

Reporte: Actividad 7.1 y 7.2

Regresión No Lineal

Para analizar la relación entre el total de gastos y el (IVA) de Calor y Control, se utilizaron los datos de las columnas TOTAL MX e IVA del archivo Gastos y Costos 20-23.csv. Estos datos corresponden a los años 2020, 2021, 2022 y 2023. Se asumió que el IVA era la variable independiente y el TOTAL MX la variable dependiente, y se ajustaron ocho modelos no lineales distintos. Los resultados mostraron que los modelos cuadráticos tenían el mejor ajuste, con un coeficiente de determinación (R^2) superior al 0.7 para los cuatro años. El modelo más preciso fue el cuadrático del año 2021, con un R^2 de 0.90. Por el contrario, los modelos cúbicos tuvieron el peor ajuste en todos los años.

Regresión Logística

Para clasificar el tipo de gasto de Calor y Control en dos categorías: COMPRAS y OTROS_GASTOS, se modificó la columna TIPO GASTO del archivo Gastos y Costos 20-23.csv. Se usaron las mismas variables independiente y dependiente que en la regresión no lineal, y se entrenaron cuatro modelos de regresión logística: uno por cada año. Se evaluaron las métricas de precisión, exactitud, sensibilidad y F1 para cada modelo. El modelo del año 2021 tuvo la mayor precisión (0.517) y sensibilidad (0.667), lo que indica que fue el más capaz de identificar correctamente las COMPRAS. El modelo del año 2020 tuvo la mayor exactitud (0.533), lo que significa que tuvo menos errores de clasificación en general. El modelo del año 2021 también tuvo el mayor valor F1 (0.583), lo que refleja un buen balance entre precisión y sensibilidad.