# “One AI Results Summary” Module Transcript

## Chapter 1

### Intro, Topics Covered, & Learning Outcomes

Hi all. My name is Hayley, and I'm on the One AI team here at One Model. In past modules, we've discussed topics like exploratory data analysis(EDA), performance, and ethical AI, highlighting the Results Summary report as a critical tool for interpreting machine learning models. This module will examine the Results Summary and demonstrate how it complements the EDA report to help you fully understand your models.

The concepts we will cover build upon several previous modules. They will be linked in the description under prerequisites for easy access so you can refer to them if you need additional information. In this module, we will cover an overview of the Result Summary report, navigation to the Results Summary in One Model, the key takeaways from each section of the report, and deciding whether to deploy or ignore models based on the Result Summary and EDA report.

After completing this module, you will recognize the importance of the Result Summary for promoting transparency and understanding model configuration and performance. You will gain the ability to interpret and modify the model's configuration based on insights from the Estimation Details section. You will know how to evaluate feature importance using techniques like feature ranking, SHAP visuals, and VIF scores. And you will use the regression or classification performance reports to make informed decisions about model deployment or further refinement.

## Chapter 2

### Overview & Significance of Results Summary

Section 2 - Overview and Significance

The One AI Results Summary report provides a detailed overview of the configuration and performance of machine learning models. It includes information on estimator configuration, feature selection, upsampling, probability calibration, prediction details, feature analysis, model performance, and relevant messages or warnings.

This report is essential for making informed decisions about manual overrides, advanced configuration, and ultimately, model deployment.

Understanding the model's configuration and selected features is really important for determining necessary adjustments before wider sharing with end users and stakeholders.

This report offers transparency and insights into the machine learning process, enabling you to deeply understand and optimize your models.

If you've used other machine learning and predictive analytics human resources tools in the past, you might not be familiar with this type of report as many models are proprietary and vendors consider sharing such details as revealing their "secret sauce". However, One Model prioritizes full transparency in machine learning and considers this report essential for your success. The Results Summary paired with the EDA report provides comprehensive insights into your model, supporting interpretability, knowledge sharing, collaboration, compliance, and ethics, as well as facilitating audits.

## Chapter 3

## Navigating to the Results Summary in One Model

Section 3 - Navigating to the Results Summary in One Model

A unique Results Summary report is automatically generated by One AI for each run of a machine learning model that is in a pending, ignored, deployed, or deployed and persisted status.

Runs that error or are canceled before completion will not have a Result Summary report because it can't be generated for incomplete models.

To access this report, click on the One AI tab in the main ribbon menu and navigate to the machine learning model you wish to view. Click the 'Runs' button and then select the status label for the iteration you are interested in. This window will automatically open to the EDA report tab, and it may take a moment to load.

You must view the EDA report first before the entire Result Summary will display because you can't fully interpret this report without information from the EDA report. From the EDA report, you can click 'View Results Summary' or the Results Summary tab.

## Chapter 4

## Key Takeaways of the Results Summary

Section 4 - Key Takeaways of the Results Summary Report by Section

## Chapter 5

## Estimation Details Section

Estimation Details

The first section of the Results Summary report is the Estimation Details, which provides a comprehensive overview of the model's configuration. This section offers transparency into the model setup, helping users understand the underlying algorithms and settings used. If you are not satisfied with the configuration selected by One AI, you can manually reconfigure anything that is displayed in this section in the model settings or One AI Query Builder. Let's head over to the demo site to take a look.

The left-hand column details the estimator configuration starting with the selected estimator, either a regressor or a classifier, followed by the parameters and hyperparameters details below.

The center column begins with feature selection information for both the filter and wrapper steps of dimensionality reduction. Models in One AI are typically presented with numerous features, sometimes hundreds. This section shows how the model removed the non-important features so that the model was only being presented with predictive features that increased the model performance.

Below that, you'll find upsampling and probability calibration details indicating which method, if any, was selected along with specific details about that method. The estimator dimensionality reduction, upsampling, and probability configuration can all be manually configured in the model settings if you wish to make changes from One AI selections.

Finally, the right hand column varies slightly depending on whether a regression or classification model was run. For classifications, it provides a quick overview of the prediction details, including the metric or column being predicted, which is the binary classification target, the number of predictions in each class outcome, and the positive label class. The target and positive label can be configured from the One AI Query Builder. For regressions, this column simply provides the regression target column.

