

09/11 – 13/11

KLARA, YORAN, OSCAR

---

**NANO – WEEK 10**

# OBJECTIVE

- ▶ VSParticle's image analysis software
- ▶ Create a model that predicts the thresholding method

VSparticle

Catalogue Current run New run Debug

B4B044k97.tif, started 2020-08-10 13:45:03

Run ID: 02c04f0, User ID: 1ab9032

Upload

Backend

Import

Scalebar

Crop

Resize

Invert

Align

Level

Normalise

Scar

Smooth

Threshold

Fill

Separate

Remove partial

Remove small

Remove large

Plot outline

Available alternatives:

»

In the threshold step, the greyscale image is converted to a black-and-white (background/not background) image. This is done by checking if the value of each pixel is below or above a certain threshold value.

Many threshold methods exist to automatically determine the threshold value. Depending on the input image, different methods are best suited. The tool attempts to identify the best method automatically.

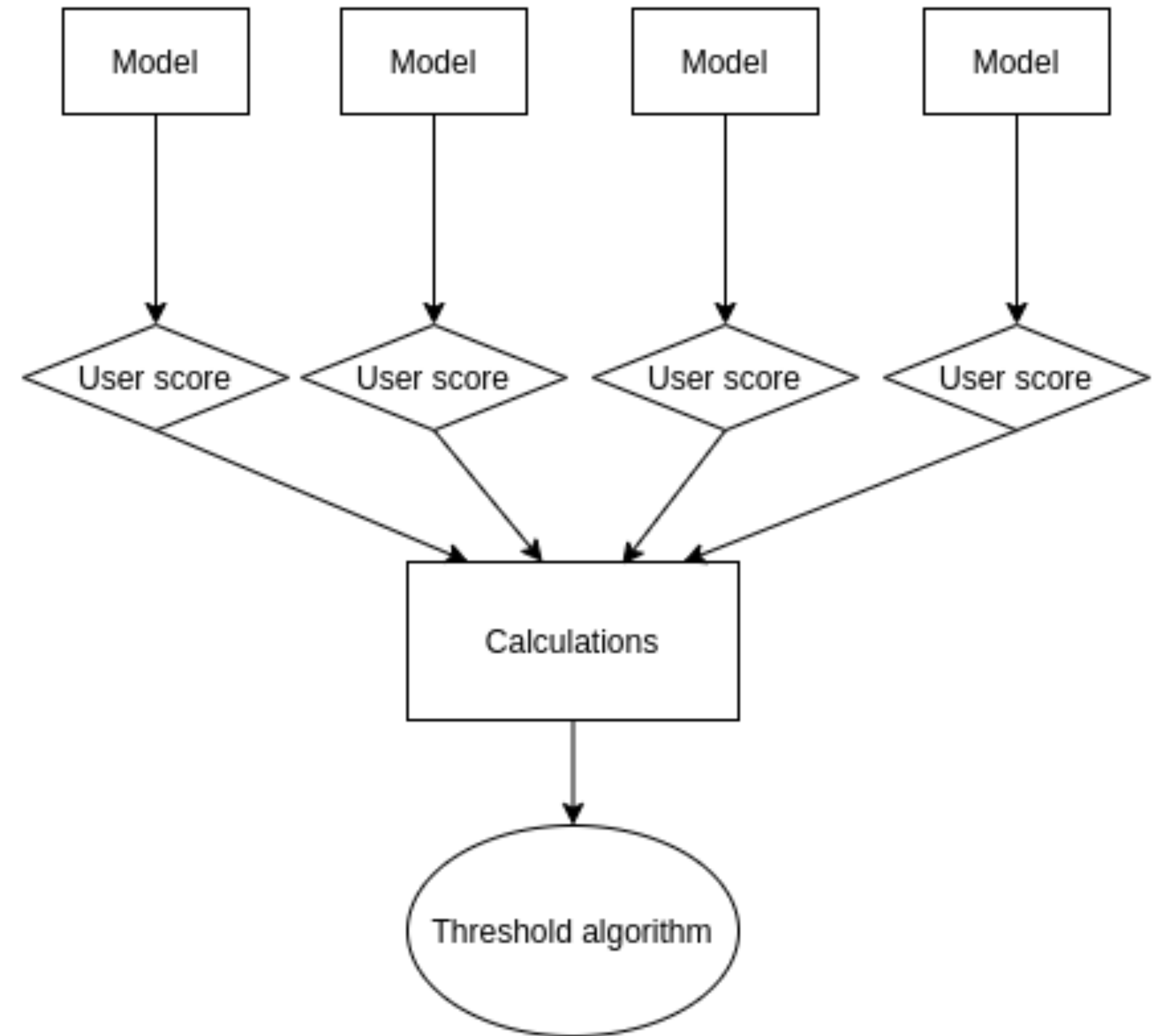
Input parameters:

