this case, a manufacturer can predict which parts are likely to fail and take proactive action before it becomes a problem.

Figure 1: Business process for preventing employee churn without DataRobot.

Figure 2: Business process for preventing employee churn with DataRobot.
Questions to Consider During the Ideation of an AI Churn Model

The goal of a churn model is ultimately to add business value. A model that has low ROI or cannot be implemented may not be worth the investment. Ask the following questions before building a model.

1. **What is the nature of your churn challenge?** Is it high volume churn where you have a lot of data to model the problem, or is it low volume churn where the data is scarce? Would an intervention make a meaningful difference as far as the bottom line is concerned? Knowing the answers to these questions will not only inform whether there will be any ROI resulting from building a churn model, but also guide you towards the most effective solution set-up.

2. **What can you do to prevent churn?** Answering this question for your business will highlight whether a churn model is worth exploring. It will also force you to consider the relevant stakeholders to include early on during solution formulation. For instance, if the only action you can take for a customer churn problem is improving customer service, then make sure the customer service team is involved with the implementation strategy.

3. **How will the model be embedded into the current business process?** Do you have an established approach to identifying and incorporating changes to business processes? If this is going to be a new program, do you have a mechanism to pilot it successfully in your organization? Far too often, churn models are built and no one is able to use them.

4. **What does churn mean in your situation?** An example definition for an employee churn problem could be “will the employee leave in the next three months?” In an online business, churn might be defined by the customer’s activity frequency. Use that frequency to determine a value below which you may consider that customer as having churned.

5. **What is the cost of churn?** Even if you have to develop a back of the envelope calculation, there should be some solid indication of ROI to justify building an AI model.

6. **Is all churn necessarily bad?** In an employee context, sometimes people just aren’t a good fit for your organization. The decision you take in response to a churn prediction should include some consideration of their performance, compensation, etc. In a customer service context, if you have customers whose expected spend is less than what you’d be spending to keep them, churn is in fact a good thing.
AI Solution for Churn

Before you implement any AI solution for churn, review the current approach your company uses to identify and predict churn. This will not only inform the baseline performance that your AI solution must achieve, but also help the business start gaining an understanding of the benefits of the AI solution for churn (See Figure 3).

A good churn prevention solution involves both a predictive model and complementary churn prevention actions that the business takes. While there are a number of ways you can set up the solution for this goal, in this playbook we are going to explore two approaches that you will be able to apply to any churn problem: using a single row per entity versus using multiple rows per entity. An entity could be an employee, customer, manufacturing equipment, patient, etc.

First, let’s define four concepts that are common among the two approaches outlined in this playbook.

<table>
<thead>
<tr>
<th>Concept</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Modeling Period</strong>:</td>
<td>This is the total time (in days/months/years) that your data was collected in. In the examples in Figures 4 and 5, the total modeling period was 5 months i.e., January - May, 2018.</td>
</tr>
<tr>
<td><strong>Observable Period</strong>:</td>
<td>Out of the total modeling period you need to pick a shorter period from which to extract information that will be used to learn the characteristics of the entity that is likely to churn or not likely to churn. Let’s call that the observable period. In Figures 4 and 5, the observable period ranges from January - February, 2018.</td>
</tr>
<tr>
<td><strong>Operationalization Gap</strong>:</td>
<td>Usually predicting churn for tomorrow may be too late to allow for taking any action on that prediction. We recommend you leave a gap that is big enough for the relevant stakeholders in your use case to be able to execute the appropriate churn prevention strategy. The exact size of this gap depends on your use case. In the examples in Figures 4 and 5, we chose the month of March as our operationalization gap.</td>
</tr>
<tr>
<td><strong>Target Period</strong>:</td>
<td>This is the period from which you create your target variable. A target variable is the variable you want AI to be able to predict. In the examples in Figures 4 and 5, we defined churn to be “an employee will churn in 2 months”. Consequently, we used the last 2 months of the total modeling period to extract the values for the target variable. You can use words such as churn/no churn as the two values in the target variable. You could also convert those words into binary values 1 and 0 to denote churn and no churn, respectively.</td>
</tr>
</tbody>
</table>

Figure 3: Using DataRobot to identify individuals with high probability of churn.
Now, let’s explore what is different. The two approaches we describe here differ in how you choose to represent the information extracted from the observable period i.e., the characteristics of the entity. This information will be used to predict the target variable.

