



Tecnológico de Monterrey

Actividad 3.2: Regresión No Lineal

Ivanna Maldonado Cervantes

Paula Simonetta Madrid Pérez

Ania Diaz Gonzalez

Miranda Eugenia Colorado Arróniz

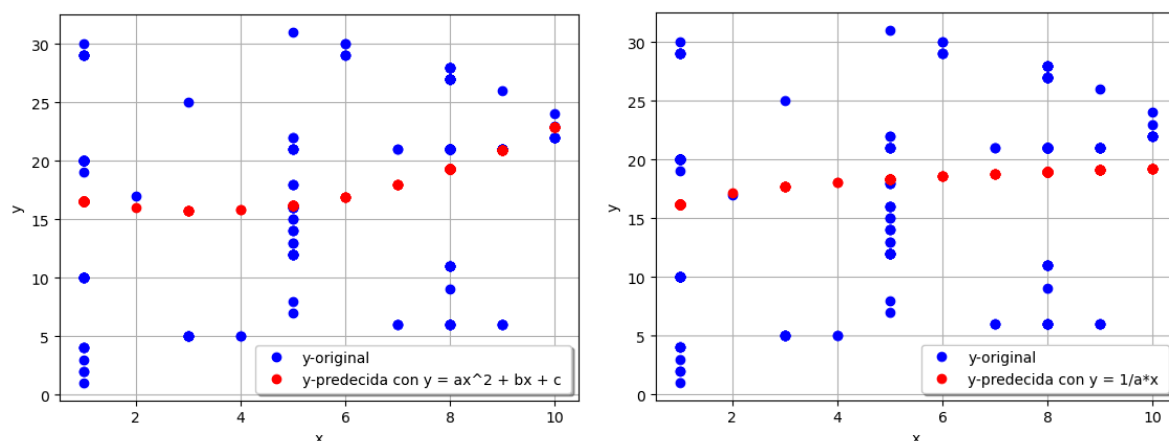
10 de octubre de 2025

Analítica de datos y herramientas de inteligencia artificial I (Gpo 101)

Dr. Alfredo García Suárez

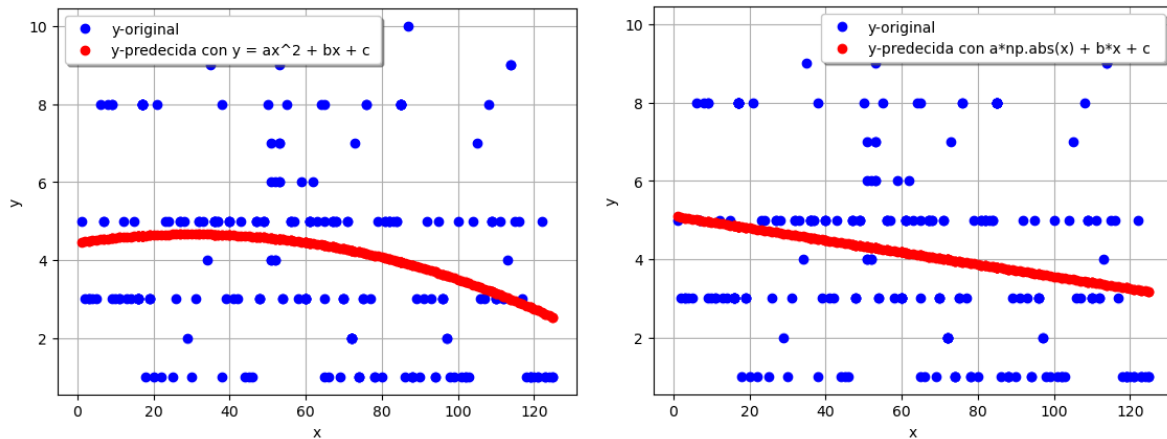
La presente actividad se centra en la aplicación de modelos de regresión no lineal para explorar las relaciones entre diferentes atributos del conjunto de datos de la empresa Forvia. En este análisis, se utilizaron enfoques basados en diferentes funciones no lineales, como la función cuadrática, logarítmica, de valor absoluto, exponencial y de cociente entre polinomios. Todo con el objetivo de evaluar la capacidad predictiva de cada modelo y comparar sus resultados con los obtenidos previamente mediante regresión lineal. A continuación, se presenta un análisis detallado de cada atributo y los hallazgos obtenidos.

