



# Tecnológico de Monterrey

**Instituto Tecnológico y de Estudios Superiores de Monterrey Campus Puebla**

## **Actividad 3.2 Regresión No Lineal Wuupi**

### **Equipo 5:**

José Manuel Morales Escalante

Josué Ibáñez Santana

Sergio Giles Bracamontes

Ismael Martinez Duran

### **Materia:**

Analítica de datos y herramientas de inteligencia artificial II

### **Fecha:**

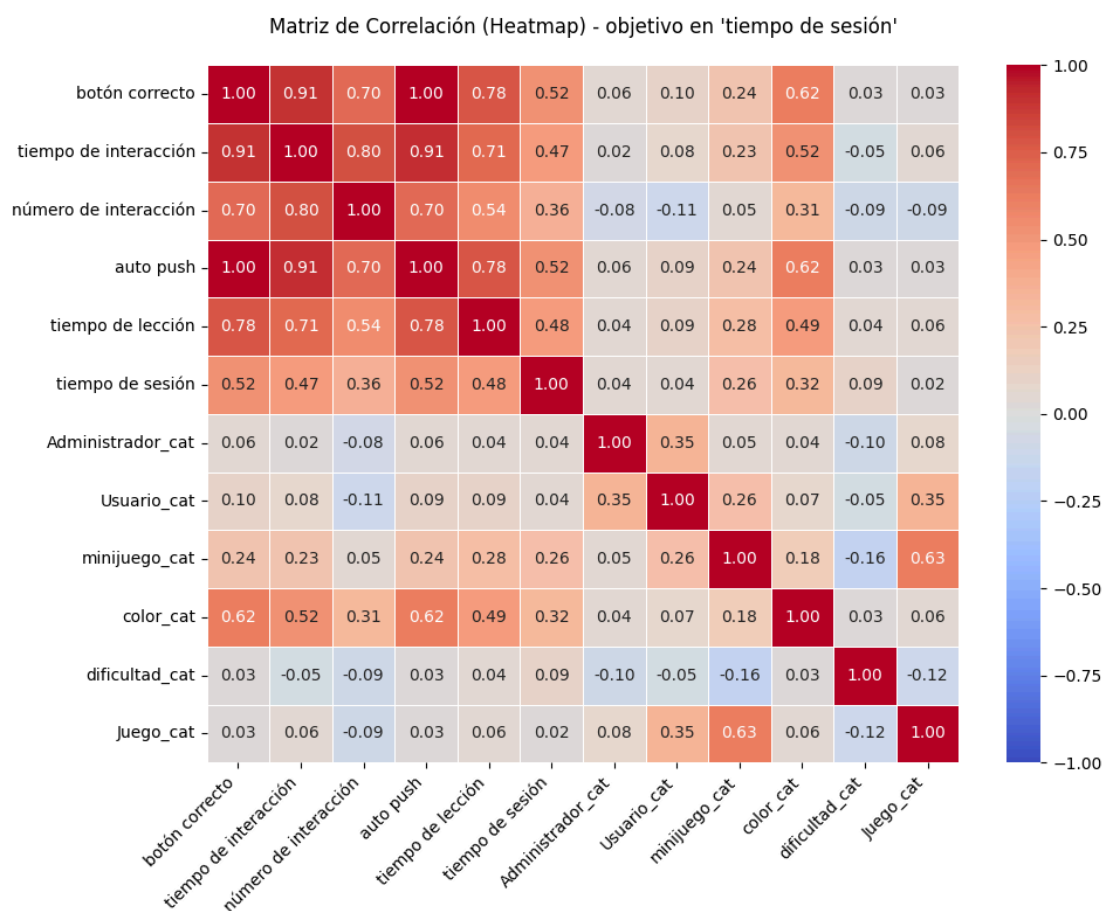
11 de abril de 2025

El estudio tuvo como objetivo principal analizar las relaciones entre variables clave en un entorno interactivo, centrándose en dos aspectos:

- Correlaciones lineales: Evaluar la relación lineal entre la variable objetivo y otras variables numéricas mediante un heatmap, esto en base a lo obtenido en el trabajo anterior.
- Modelos no lineales: Proponer alternativas no lineales (como regresión polinomial, SVM y función cuadrática) para superar las limitaciones de los modelos lineales y mejorar la precisión predictiva, adaptándose a comportamientos específicos de los usuarios.

El enfoque busca mejorar la precisión predictiva y adaptarse a comportamientos específicos, optimizando estrategias basadas en datos.