- threshold = None
- threshold\_method =

## DATASET

Contains:

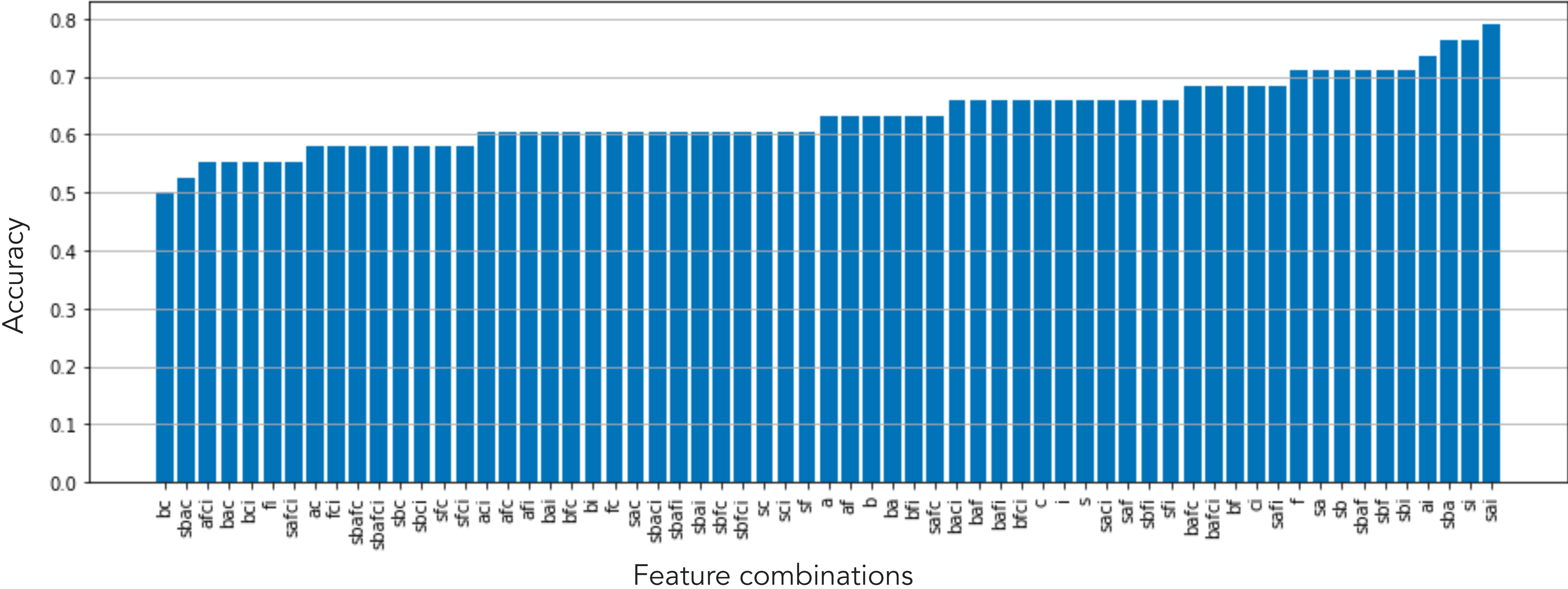
- ▶ Used threshold algorithm
- ▶ Used algorithm feature scores
- ▶ User score