En el caso del atributo "Project Type", se aplicaron tanto la regresión cuadrática como la logarítmica para modelar su relación con "Project Organization". La regresión cuadrática mostró un mejor desempeño en comparación con la regresión lineal, con un coeficiente de determinación (R^2) superior al 0.22 obtenido en el modelo lineal. Esto sugiere que la relación entre estas variables no es estrictamente lineal y que un modelo cuadrático puede capturar mejor las variaciones en los datos. Por otro lado, la regresión logarítmica presentó un desempeño inferior, con un R^2 de 0.17, lo que indica que este modelo no es tan adecuado para describir la relación entre estas variables. En general, se concluye que el modelo cuadrático es más apropiado para este caso.



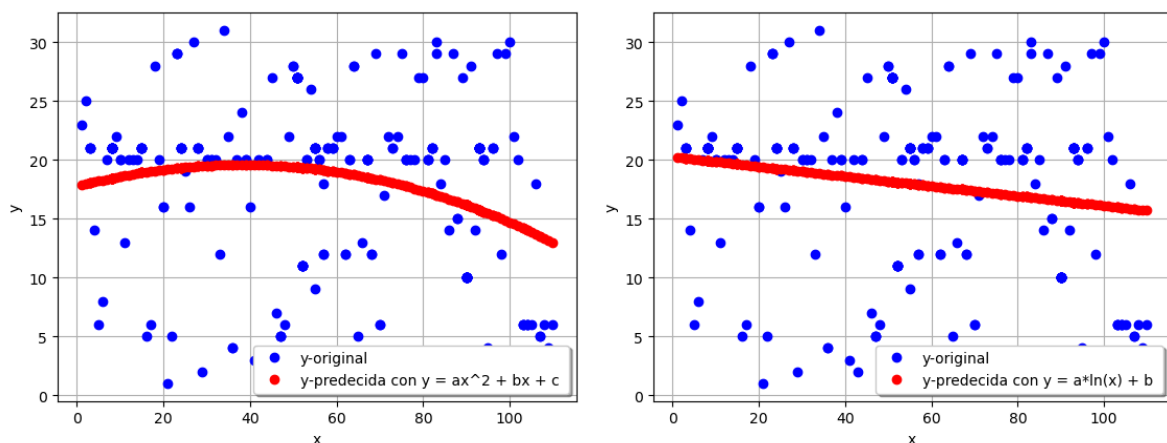
Gráfica 1 y 2. Diagramas de dispersión de los modelos "Cuadrático" y "Logarítmico" del atributo "Project Type" con relación con "Project Organization"

Para el atributo "Geographical Scope", se exploró su relación con "BG" utilizando los modelos cuadrático y de valor absoluto. En este caso, la regresión cuadrática logró un R^2 de 0.25, superando el 0.23 obtenido en la regresión lineal. Esto refuerza la idea de que la relación entre estas variables es mejor representada por un modelo no lineal. Sin embargo, la regresión de valor absoluto no logró superar el desempeño del modelo lineal, lo que sugiere que la naturaleza de la relación entre estas variables no se ajusta bien a una función logarítmica. En este contexto, el modelo cuadrático se posiciona como la mejor opción para describir la relación entre "Geographical Scope" y "BG".



