OneModel Academy

One AI

Refining Models

Machine Learning

Hayley Bresina
One AI Client Enablement

07 Jun 2024

Topics Covered

- Introduction to machine learning model refinement
- Using the EDA & Results Summary reports inform model refinement
- Easy yet powerful methods for refining models in One AI
    - Recipe reconfiguration
    - Create separate models for different parts of the organization
    - Adjust global settings
    - Perform per column interventions

Learning Outcomes

You will:

- Understand the purpose & goals of model refinement in machine learning to prepare for advanced model configuration
- Leverage the EDA & Results Summary reports to determine the appropriate refinement techniques for different models
- Master the configurations & overrides in the One AI tool to refine your models effectively
- Understand their impacts to guide future model refinement efforts

Introduction to Model Refinement

Overview

- Machine learning is iterative; models often need multiple refinements before they’re deployment-ready
- **Model refinement** is the process of improving the performance & accuracy of a ML model after its initial development
    - Involves making small adjustments or modifications to enhance the model’s ability to make accurate predictions or fit to a specific task or dataset
    - **Purpose**
        - Enhance relevance, scalability, & generalization
        - Allows for targeted improvements to align with business goals
        - Increase interpretability

EDA & Results Summary Reports Inform Refinement

EDA & Results Summary Reports

Most refinement occurs after running the model, allowing use of the EDA & Results Summary reports to inform the best-suited refinements for the specific model

- **EDA**
  - Guides adjustments to settings to allow different features to be used by the model
- **Results Summary**
  - Guides adjustments to the configuration of the model & number of features used
  - Provides detailed performance scores to help analyze the impact of the changes
- Make adjustments one at a time to observe the impact on performance & output
- Some changes will decrease performance to improve interpretability & usefulness

Recipe Reconfiguration

Adjust History Training Intervals

Improve Core Attribute Selection

Improve Generative Attribute Selection

Create Separate Models

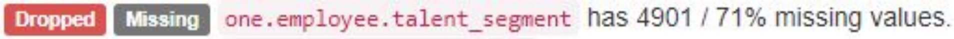
Creating Separate Models

- Highly effective due to varying behaviors across subgroups
  - Focusing on smaller subgroups tailors the model to their unique characteristics
- Employee motivations & behaviors differ based on several factors (location, department, etc.)
  - Consider these differences when grouping your model population into separate models
- Made easy by copying models in One AI

Adjust Global Settings

Global Settings

- The overarching configuration parameters & hyperparameters that affect the behavior & performance of the entire model, not individual columns
    - These are the rules for which columns **One AI automatically drops**
    - The most common reason for a column to be automatically dropped is excessive missing data - determined by the **null drop threshold**
- The **null drop threshold** specifies the percentage of missing values in a column that will trigger its exclusion from the model
    - Default threshold: 0.05
    - EDA report example:



Dropped Missing `one.employee.talent_segment` has 4901 / 71% missing values.

Per Column Interventions

Per Column Interventions

- Making tailored changes specifically for individual columns or features in a dataset
  - Different columns may have distinct characteristics - varying data types, scales, distributions, & degrees of missing data or outliers
  - Allows for tailored preprocessing, cleaning, & transformation by column
    - Optimizes the model's ability to learn & generalize patterns from the data
- One AI offers a few types of per column interventions; we will examine droppability & null filling

OneModel Academy

Thanks for watching!

OneAI