# MODEL: BINARY DECISION TREE

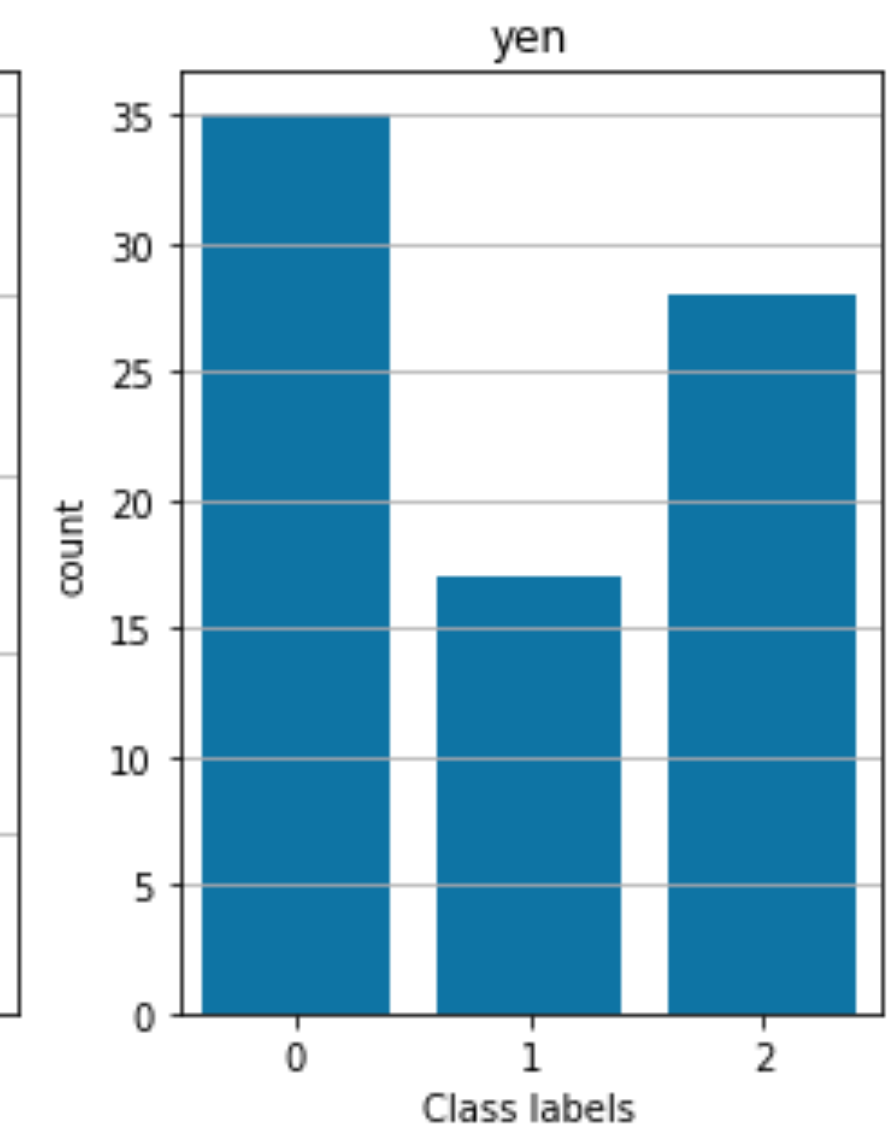
