



Momento de Retroalimentación: Módulo 2 - Uso de framework o biblioteca de aprendizaje máquina para la implementación de una solución

Raúl Youthan Irigoyen Osorio A01750476

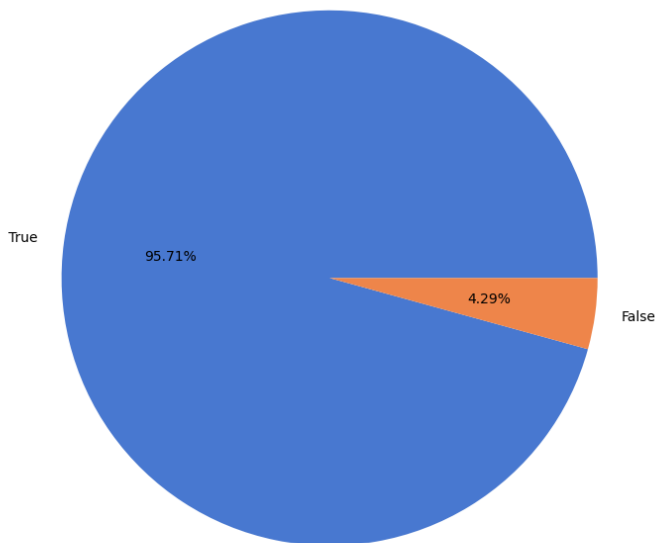
13 de septiembre del 2022

Inteligencia artificial avanzada para la ciencia de datos

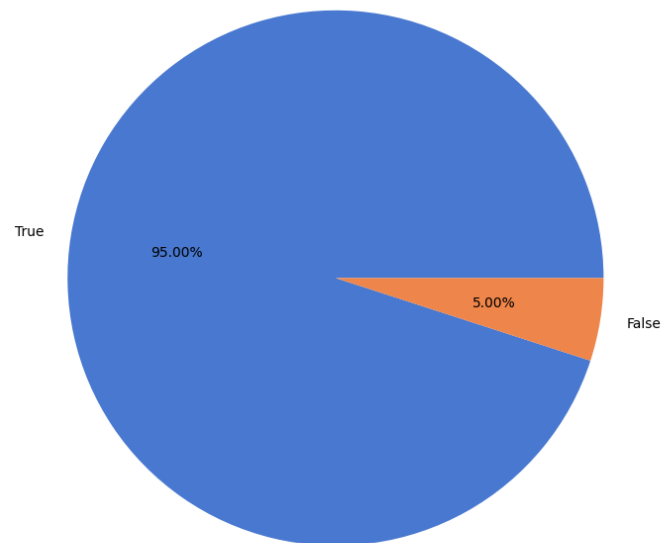
Las pruebas mostradas a continuación pretenden demostrar el cambio en la precisión del modelo al utilizar diferentes métodos de entrenamiento para la regresión logística, así como diferentes estrategias de penalización:

Solucionador por Método de Newton:

newton-cg solver with l2 penalty accuracy by iterations

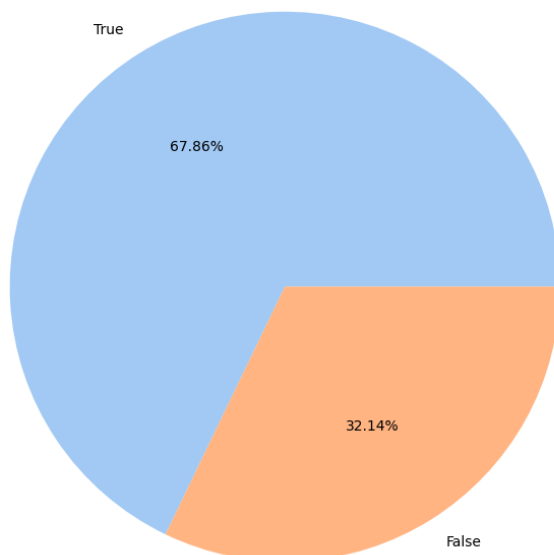


newton-cg solver with none penalty accuracy by iterations

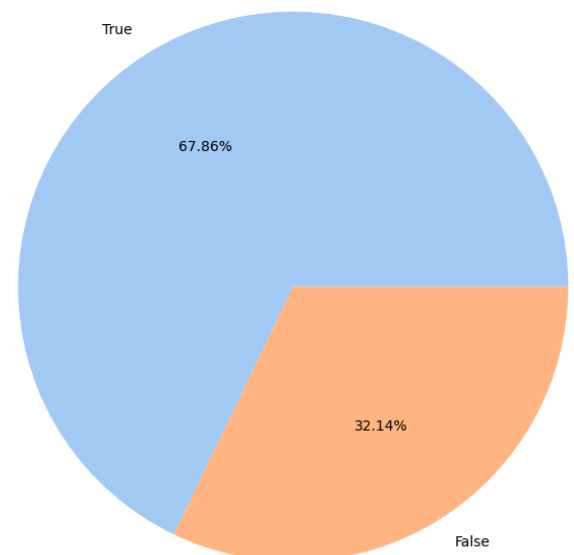


Solucionador de Algoritmo de Memoria Limitada Broyden-Fletcher-Goldfarb-Shanno:

lbfgs solver with l2 penalty accuracy by iterations

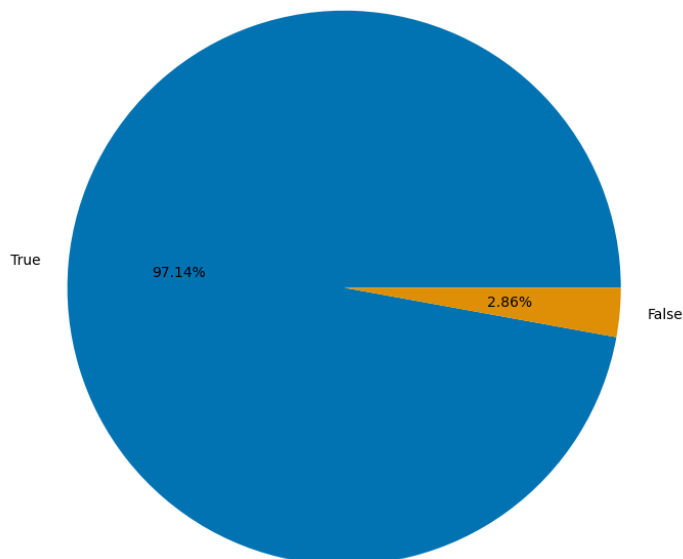


lbfgs solver with none penalty accuracy by iterations

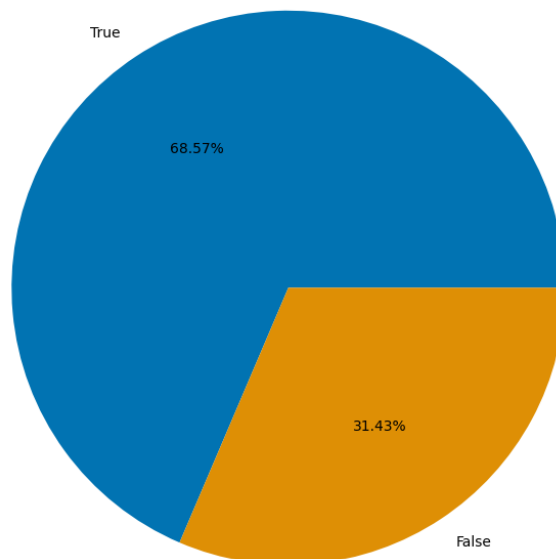


Solucionador de Clasificación Lineal de Gran Escala:

liblinear solver with l1 penalty accuracy by iterations

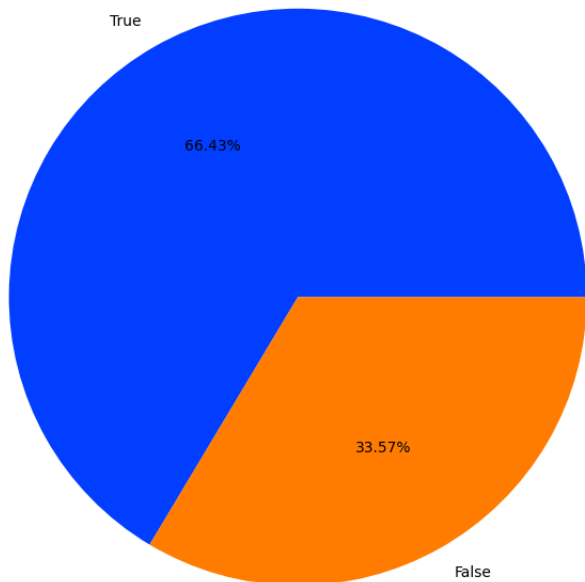


liblinear solver with l2 penalty accuracy by iterations

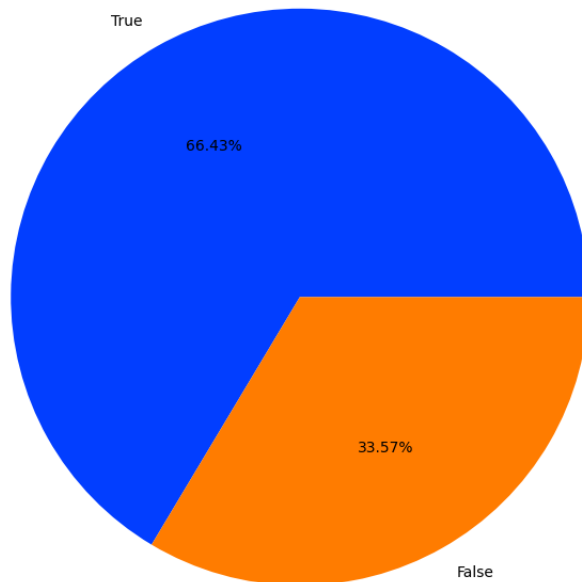


Solucionador de Gradiente Estocástica por Promedio:

sag solver with l2 penalty accuracy by iterations

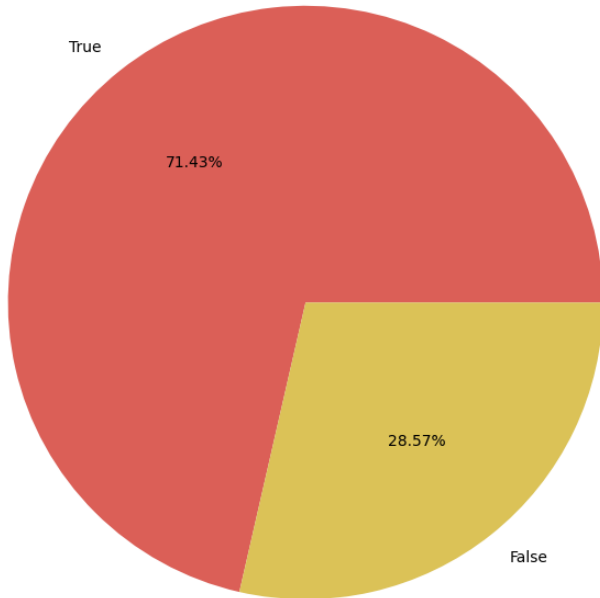


sag solver with none penalty accuracy by iterations

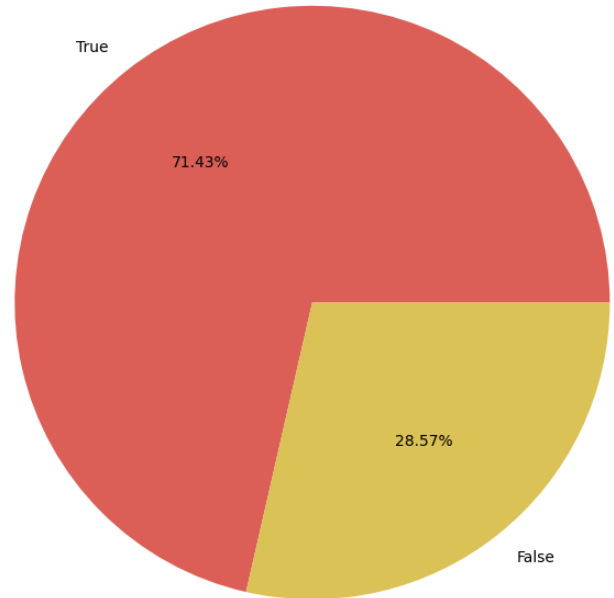


Solucionador de Gradiente Estocástica (Variante):

saga solver with l2 penalty accuracy by iterations



saga solver with none penalty accuracy by iterations



Dada la evidencia anterior es visible que para casos donde hay un número relativamente alto de atributos a considerar para el resultado de la clasificación, resulta mejor utilizar el solucionador por clasificación lineal de gran escala. Lo anterior es cierto dado que este método utiliza Support Vector Machines para realizar los cálculos, lo cual, junto con la penalización L1 (*que utiliza el valor absoluto de la magnitud del error para penalizar*), permite obtener un resultado casi perfecto para este tipo de datasets.