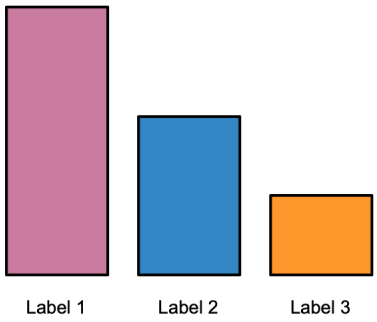
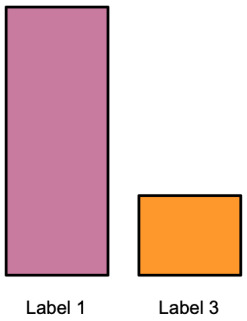
Alberto Andrés Valdés González.
Degree: Mathematical Engineer.
Work position: ML-Engineer.
Mail: anvaldes@uc.cl/alberto.valdes.gonzalez.96@gmail.com
Location: Santiago, Chile.

Oversampling

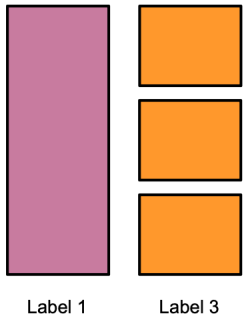
Step 0: We have imbalanced datasets.



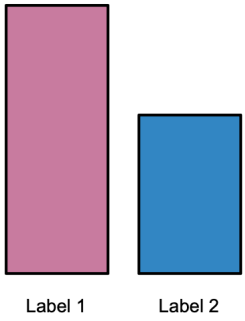
Step 1: We take the bigger dataset and we going to iterate with the other datasets.



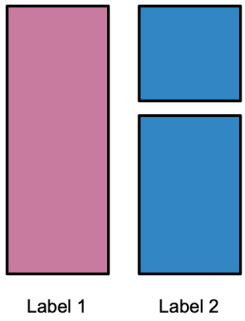
Step 2: Repeat and sample the smaller dataset till get the same size of the bigger dataset.



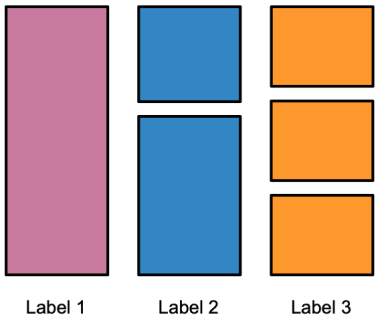
Step 3: We iterate with the other dataset.



Step 4: Repeat and sample the smaller dataset till get the same size of the bigger dataset.

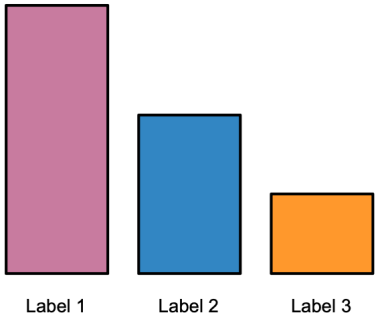


Step 5: We finally have the final datasets.

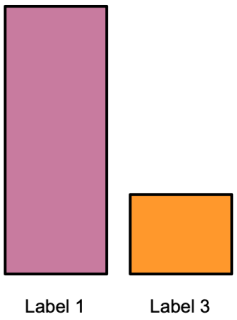


Undersampling

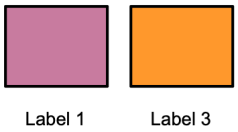
Step 0: We have imbalanced datasets.



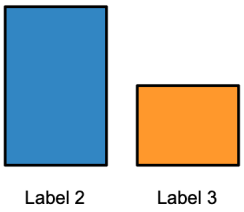
Step 1: We take the smaller dataset and we going to iterate with the other datasets.



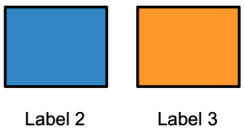
Step 2: Sample the bigger dataset till get the same size of the smaller dataset.



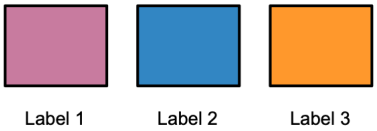
Step 3: We iterate with the other dataset.



Step 4: Sample the bigger dataset till get the same size of the smaller dataset.



Step 5: We finally have the final datasets.



SMOTE

We can use this method only for structured data and we only have to apply the same strategy of **Oversampling**.