Heatmap del dataframe general:



## Correlaciones Fuertes

Estas variables tienen una relación lineal significativa con la variable objetivo u otras variables clave:

- **Botón correcto y auto push:**
  - Correlación con "tiempo de sesión": 0.52 (moderada).
- **Tiempo de interacción:**
  - Correlación alta con "botón correcto" (0.91) y "auto push" (0.91).
  - Correlación con "tiempo de sesión": 0.47 (moderada-baja).
- **Color\_cat:**
  - Correlación notable con "botón correcto" (0.62) y "auto push" (0.62).
  - Correlación con "tiempo de sesión": 0.32 (débil).

**Interpretación:** Las acciones del usuario ("botón correcto", "auto push") y el tiempo dedicado a interacciones están vinculadas al tiempo de sesión, aunque no de manera extremadamente fuerte.

## Correlaciones Moderadas

- **Tiempo de lección:**
  - Correlación con "tiempo de sesión": 0.48.
  - También relacionada con "botón correcto" (0.78) y "auto push" (0.78).
- **Minijuego\_cat:**
  - Correlación con "tiempo de sesión": 0.26.
  - Relación fuerte con "juego\_cat" (0.63), lo que sugiere que el tipo de minijuego puede influir en otras categorías.

**Interpretación:** El diseño del juego (lecciones, minijuegos) tiene un impacto moderado en el tiempo de sesión.

## Modelo No Lineal propuesto:

Se probaron 4 modelos iniciales: ( $R^2$ )

Cociente de polinomios: 0.5191422153906852

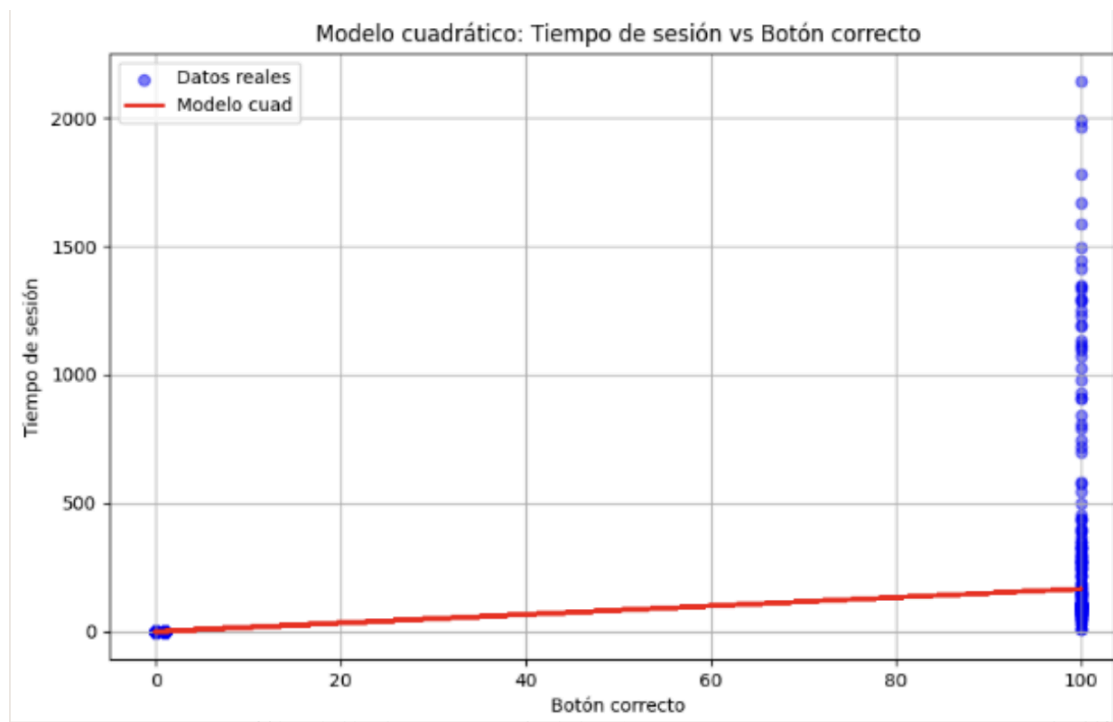
Exponencial: 0.17840017598095675

Cuadrática: 0.5191422165528985

Polinomio inverso: 0.5191422164758301

Modelo elegido: Cuadrático 

- Mayor poder predictivo: Explica el 51.9% de la varianza ( $R^2$  más alto).
- Simplicidad y eficiencia: Igual desempeño que modelos complejos, pero más interpretable.
- Captura no linealidad: Ajusta mejor la relación entre botones correctos y tiempo de sesión.