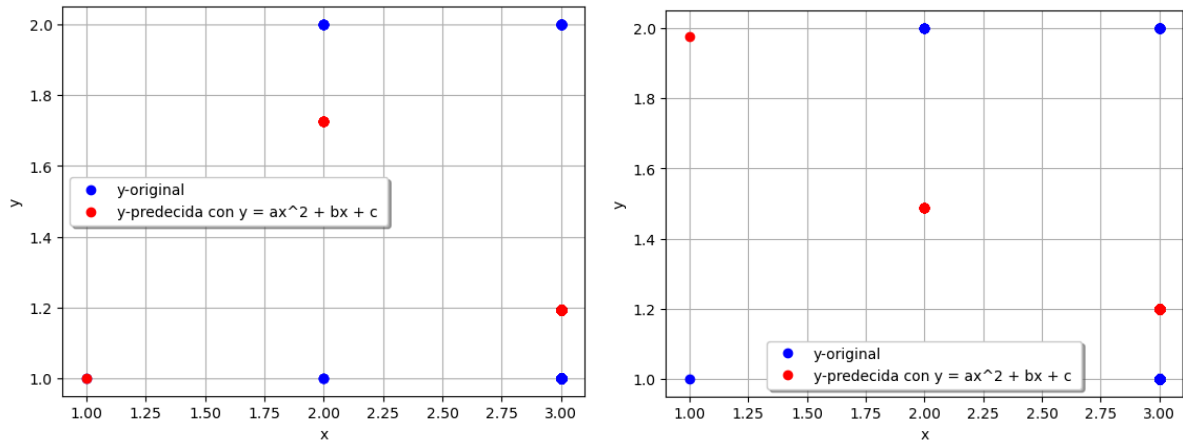
Gráfica 3 y 4. Diagramas de dispersión de los modelos "Cuadrático" y "De valor absoluto" del atributo "Geographical Scope" con relación con "BG"

En el análisis del atributo "Project Manager", se evaluó su relación con "Project Organization". La regresión cuadrática nuevamente mostró un mejor desempeño en comparación con la regresión lineal, con un R^2 superior al 0.18 obtenido en el modelo lineal. Esto indica que el modelo cuadrático puede capturar mejor las variaciones en los datos y proporcionar una representación más precisa de la relación entre estas variables. Por otro lado, la regresión exponencial no logró superar el desempeño del modelo lineal, lo que sugiere que este enfoque no es adecuado para este caso. En general, se concluye que el modelo cuadrático es más efectivo para modelar la relación entre "Project Manager" y "Project Organization".



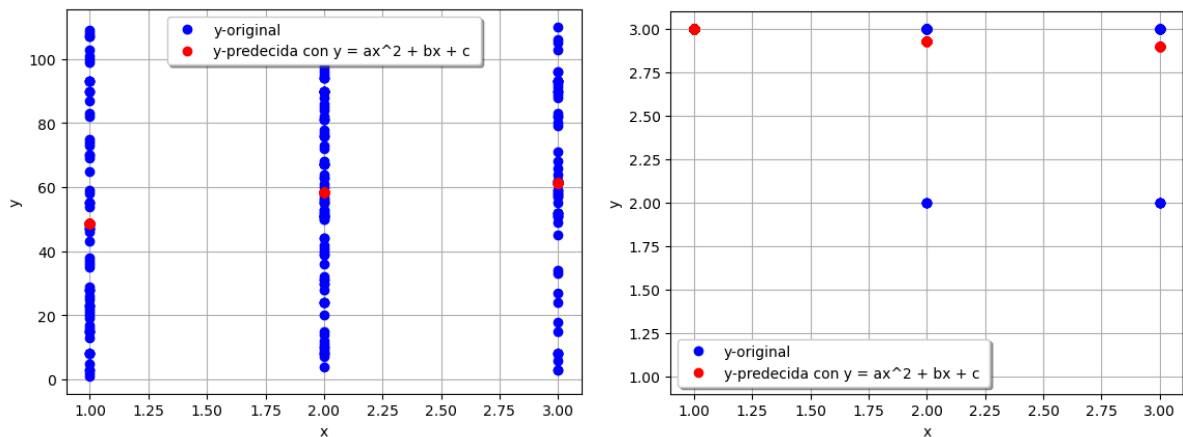
Gráfica 5 y 6. Diagramas de dispersión de los modelos "Cuadrático" y "Exponencial" del atributo "Project Manager" con relación con "Project Organization"

