

Reporte Comparativo

Diego Garza Solís, A01705984

Se utilizaron cinco modelos de regresión no lineal para determinar las relaciones entre cinco pares de variables. Se escogieron los métodos de regresión polinómica, *random forest*, SVM, KNN y *decision trees*. Se encontró, para cada caso, lo siguiente:

1. Riesgo vs. Score de Buró

Regresión Polinómica: $R^2 = 0.85$

Random forest: $R^2 = 0.90$

SVM: $R^2 = 0.86$

KNN: $R^2 = 0.88$

Decision trees: $R^2 = 0.90$

En este caso, los modelos de *decision trees* y *random forest* mostraron el mejor ajuste a los datos con el coeficiente de determinación más alto.

2. Riesgo vs. Porcentaje de Enganche

Regresión Polinómica: $R^2 = 0.45$

Random forest: $R^2 = 0.66$

SVM: $R^2 = 0.50$

KNN: $R^2 = 0.61$

Decision trees: $R^2 = 0.67$

En este caso, el modelo de *decision trees* mostró el mejor ajuste con el coeficiente de determinación más alto.

3. Riesgo vs. Límite de Crédito

Regresión Polinómica: $R^2 = 0.61$

Random forest: $R^2 = 0.76$

SVM: $R^2 = 0.46$

KNN: $R^2 = 0.71$

Decision trees: $R^2 = 0.77$

Se repite, *decision trees* con el mejor resultado.

4. Score de Buró vs. Porcentaje de Enganche

Regresión Polinómica: $R^2 = 0.37$

Random forest: $R^2 = 0.47$

SVM: $R^2 = 0.39$

KNN: $R^2 = 0.34$

Decision trees: $R^2 = 0.47$

Aquí, *decision trees* y *random forest* con el mejor resultado.

5. Score de Buró vs. Límite de Crédito

Regresión Polinómica: $R^2 = 0.54$

Random forest: $R^2 = 0.61$

SVM: $R^2 = 0.45$

KNN: $R^2 = 0.50$

Decision trees: $R^2 = 0.61$

Aquí, *decision trees* y *random forest* también con el mejor resultado.

Comentario final

En este análisis, se ha observado que, en la mayoría de los casos, el modelo de *decision trees* ha proporcionado el mejor ajuste a los datos, aunado al modelo de *random forest*, dado el coeficiente de determinación que han arrojado como resultado. En este tenor, se puede concluir que este modelo puede ser una elección óptima para predecir relaciones no lineales entre variables en el marco de esta base de datos en específico, siempre recordando que cada caso es único en su totalidad y, por tanto, diferentes conjuntos de datos podrán tener características que los vuelva más afines a otros modelos.