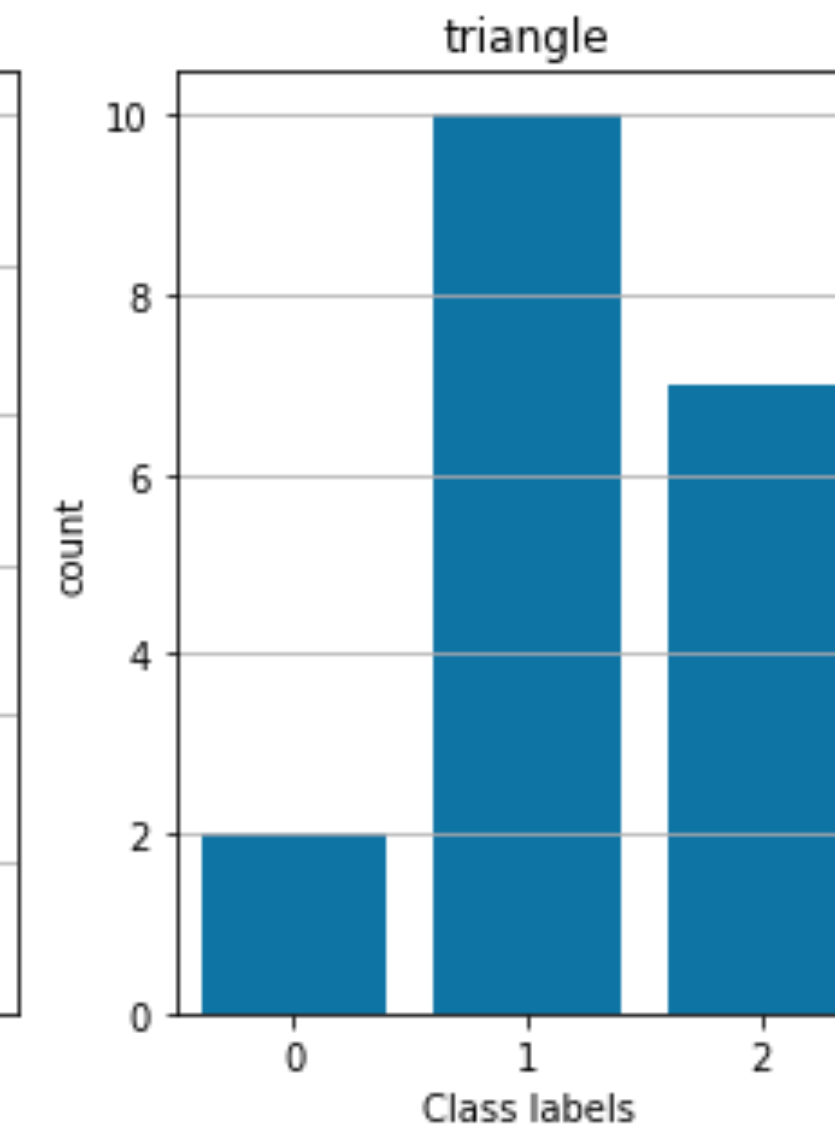