En el caso del atributo "State", se exploró su relación con "On-hold". La regresión cuadrática logró un R^2 de 0.28, superando significativamente el 0.22 obtenido en la regresión lineal. Esto indica que la relación entre estas variables es mejor representada por un modelo no lineal. Sin embargo, la regresión logarítmica no logró superar el desempeño del modelo lineal, lo que sugiere que este enfoque no es adecuado para este caso. En este contexto, el modelo cuadrático se posiciona como la mejor opción para describir la relación entre "State" y "On-hold".



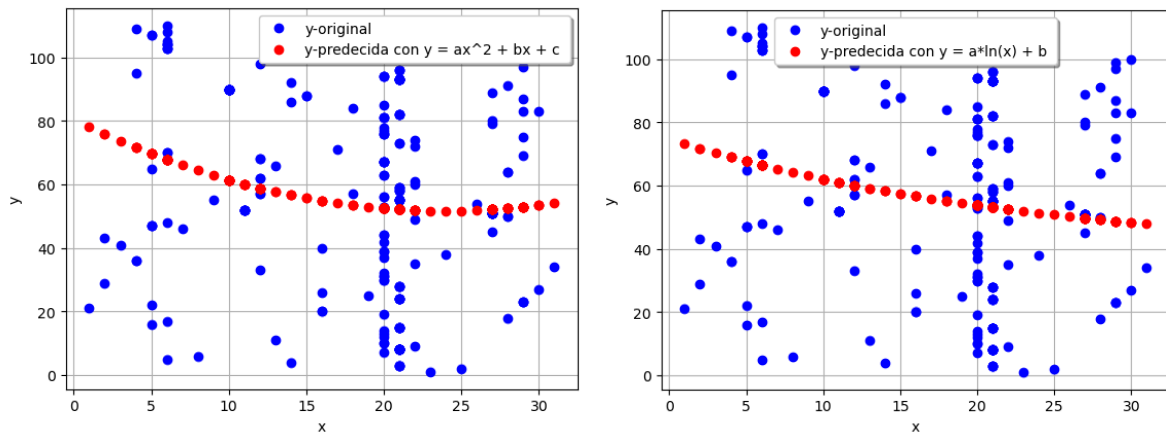
Gráfica 7 y 8. Diagramas de dispersión de los modelos "Cuadrático" y "Logarítmico" del atributo "State" con relación con "On-hold"

Para el atributo "Project Size", se evaluó su relación con "Project Manager" y "State". En ambos casos, la regresión cuadrática mostró un mejor desempeño en comparación con la regresión lineal, con R^2 superiores a los obtenidos en los modelos lineales. Esto refuerza la idea de que las relaciones entre estas variables son mejor representadas por modelos no lineales. En general, se concluye que el modelo cuadrático es más efectivo para modelar las relaciones entre "Project Size" y las variables dependientes analizadas.



