

Algoritmo: α Distance Borderline-ADASYN-SMOTE

Algoritmo: AR-ADASYN (Angle-Radius Adaptive Synthetic Sampling)

Referencia: Park & Kim, 2024.

1. Cálculo de pesos adaptativos por muestra minoritaria:

Para cada muestra minoritaria x_i , calcular la proporción de vecinos mayoritarios entre sus k vecinos más cercanos:

$$w(i) = \frac{\# \text{ vecinos mayoritarios}}{k}$$

Normalizar los pesos:

$$\hat{w}(i) = \frac{w(i)}{\sum_j w(j)}$$

Calcular la cantidad de ejemplos sintéticos a generar para cada muestra:

$$n_{syn}(i) = \hat{w}(i) \cdot n_{syn}$$

donde n_{syn} es la cantidad total deseada de ejemplos sintéticos.

2. Definición del área segura (Safe Region):

Seleccionar dos vecinos minoritarios cercanos x_{nn1} y x_{nn2} de x_i . Calcular:

- El ángulo entre ellos:

$$\theta' = \arccos \left(\frac{(x_{nn1} - x_i) \odot (x_{nn2} - x_i)}{\|x_{nn1} - x_i\| \cdot \|x_{nn2} - x_i\|} \right)$$

- El ángulo mínimo para evitar ambigüedad:

$$\theta = \min(\theta', \pi - \theta')$$

- El radio del área segura:

$$r = \max(\|x_i - x_{nn1}\|, \|x_i - x_{nn2}\|)$$

3. Generación de datos sintéticos dentro del área segura:

Para cada $j = 1, \dots, n_{syn}(i)$:

- Elegir un ángulo aleatorio $\alpha \in [0, \theta]$
- Construir matriz de rotación:

$$R(\alpha) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

- Calcular vector base $v = x_{nn1} - x_i$
- Elegir un radio aleatorio $\beta \in [0, r]$
- Generar el nuevo punto sintético:

$$x_{\text{syn}}^{(j)} = x_i + \beta \cdot R(\alpha) \cdot \frac{v}{\|v\|}$$