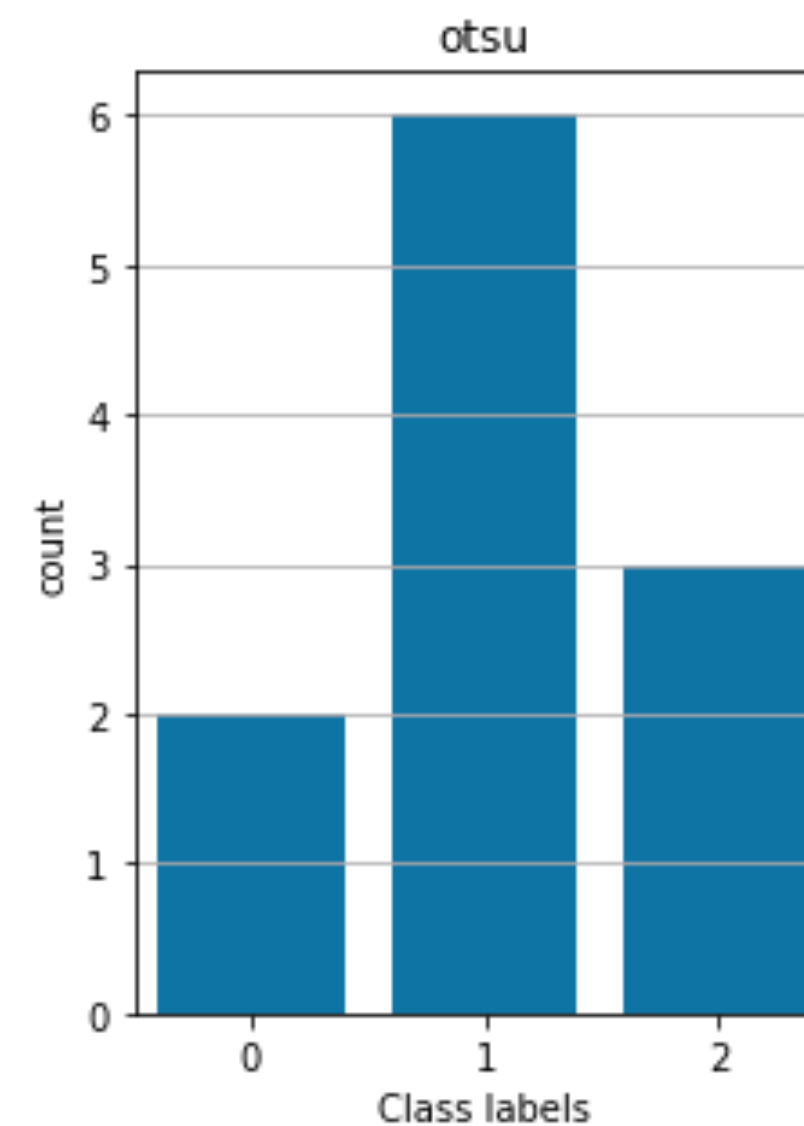
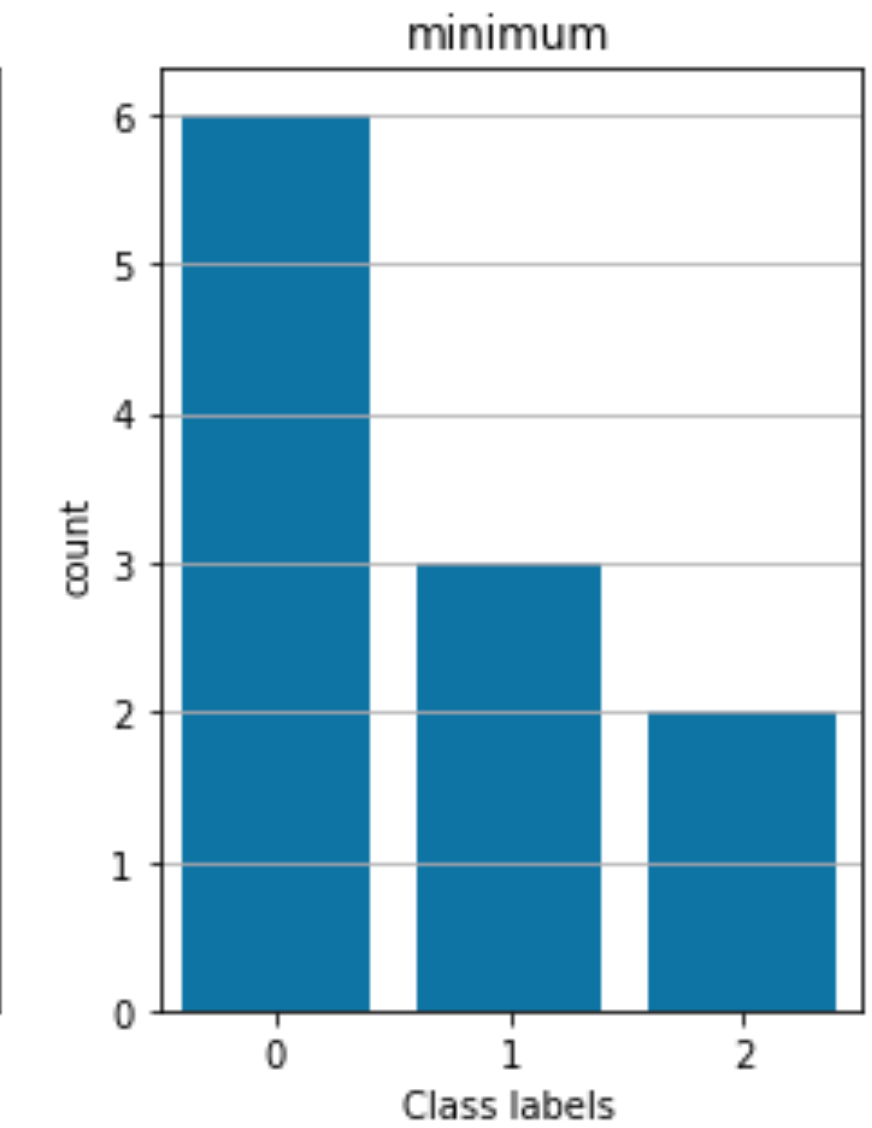
