|  |
| --- |
| **Dataset HEART (densidad=25, riesgo=50, pureza=entropia)** |

**1. Identificación general**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **dataset** | **clase** | **train\_original** | **objetivo\_balance** | **estado** | **motivo\_sin\_sinteticas** |
| heart | 0 | 128 | 102 | no se sobremuestrea | sin\_faltante(actual>=objetivo) |
| heart | 1 | 43 | 102 | sobremuestreada | ok |
| heart | 2 | 28 | 102 | sobremuestreada | ok |
| heart | 3 | 28 | 102 | sobremuestreada | ok |
| heart | 4 | 10 | 102 | sobremuestreada | ok |

**2. Resultados de tamaño y ratios**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **clase** | **muestras\_sinteticas\_generadas** | **total\_original** | **total\_resampled** | **ratio\_original** | **ratio\_resampled** |
| 0 | 0 | 237 | 237 | 0.540084 | 0.540084 |
| 1 | 59 | 237 | 296 | 0.181435 | 0.344595 |
| 2 | 74 | 237 | 370 | 0.118143 | 0.275676 |
| 3 | 74 | 237 | 444 | 0.118143 | 0.229730 |
| 4 | 92 | 237 | 536 | 0.042194 | 0.190299 |

**3. Métricas internas de riesgo/densidad**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **clase** | **n\_candidatas** | **n\_filtradas** | **riesgo\_medio** | **riesgo\_std** | **densidad\_media** | **vecinos\_validos\_promedio** |
| 1 | 43 | 43 | 0.7628 | 0.2124 | 0.0651 | 3.0 |
| 2 | 28 | 28 | 0.7500 | 0.1742 | 0.0429 | 3.0 |
| 3 | 28 | 28 | 0.8214 | 0.1800 | 0.0143 | 3.0 |
| 4 | 10 | 7 | 1.0000 | 0.0000 | 0.0000 | 3.0 |

**4. Parámetros de configuración**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **clase** | **umbral\_densidad** | **umbral\_entropia** | **percentil\_densidad** | **percentil\_riesgo** | **criterio\_pureza** | **tecnica\_sobremuestreo** | **factor\_equilibrio** | **random\_state** | **modo\_espacial** | **timestamp** |
| 0 | — | — | 25.0 | 50.0 | entropia | PCSMOTE | 0.8 | 42 | 2d | 2025-08-20T00:20:21.672483 |
| 1 | 0.0 | 0.970951 | 25.0 | 50.0 | entropia | PCSMOTE | 0.8 | 42 | 2d | 2025-08-20T00:20:21.948341 |
| 2 | 0.0 | 0.970951 | 25.0 | 50.0 | entropia | PCSMOTE | 0.8 | 42 | 2d | 2025-08-20T00:20:22.062857 |
| 3 | 0.0 | 0.970951 | 25.0 | 50.0 | entropia | PCSMOTE | 0.8 | 42 | 2d | 2025-08-20T00:20:22.158406 |
| 4 | 0.0 | 0.541446 | 25.0 | 50.0 | entropia |  |  |  |  |  |

**Tabla 1 – Datos generales de configuración**

Contiene la información básica de cada dataset, el tipo de muestreo, la técnica aplicada y los parámetros principales de PCSMOTE.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **dataset** | **tipo** | **tecnica** | **densidad** | **riesgo** | **pureza** | **n\_train** |
| heart | aumentado | pcsmote | 25.0 | 50.0 | entropia | 536 |

**Tabla 2 – Detalles del modelo y configuración óptima**

Muestra la cantidad de test, número de features, folds de CV, iteraciones, modelo usado y la mejor configuración de hiperparámetros encontrada.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **n\_test** | **n\_features** | **es\_grande** | **cv\_splits** | **n\_iter** | **modelo** | **mejor\_configuracion** | **cv\_f1\_macro** |
|  | 60 | 13 | False | 5 | 10 | LogisticRegression | {'classifier\_\_C': np.float64(0.4042872735027331), 'classifier\_\_penalty': 'l1', 'classifier\_\_solver': 'liblinear'} | 0.512878 |

**Tabla 3 – Resultados de métricas de validación y test**

Presenta los resultados de validación cruzada y test en todas las métricas calculadas.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **cv\_balanced\_accuracy** | **cv\_mcc** | **cv\_cohen\_kappa** | **test\_f1\_macro** | **test\_balanced\_accuracy** | **test\_mcc** | **test\_cohen\_kappa** | **search\_time\_sec** | **n\_jobs\_search** |
| 0.541257 | 0.448448 | 0.439585 | 0.400758 | 0.416748 | 0.408567 | 0.405685 | 0.168 | 4 |