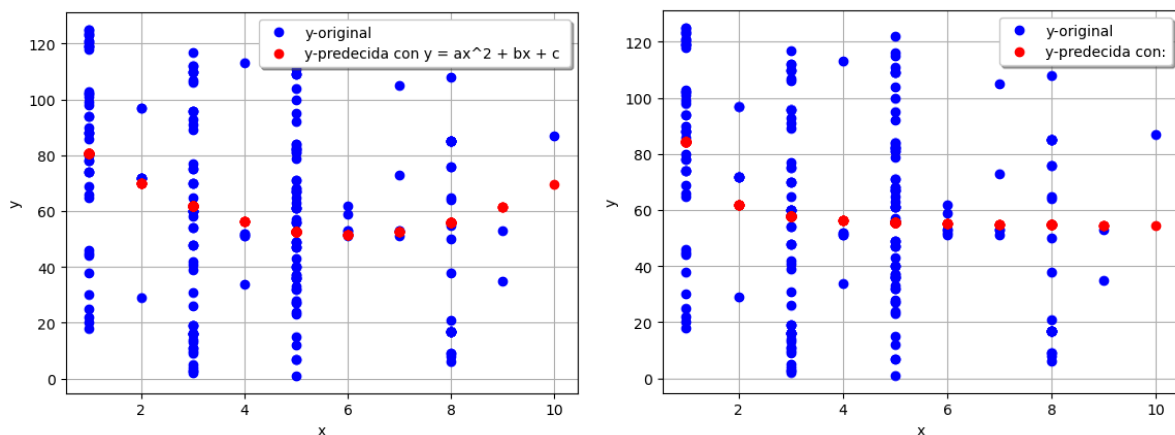
Gráfica 9 y 10. Diagramas de dispersión del modelo "Cuadrático" respecto al atributo "Project Size" con relación con "Project Manager" y "State"

En el análisis del atributo "Project Organization", se exploró su relación con "Project Manager". La regresión cuadrática logró un R^2 de 0.20, lo que indica un desempeño ligeramente inferior al 0.22 obtenido en la regresión lineal. Sin embargo, este modelo sigue siendo una de las mejores opciones para describir la relación entre estas variables. Por otro lado, la regresión exponencial no logró superar el desempeño del modelo lineal, lo que sugiere que este enfoque no es adecuado para este caso.



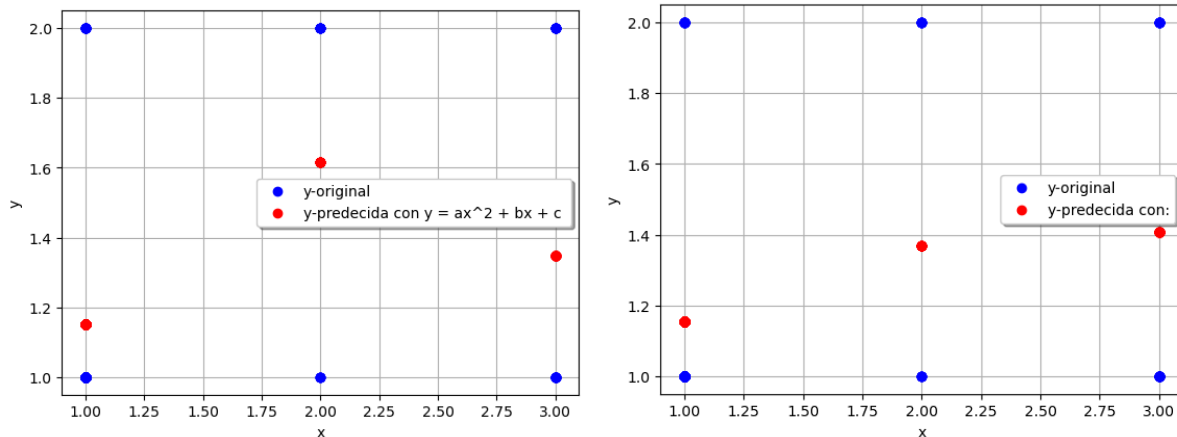
Gráfica 11 y 12. Diagramas de dispersión de los modelos "Cuadrático" y "Exponencial" del atributo "Project Organization" con relación con "Project Manager"

En el caso del atributo "BG", se evaluó su relación con "Geographical Scope". La regresión cuadrática logró un R^2 de 0.29, superando el 0.23 obtenido en la regresión lineal. Esto indica que la relación entre estas variables es mejor representada por un modelo no lineal. Además, se exploró una función cociente entre polinomios, que logró un R^2 de 0.32, posicionándose como el mejor modelo para describir esta relación. Estos resultados destacan la importancia de considerar modelos más complejos para capturar las relaciones entre estas variables.



