“Voluntary Attrition Risk Model” Module Transcript

Chapter 1

Intro, Topics Covered, & Learning Outcomes

Hello. My name is Hayley and I'm on the One AI team here at One Model. In past modules, we covered machine learning foundations and model-building best practices.

In this module, we will bring all of that together and go through building a voluntary attrition risk model using a recipe in One AI. This is our most popular model as well as a great starting point for beginners. We will cover an overview of the voluntary attrition risk model, important considerations before beginning the model building process, step-by-step instructions for building a voluntary attrition risk model in One AI with a recipe, and then we will close with the insights that can be drawn from a voluntary attrition risk model.

After completing this module, you will understand what a voluntary attrition risk model is and its purpose, gain insights into key considerations before starting the model-building process, know the step-by-step instructions to build these models in One AI, and be equipped to interpret insights from voluntary attrition risk models and understand how to go beyond the predictions towards drivers and other interesting insights.

Chapter 2

Overview & Purpose

Section 2 - Overview and Purpose

The voluntary attrition risk model, which is also known as a flight risk model, predicts the likelihood of an employee voluntarily terminating from your organization within a specified time period.

By analyzing attributes such as employee demographics, job satisfaction, performance metrics, engagement, and compensation, this binary classification model identifies individuals at higher turnover risk. The two possible outcomes that the model will classify employees into are voluntary termination and no voluntary termination, which you can think of as voluntary retention.

Instances are individual employees typically identified by their unique person IDs from the employee table.

The model creator selects the model population, the prediction period, and the input features used to train the model to make its prediction during the model creation.

These models are commonly applied to retention planning, resource allocation, succession planning, and cost reduction.

Overall, flight risk models help organizations understand workforce dynamics, specifically what is driving employees' decisions to voluntarily stay or go. These insights are used to develop proactive strategies to retain valuable talent.

Chapter 3

Important Considerations Before Beginning the Model-Building Process

Section 3 - Important Considerations Before Beginning the Model-Building Process

Before creating a flight risk model, there are some important factors that should be carefully considered to ensure its effectiveness and relevance to the organization's needs.

First, clearly outline the goals of the model, such as reducing turnover rates, improving employee retention, or optimizing resource allocation. Determine if the model aims to confirm hypotheses, conduct exploratory analysis, or replicate a previously built model outside of One Model. This will guide your approach. For example, exploratory models might use various input features and historical data lengths, while a hypothesis-confirming model would use specific features believed to drive turnover.

Next, ensure you have accessible and reliable data on employee demographics, performance metrics, job satisfaction, and turnover history. Verify if the necessary data sources are integrated and validated in One Model. A solid baseline is crucial for starting the model.

Consider if there is sufficient historical data on voluntary turnover to make accurate predictions. For example, if your organization has only a 1% voluntary turnover rate, there likely won't be enough instances of voluntary turnover in the historical data for the model to effectively learn to make accurate predictions for both retention and termination classes.

And finally, we must ensure you have the resources to both build and maintain the model. One AI simplifies the model-building and visualization process, but ongoing maintenance is essential. This includes rerunning and refining the model with new data, creating effective visualizations and storyboards, incorporating feedback from stakeholders, and monitoring performance. The maintenance effort will vary based on whether the model is for long-term use or for a one-time analysis.

Next, we will move over to One Model for a demo on building flight risk models in One AI.

Chapter 4

Building a Voluntary Attrition Model

Section 4 - Step-by-step Instructions on How to Build a Voluntary Attrition Model in One AI

To get started, click 'One AI' in the main ribbon menu. You need access to the 'CanAccessOneAIMenu' application access role to access this page. This can be granted by your site administrator.

This page is where you can edit or create machine learning models. If your organization has previously created any models, they will be listed here. To add a new model, click 'Add Machine Learning Model' in the upper right hand corner.

Then give your model a unique display name in the display name field. Once the model has been created and saved, the display name cannot be changed, so please choose wisely.

Under the 'Using Data From', click 'One AI Recipe' and then 'Configure One AI Recipe'.

This will bring you to the One AI Query Builder screen. Under "What are you interested in predicting?", select 'Voluntary Attrition' from the list of available recipes.

Once selected, the problem statement is populated with the information we need to provide to build the model. One AI simplifies most settings by automatically selecting defaults, so most of the configuration focuses on data framing.

In step one, select the column or metric to define voluntary terminations.

Only use a column if it was explicitly created by your data engineer for these purposes.

Otherwise, use a terminations metric since it will be much more straightforward.

To select a metric, you can either type in this box or you can scroll through the list to find the metric you're looking for.

You can filter the termination metric as needed. For example, you could use the Event Reason dimension to exclude retirements from your analysis like so.

Because I selected a voluntary terminations metric, this model will now learn from past voluntary terminations to make its predictions.

Next, we will select the Headcount Population, which is the group of employees for which One AI will predict voluntary attrition risk. You can use a headcount end of period metric or a specific headcount metric like managers, engineers, or California if that is the group that you are targeting in your model.

Next, you will choose a unique identifier for each instance in the model. Typically person ID from the employee table is the best choice here, but employee ID, worker ID, or any other unique ID will work. One AI will verify the IDs' uniqueness for the model. Next, you will select a population date, which anchors the query with the predict frame and test train dataset offset from it. Common choices are 'Today', the 'End of Last Month', or a static date. Avoid dates too far in the past since we already know the outcomes for those.

And lastly, you can use filters to refine your headcount population.

Instead of running the model on the entire organization, segment it by factors like department, country, or job level. This is best practice because motivations and drivers may vary drastically by group.

For example, I'm going to use a managerial dimension to grab only managers for this model.

