

Elemento	Regla / transformación	Motivación (calidad + forense)	Resultado esperado
Coherencia semántica	Si attack_cat = 'Normal' $\Rightarrow$ label = 0	Evita inconsistencias que invalidan métricas y complican defensa pericial	Tráfico benigno siempre etiquetado como 0
Coherencia semántica	Si attack_cat $\neq$ 'Normal' $\Rightarrow$ label = 1	Garantiza que 'ataque' sea consistente en todos los datasets	Tráfico malicioso siempre etiquetado como 1
Normalización de attack_cat	Trim, limpieza de nulos y categorías vacías; sustitución de 'nan/None/' por 'Normal' cuando proceda	Mantener taxonomía estable y auditável	attack_cat sin nulos y con valores controlados
Normalización de label	Conversión a entero, coerción de errores a 0, recorte a {0,1}	Evita valores fuera de rango y 'ruido' de etiquetado	label estrictamente binario
Normalización de columnas	Renombrado a minúsculas y sin espacios/caracteres especiales	Reproducibilidad (mismo pipeline en todos los datasets)	Esquema homogéneo para scripts y trazabilidad
Control de integridad	Revisión de 'Normal con label=1' y 'Ataque con label=0'	Detectar fugas o corrupción del dataset	Dataset final consistente y defendible