## Chapter 6

## Feature Analysis

Feature Analysis

Next is the feature analysis section, which provides valuable insights into the importance and impact of features on the model's performance. This section includes various tabs offering different perspectives on feature analysis tailored to meet the diverse needs of users. Let's return to the demo site to see.

First and most importantly, the Feature Importances or Coefficients tab ranks and numerically scores features based on their significance in influencing model outcomes. Understanding which features have the most significant impact allows users to prioritize key variables for further analysis or intervention. The format of these values varies depending on the estimator selected, but generally speaking, a higher score indicates greater feature importance. If a logistic regression classifier was selected, like in this model that I just toggled over to, or a lasso regressor, feature coefficients, which are the weights by which the features are multiplied in the model, are displayed in place of feature importances in a "Coefficients" tab.

If SHAP was enabled before the model was run, the next tabs include the SHAP beeswarm chart, which visualizes feature impact by providing numeric importance values for every feature across individual predictions, and the SHAP average bar chart, which shows the average absolute value of the SHAP values for each feature. We examine SHAP and how to best interpret these charts in the SHAP module.

The final tab, the VIF Scores tab, presents variance inflation factor scores for each selected feature in the train, holdout, and predict sets, indicating the degree of correlation between features. A VIF score of 1 is the lowest, while a score of 8 or more indicates a high degree of correlation with another column. Identifying highly correlated features helps mitigate multicollinearity issues and optimize model performance.

## Chapter 7

## Classification or Regression Report

Classification or Regression Report

The third section of the Result Summary is the classification or regression report, which provides a comprehensive evaluation of the model's performance tailored to either

classification or regression tasks depending on the type of model you have run. This section includes various metrics and visualizations to assess accuracy, precision, recall, F1, and other important performance indicators. For a detailed explanation of these visualizations, refer to section 3 of the "Classification and Regression Model Performance" modules.

Now let's return to One Model to examine this section further. For classification models, the first tab is the classification report, which displays the table with the model's F1, precision, and recall scores. This table provides insights into the model's performance in classifying instances correctly.

Generated against the holdout training set, it helps users understand the model's ability to identify positive and negative instances, indicating whether the model is ready for deployment, sharing, and strategic decision making, or if more configuration and work is needed. You can scroll to the other tabs for additional performance metrics.

For regression models, the regression report is displayed at the top of this section to provide crucial information about model performance. The explained variance score indicates the proportion of variance in the dependent variable explained by the independent variables in the model, offering insights into the model's fit. The mean squared error quantifies the average squared difference between the actual and predicted values, measuring the model's accuracy and predicting the target variable. Together, these metrics help users assess the model's predictive capability, and identify areas for improvement. You can scroll through the visualization tabs for additional performance metrics and visualizations.

## Chapter 8

## Messages & Warnings

Messages and Warnings

The last section of the Results Summary report is the messages and warnings, which aims to enhance transparency about what happens when a model is run. This section displays information about the models created during the model selection process and any setting overrides applied during configuration.

Specifically, it includes which algorithms were tested and scored for model selection. For each algorithm, was it run by default settings, a heuristic, or a user override? And were there any setting combinations rejected because the data fed to the model did not meet certain criteria? Here's an example message from this section.

## Chapter 9

## Deploying or Ignoring Models

Section 5 - Deploying or Ignoring Models

Armed with the knowledge from the EDA report and Results Summary, you can choose to ignore, deploy, or deploy and persist the model using the options in the bottom right of the Results Summary.

Deploying a model pushes the data from the model to tables and storyboards in the front end of One Model. We discuss each of these actions and what they entail in the model deployment module.

At the bottom of the Results Summary, you can also download several files, such as grid search metadata, label predictions, and prediction explanations for further analysis outside of One Model. You can select specific files or download all available files.

## Chapter 10

## Conclusion & Thanks

Understanding the Results Summary report is essential for leveraging the full potential of Machine Learning models in One AI. By exploring the model configuration, feature analysis, and performance assessment detailed in this report, you can gain actionable insights to refine models, optimize predictive capabilities, and ensure compliance and transparency.

This knowledge empowers you to make informed decisions on model deployment and management, fostering a data driven approach that drives strategic outcomes and maximizes the value of your machine learning efforts. Happy modeling!