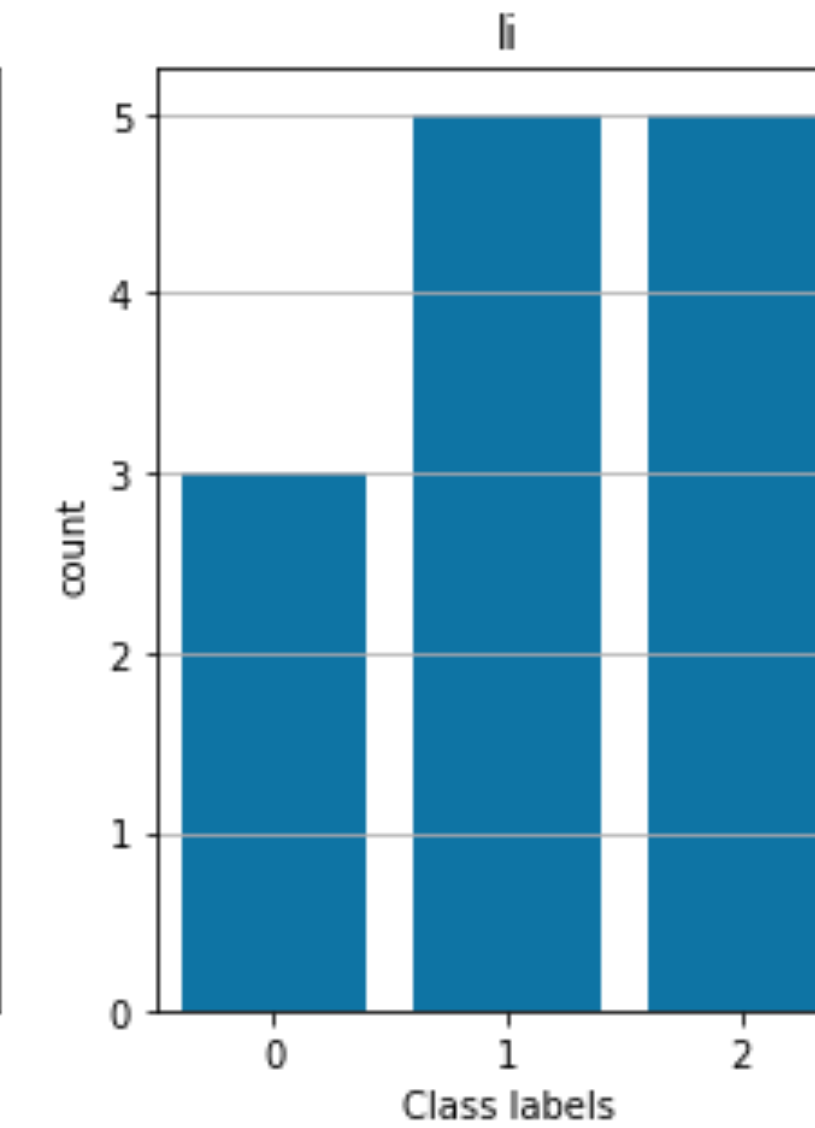