In the next step, you will select how far into the future you want to predict from the population date. Without a defined time frame, everybody would be predicted to terminate, which isn't very useful. One year is the most common choice because it's helpful to know who will terminate relatively soon, but you can adjust this to fit your needs.

Then you will choose how much historical data you want to use to train the model. The previous step also determines the increments for how far back in time we will go to train the model. For example, if you selected 1 year here, then training intervals are in 1-year increments.

Meaning, down here, 1 equals 1 year or 2 would equal 2 years, etcetera. For flight risk models, we recommend starting with 1 year and increasing to 2 or 3 years to see if it improves performance. Usually, less is more as training on all available data may not be beneficial due to changing organizational dynamics and employee behaviors.

This next step, the override step, is optional, but highly recommended as it will simplify reading your result summary and storyboards.

In this model, null values or 0s indicate no voluntary termination, while values greater than 0 indicate voluntary termination.

Click 'Load Target Metric Values' to set your override labels.

You can name the groups whatever you'd like, such as 'No Termination' and 'Termination'.

Just ensure that there are only two unique labels. Without this step, predictions will be sorted into 1s and 0s, making results and the storyboards difficult to interpret for end users who don't know which label corresponds to which value.

In the next step, we will select core attributes, which are the input variables that form the dataset the model learns from and uses to make predictions.

We want to aim for balance here. We want enough attributes for robust predictions and great learning data, but not so many that the results become difficult to interpret.

There are 4 defined scopes available for you to choose from.

First, we have 'None'' where only the unique identifier is included automatically. You must manually add all other attributes. Next is 'Narrow', which includes all attribute columns from the table where the unique identifier originated, which is the employee table in this model.

Next is 'Balanced', which includes all attribute columns within one join of the employee table. And finally, we have 'Broad', which includes all attribute columns with a possible join path to the employee table.

We recommend starting with narrow and moving to a balanced scope as you become more comfortable interpreting the results summary and EDA report. Broad can include attributes that may be harder to relate to the model, which can degrade model interpretability.

If none of these scopes fully meet your needs, you can manually include or exclude columns using the check mark button to include additional columns in the excluded section or the X button to exclude columns from the included section.

You can use the search bar to find specific columns to include or exclude or scroll through either section to find the columns you're looking for. If a table is in the unavailable section, it means using it would cause duplication or there is no join path. The reason is explained in the reason column. In some cases, you can work with your customer success lead and data engineer to make unavailable tables available.

Next, you will select the generative attributes that you want to use in your prediction. Use the checkbox to the left of each attribute to include it as an input variable, or you can select all by checking the box at the top.

This step is optional but highly recommended, as generative attributes often enhance model performance and robustness.

You can also create, edit, and delete generative attributes in this step. For more information on how they work and how to build them, check out our module on generative attributes.

In the next step, click 'Generate Data Statistics' to verify that all of your previous selections are valid. This helps identify errors that would cause the model to fail while running.

You will also receive valuable data exploration information about the train and test dataset such as row count, standard deviations, and min and max values. This helps identify columns you may want to exclude from your dataset.

We received the green light "Success" so we can proceed. If One AI finds a problem, it will display action needed with details about the errors in your configuration.

This step is also optional, but highly recommended.

In the final step, you can download the train and test dataset, predict dataset, or both that were generated by this recipe. This is useful for analyzing the data in external tools or for investigating errors encountered in the verification step.

Before saving, you should review your selections in each step. When you click the save icon in the upper right hand, you will return to the machine learning model screen. Here, you can manually configure model settings, refer to our modules on advanced configuration and global settings if interested.

At minimum, you should generate SHAP in order to use this data in storyboards.

Once complete, scroll to the bottom and click 'Create'. If you do not click create and navigate elsewhere in One Model, you will lose your work from the One AI query builder screen.

Following these steps, you can now create predictive voluntary attrition risk models in One AI. The next step is to run your model, evaluate your results, refine your model based on the results, and then deploy and share the insights.

Chapter 5

Insights from Voluntary Attrition Risk Model

Section 5 - Insights to be Drawn from a Voluntary Attrition Risk Model

Voluntary attrition risk models provide several insights. First and foremost, you will receive data on the model's predictions made at the instance level. These can and should be aggregated to larger groups for better accuracy.

Without creating any visualizations in Explore or Storyboards, you can view individual predictions by dataset ID in the Results Explorer.

You can also use storyboards to analyze individual employees or groups examining which features drive their behaviors, how it compares to the model dataset mean, and an explanation of these drivers.

You can also analyze correlation data from the EDA report to see how your input features interact. This can help you understand how your data is related to each other and uncover interesting patterns.

You can also understand and visualize the factors influencing employees to either voluntarily terminate or remain employed. If SHAP is enabled in your model and incorporated into storyboards, you can view the top drivers for both classes in order of importance.

These charts can be filtered down to any groups within your model population.

You can also analyze risk by groupings within your model population. Any dimension that can be used to filter the headcount metric can break out your model population to identify which groups have the highest attrition risk.

For example, using the Org Unit dimension, you can see here that Management Accounts and Technical Product Managers have the largest population of high risk employees for voluntary termination.

These are just a few examples of insights that can be drawn from flight risk models. Many more are available depending on your objectives and what questions you are using the model to try to answer.

Chapter 6

Conclusion & Thanks

In conclusion, mastering the concepts and techniques covered in this module empowers you to effectively use voluntary attrition risk models. By understanding model building, interpreting insights, and applying findings to real world scenarios, you can shape talent management strategies in your organization. Proactively identifying and mitigating attrition risk helps cultivate an engaged workforce, optimize resource allocation, and drive sustained success and competitiveness. Happy modeling!