Topics Covered

- Overview & significance of upsampling

- The different upsampling methods & ratios available in One AI

- The default upsampling settings in One AI

- How to configure upsampling for One AI models

Learning Outcomes

You will:

- Understand how upsampling addresses imbalanced model datasets & improves performance by ensuring that the majority & minority classes are represented

- Differentiate between available upsampling methods to determine when to use each method based on the characteristics of your dataset

- Configure upsampling settings in One AI, including selecting appropriate methods & ratios, & understand how to save & apply these settings to optimize your models

Overview & Significance

Overview & Significance

- Also known as oversampling

- Technique used to address the issue of imbalanced model datasets
  - Most One AI classification recipes result in imbalanced datasets

- Involves increasing the number of instances in the minority class to balance the class distribution
  - Helps avoid biased training where the model favors the majority class & neglects the minority class
  - Results in better learning & model performance

OneModel Academy

One AI

Upsampling Ratios & Methods

- Classes do not need to be perfectly balanced
  - Focus should be on reducing the imbalance sufficiently so that the model can learn effectively from all of the classes
  - Degree of balance is determined by the **upsampling ratio**
  - Goal is to strike a balance that improves model performance without overfitting
- Types of **upsampling methods**
  - Random oversampling
  - Synthetic data generation
    - One AI only offers synthetic data generation methods

Methods & Ratios Available in One AI

Upsampling Methods & Ratios in One AI

- Methods
  - ADASYN, SMOTE, & SMOTENC
- Ratios
  - Numerical values
    - Decimal value for the percentage of minority class/majority class desired
    - Ex: 1.0 upsamples the minority class to achieve a 50/50 balance with the majority class
  - Text Strings: auto, minority, not minority, not majority, all
- You can select multiple methods & ratios. One AI will try each combination & select the one that results in the best model performance & fit

Default Settings in One AI

Default Upsampling in One AI

- Upsampling method: One AI tries SMOTE & None

- Upsampling ratio (if SMOTE is selected): 'auto'

  - Upsample all classes except the majority class

Configuring Upsampling in One AI