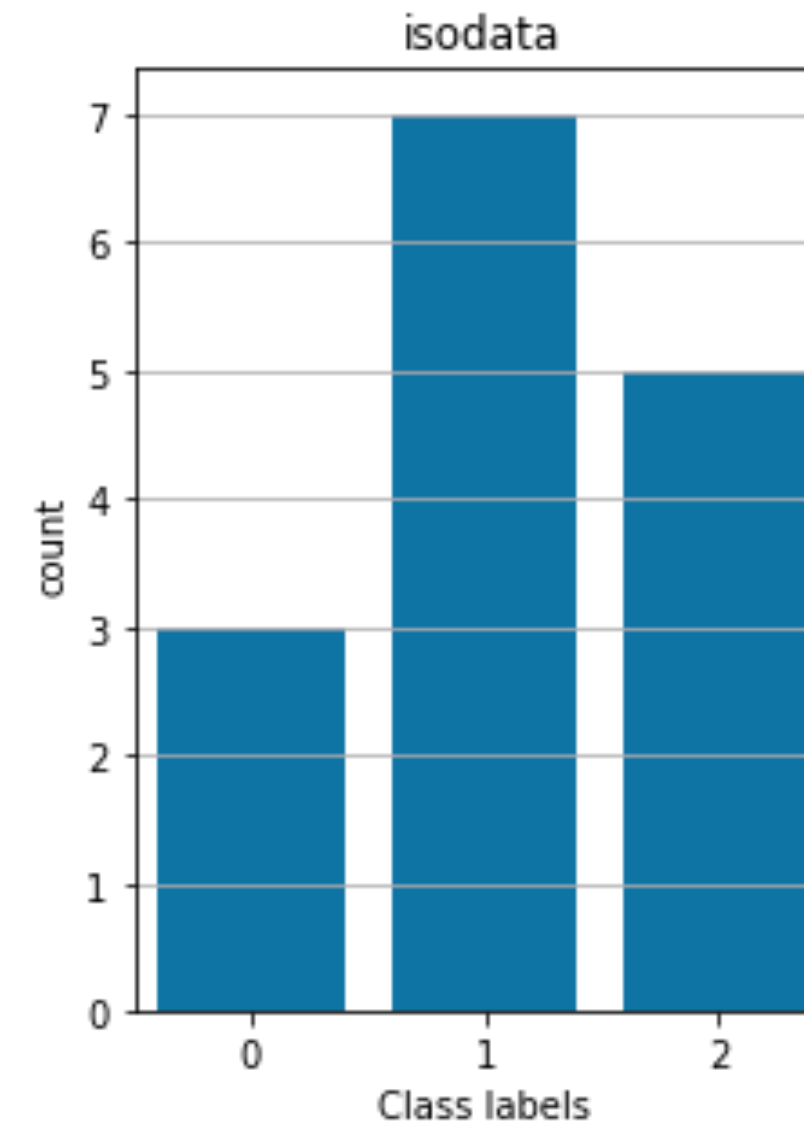
- Separation, Area spread, Intensity