### Single Row Per Entity

Here you capture all data for a given entity at the end of the observable period and record it in one row. Figure 4 shows how this could be achieved for an employee churn use case. Each row in the data stores the cumulative information for each employee. The features in the training data will be engineered based on information available up to February 2018, the end of the observable period.

Single row per entity approach is useful if you have a high volume churn problem where you have a lot of data available for model building. It is also the recommended approach for anybody who is new to churn modeling because it is simple and fast to implement.

![Figure 4: DataRobot uses cumulative information for each staff during the observable period to learn to predict the future.](image-url)
Multiple Rows Per Entity

Alternatively, you can represent an entity by multiple rows extracted from the observable period. One way to do this is illustrated in Figure 5 where we record data from the observable period such that each row of data represents an employee’s information for one month giving each employee multiple entries in the data set.

The multiple rows per entity approach is more complex than the single row per entity approach. It tends to be used in use cases where there isn’t enough data to model churn. Breaking the data extraction period into months can help increase the amount of data available. You should also consider this approach if you are interested in capturing emerging trends over time.

Figure 5: DataRobot uses monthly information for each staff during the observable period to learn to predict the future.
What Data is Best Suited for an AI Churn Solution?

The data that is best suited for a churn model depends on the domain. We recommend you incorporate advice from domain experts (your HR team, legal department, onsite managers, marketing department, etc.) and collect features from various sources. Based on the knowledge obtained from these experts, grow a tree of ideas about variables that could potentially cause churn. In Figure 6 we demonstrate this using the employee churn problem.

Take this opportunity to get feedback from your stakeholders on what features they can act on and those that they cannot (See Figure 7). Make sure your data has enough variables the business can use to execute a successful churn prevention plan. You don’t have to eliminate variables that the business can’t operationalize but are useful for prediction. If you have access to those, include them in your variable list as they may enhance the model’s performance.

One thing to remember in the context of employee churn models is that many countries have rules in place that protect employees against discrimination based on demographic characteristics. Even if gender or age are predictive, clear their use with your Legal Department to ensure the models are ethical and compliant with legal guidelines (in the US, these guidelines come from the EEOC).
Tips for Better Churn Predictions in DataRobot

A standard model building approach within DataRobot will follow the steps outlined in Figure 8: import your data into the DataRobot platform, configure your modeling settings, start Autopilot, and make predictions.

1. Submit Data
   - Submit labeled data indicating a churn/non-churn state
   - Use data about employees regardless of whether or not they have churned

2. Configure Settings in DataRobot
   - Configure settings in DataRobot
     • Target (churn flag)
     • Partition
     • Group ID

3. Autopilot
   - DataRobot automatically builds a model

4. Make a Prediction
   - Make a prediction!
     • Batch prediction
     • Model deployment

In this section we are going to walk through a few things to watch out for when executing this use case using DataRobot.
Use DataRobot Model Output to Seek Feedback from Domain Experts

While modeling can reveal valuable new insights and relationships in your data, it is critical that you continue to include domain experts so they can inform you whether these insights make sense.

One way to achieve that is by showing business stakeholders the feature impact graph (Figure 9). Ask if they agree with the ranking of features. Is an unexpected but convincing feature highly ranked? Should you invest time to explore the possibility of a new idea? Sometimes an unconvincing feature gets a high rank and/or seems to influence the predictive capability of the model too strongly (See Figure 10 for an example). Find out from domain experts what might be driving this result. Do you need to remove some features due to feature leakage?

Figure 9: Feature impact graph depicting multiple variables influencing the performance of the model.

Figure 10: Example feature impact graph that depicts one variable having a strong influence on the model compared to all other variables.
Prediction explanations are another model output that business stakeholders can react to. The prediction explanations are available under the "Understand" tab (See Figure 11). Confirm that the explanations for a given individual’s churn probability mostly align with the experts’ knowledge and experiences. For instance, Figure 11 shows that being a Sales Representative had the biggest influence on why the model predicted a high probability of churn for several employees. Does this align with what the HR department is already observing?

![Prediction Explanations](image)

**Figure 11: Prediction explanations from one of the churn models.**
DataRobot also provides a Feature Effects graph, which shows the effect of each individual feature on the target feature. It utilizes a partial dependence plot, which shows how changing the values of a given feature, while keeping all others the same, changes the probability of churn. Figure 12 is a representation of this. It shows that employees who are Sales Representatives have a much higher probability of churn relative to other job roles. Likewise, Figure 13 shows that keeping all other variables the same, employees with less than 15 years of total working years are more likely to churn than those who have more than 15 total working years. Do these two features behave as expected by the experts?
Assess Model Accuracy

While there are many ways of assessing a model’s accuracy, we recommend you pay close attention to the DataRobot Leaderboard, Lift Chart, and ROC curve. Dig deeper into models that are ranked higher up on the Leaderboard. Specifically, look at those that show top optimization scores consistently across the validation, cross-validation and holdout data sets (See Figure 14).