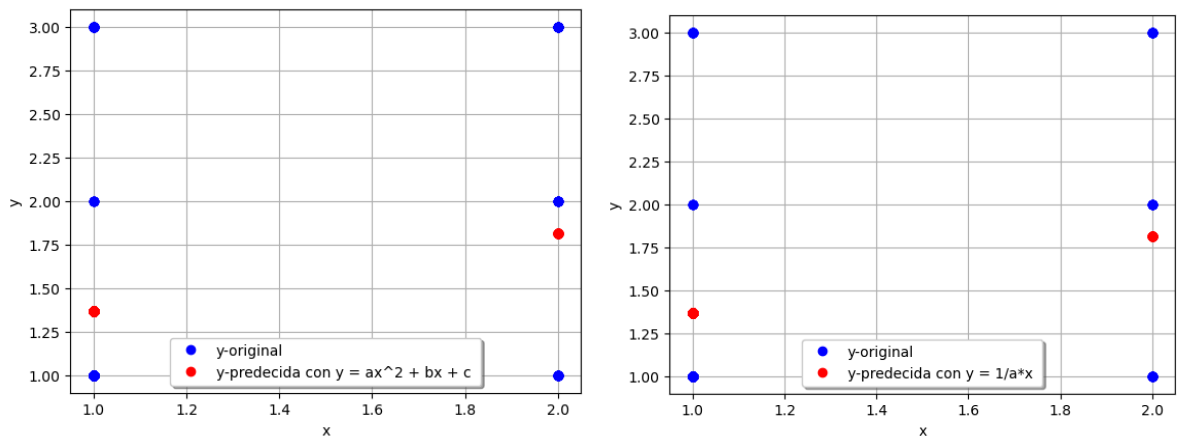
Gráfica 13 y 14. Diagramas de dispersión de los modelos "Cuadrático" y "De cociente entre polinomios" del atributo "BG" con relación con "Geographical Scope"

En el análisis del atributo "Project Health", se exploró su relación con "On-hold". La regresión cuadrática logró un R^2 de 0.30, superando el 0.23 obtenido en la regresión lineal. Esto indica que la relación entre estas variables es mejor representada por un modelo no lineal. Por otro lado, la función cociente entre polinomios logró un R^2 de 0.26, lo que sugiere que este modelo también puede ser una opción adecuada para describir esta relación.



Gráfica 15 y 16. Diagramas de dispersión de los modelos "Cuadrático" y "De cociente entre polinomios" del atributo "Project Health" con relación con "On-hold"

Finalmente, en el análisis del atributo "On-hold", se evaluó su relación con "Project Health". La regresión cuadrática mostró un menor desempeño en comparación con la regresión lineal, con un R^2 de 0.22, menor al 0.23 obtenido en el modelo lineal. Esto indica que el modelo cuadrático puede no ser el ideal para capturar las variaciones en los datos. Por otro lado, la regresión logarítmica tampoco logró superar el desempeño del modelo lineal con un R^2 de 0.22, lo que sugiere que este enfoque tampoco es el adecuado para este caso. En general, se concluye que el modelo lineal es más efectivo para modelar la relación entre "On-hold" y "Project Health".



Gráfica 17 y 18. Diagramas de dispersión de los modelos "Cuadrático" y "Logarítmico" del atributo "On-hold" con relación con "Project Health"

En conclusión, el análisis realizado demuestra que las relaciones entre los atributos del conjunto de datos son, en muchos casos, mejor representadas por modelos no lineales. La regresión cuadrática se destacó como el modelo más efectivo en la mayoría de los casos, superando consistentemente el desempeño de la regresión lineal. Sin embargo, en algunos casos específicos, como la relación entre "BG" y "Geographical Scope", otros modelos no lineales, como la función cociente entre polinomios, lograron un mejor desempeño. Estos resultados resaltan la importancia

de explorar diferentes enfoques de modelado para capturar de manera más precisa las relaciones entre las variables y mejorar la capacidad predictiva de los modelos.