Decision Tree Feature Combinations



# IMBALANCED DATA SOLUTIONS

- ▶ Random undersampling
- ▶ Random oversampling
- ▶ Cluster-based oversampling
- ▶ (M)SMOTE
- ▶ Non-linear classes

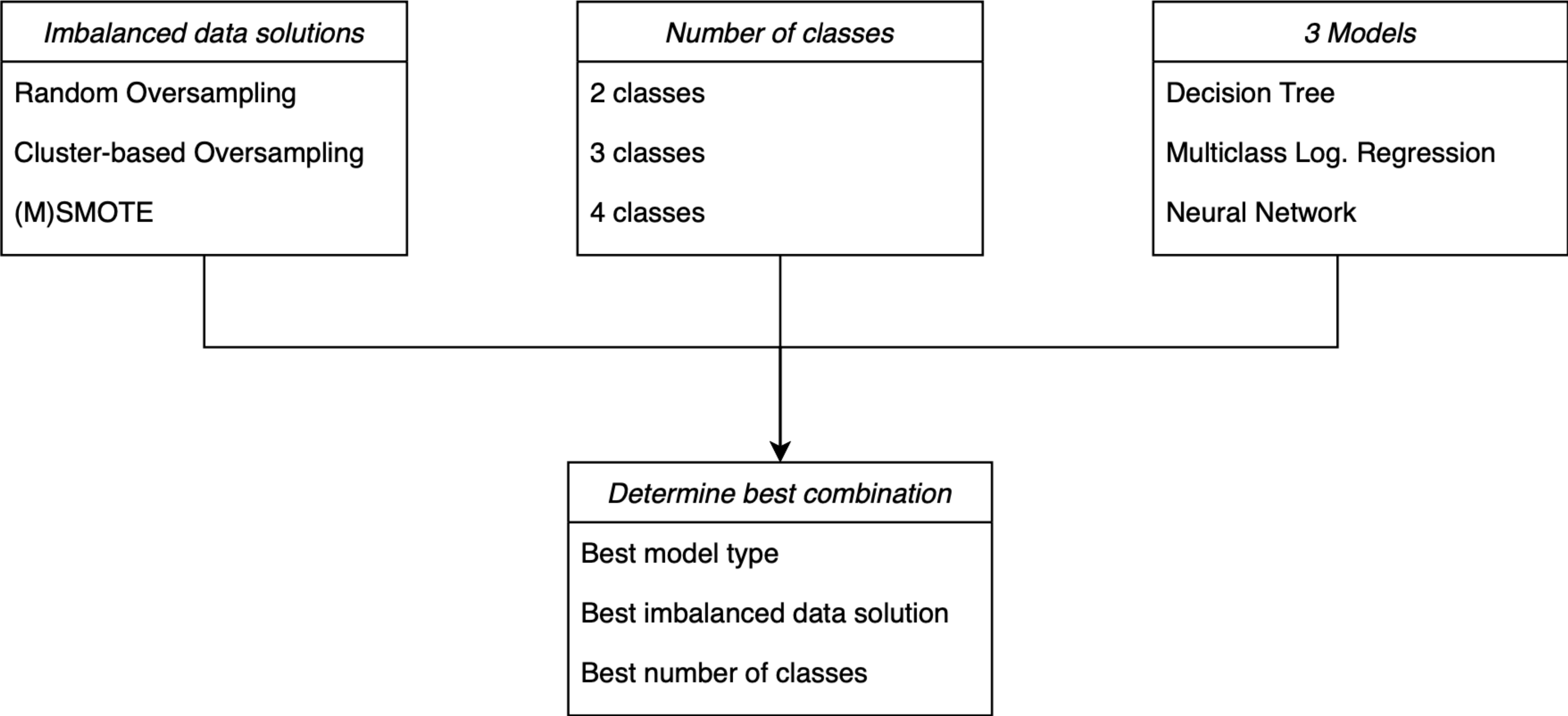




## IMBALANCED DATA SOLUTIONS

- ▶ Random undersampling → Not applicable
  - ▶ Random oversampling
  - ▶ Cluster-based oversampling
  - ▶ (M)SMOTE
- Determine best one
- ▶ Non-linear classes

APPROACH



APPROACH

Possible outcomes:

Random Oversampling 2 classes Decision Tree	Random Oversampling 2 classes Neural Network	Random Oversampling 2 classes Multiclass Log. Regression
Cluster-based oversampling 2 classes Decision Tree	Cluster-based oversampling 2 classes Neural Network	Cluster-based oversampling 2 classes Multiclass Log. Regression
(M)SMOTE 2 classes Decision Tree	(M)SMOTE 2 classes Neural Network	(M)SMOTE 2 classes Multiclass Log. Regression
Random Oversampling 3 classes Decision Tree	Random Oversampling 3 classes Neural Network	Random Oversampling 3 classes Multiclass Log. Regression
Cluster-based oversampling 3 classes Decision Tree	Cluster-based oversampling 3 classes Neural Network	Cluster-based oversampling 3 classes Multiclass Log. Regression
(M)SMOTE 3 classes Decision Tree	(M)SMOTE 3 classes Neural Network	(M)SMOTE 3 classes Multiclass Log. Regression



## RANKING PROBLEM INSTEAD OF PREDICTION PROBLEM

- ▶ Sounds good initially
- ▶ Not enough comparisons in the dataset
  - ▶ Mostly only between yen and one other algorithm for the same image
  - ▶ Example: yen + otsu, yen + isodata, yen + li are there, but not li + isodata

## PROBLEMATIC DATASET

- ▶ Dataset has runs with same values but different user score
- ▶ Before we can do anything else, we need to fix it!

## COMING WEEK

- ▶ Create new dataframe from json file
- ▶ Multiclass Logistic Regression model
- ▶ Combine models with number of classes and imbalanced data solutions

THANK YOU!

---

**ANY QUESTIONS OR FEEDBACK?**