The ROC tab in Figure 15 allows you to check how well the prediction distribution captures the model separation. You can use the provided confusion matrix to explore the appropriate threshold for separating the positive group from the negative group. For our purpose, the positive group contains entities that are predicted to churn, while the negative group contains entities predicted not to churn. The lift chart in Figure 16 shows how well a model segments the target population and how capable it is of predicting the target, letting you visualize the model’s effectiveness.

Figure 14: DataRobot leaderboard showing a set of models ranked by their LogLoss scores.
Figure 15: DataRobot ROC curve tab depicting a number of concepts useful for evaluating the accuracy of this model.

Figure 16: DataRobot lift chart tab showing that this model is better at discriminating target values in the upper bins compared to the lower bins.
Use Group Partitioning

If you choose the Multiple Rows per Entity (Figure 5) solution, then make sure you set up Group Partitioning by entity ID. This ensures that information about a given entity is not fragmented across multiple partitions during training. In Figure 17a if stratified sampling or other methods were used to divide this data, rows 1-3 could be used for training and row 4 for validation. This would result in Target Leakage.

<table>
<thead>
<tr>
<th>Date</th>
<th>Name</th>
<th>Gender</th>
<th>Age</th>
<th>Dept.</th>
<th>Months of Service</th>
<th>SV</th>
<th>SV Flag</th>
<th>Office</th>
<th>Hourly Pay</th>
<th>...</th>
<th>Churn</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018/1</td>
<td>A</td>
<td>Male</td>
<td>18</td>
<td>Support</td>
<td>25</td>
<td>Nakajima</td>
<td>0</td>
<td>Tokyo</td>
<td>2250</td>
<td>...</td>
<td>F</td>
</tr>
<tr>
<td>2018/2</td>
<td>A</td>
<td>Male</td>
<td>18</td>
<td>Support</td>
<td>26</td>
<td>Nakajima</td>
<td>0</td>
<td>Tokyo</td>
<td>2250</td>
<td>...</td>
<td>F</td>
</tr>
<tr>
<td>2018/1</td>
<td>B</td>
<td>Female</td>
<td>37</td>
<td>Service</td>
<td>37</td>
<td>Naoi</td>
<td>1</td>
<td>Tokyo</td>
<td>2700</td>
<td>...</td>
<td>T</td>
</tr>
<tr>
<td>2018/2</td>
<td>B</td>
<td>Female</td>
<td>37</td>
<td>Service</td>
<td>38</td>
<td>Naoi</td>
<td>1</td>
<td>Tokyo</td>
<td>2700</td>
<td>...</td>
<td>T</td>
</tr>
</tbody>
</table>

Figure 17a: An example of assigning an entity to more than one partition.

Advanced Options

Partitioning | Smart Downsampling | Time Series | Feature Constraints | Additional
---|---|---|---|---
Select partitioning method:
- Random
- Partition Feature
- Group
- Date/Time
- Stratified

All rows with the same single value of the selected feature are in the same partition. Each partition can contain multiple values of the feature.

Group ID feature:
This must be a feature from the dataset.

Figure 17b: The group partition method for ensuring that all rows for each distinct entity end up in the same partition.

DataRobot allows you to specify a variable used to assign rows to partitions during training (See Figure 17b). You do this before you press the “Autopilot” button by going into Advanced Options and clicking on the “Group” button under the “Partitioning” tab. In the illustration in Figure 9a, you would specify “Name” as the Group ID feature when setting a group partition inside DataRobot. With this feature, all data for staff person B will end up in the same partition rather than becoming segmented into two different partitions as illustrated in Figure 17a.
In summary, churn can be a costly problem. Modeling churn can be very time consuming. DataRobot provides a way to use AI to model churn speedily and efficiently. In this playbook we have outlined a number of questions to think about before and during the process of using AI to model churn. We have also described two solutions that could be used to implement an AI churn model and outlined a number of tips for better churn modeling in DataRobot. The key take-away from all this is to make sure you keep all key stakeholders in the loop during the whole churn modeling process, make sure your model makes sense before you deploy it, and reach out to your DataRobot account team for help if needed.