Zona inicial (Aciertos bajos):

Si el usuario acierta poco (botón correcto bajo), el tiempo de sesión puede ser alto (dificultad o indecisión) o bajo (desinterés). Ejemplo: Si la curva es U invertida, al inicio el tiempo aumenta con más aciertos (mejora en la interacción)

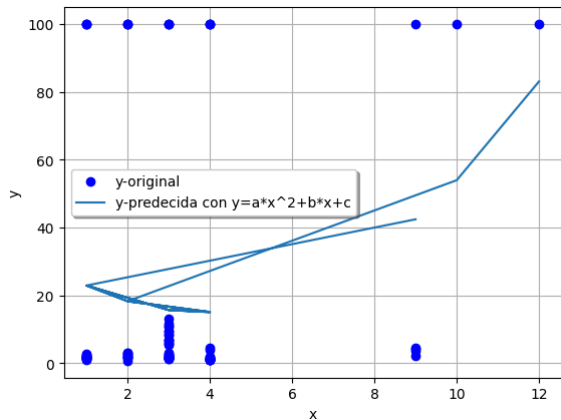
Punto crítico (Máximo/Mínimo): Indica el óptimo de eficiencia: Si hay un máximo, existe un número de aciertos donde el tiempo de sesión es más largo (ej: mayor engagement). Si hay un mínimo, cierto nivel de aciertos reduce el tiempo (ej: usuarios rápidos y precisos).

Zona final (Aciertos altos): Si el usuario acierta mucho (botón correcto alto), el tiempo puede:

Decrecer (por dominio de la tarea, eficiencia).

Creecer (por complejidad añadida o fatiga).

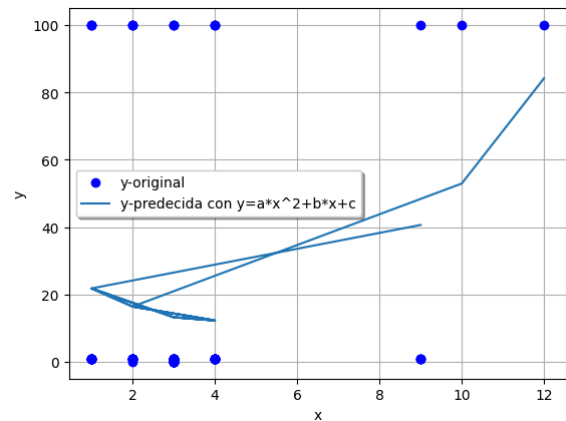
José Ian:



**tiempo de interacción**

Modelo lineal: 0.19

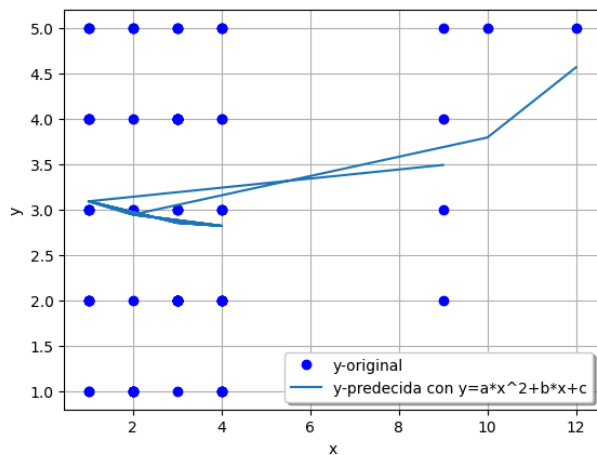
Función cuadrática con minijuego\_cat: 0.275



**botón correcto**

Modelo lineal: 0.19

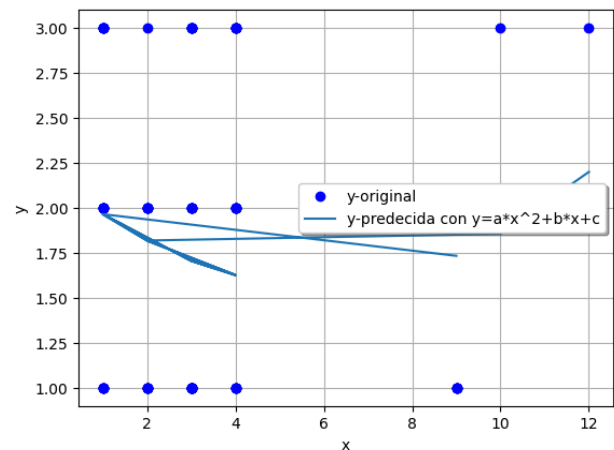
Función cuadrática con minijuego\_cat: 0.282



**color (categorizado)**

Modelo lineal: 0.12

Función cuadrática con minijuego\_cat: 0.187

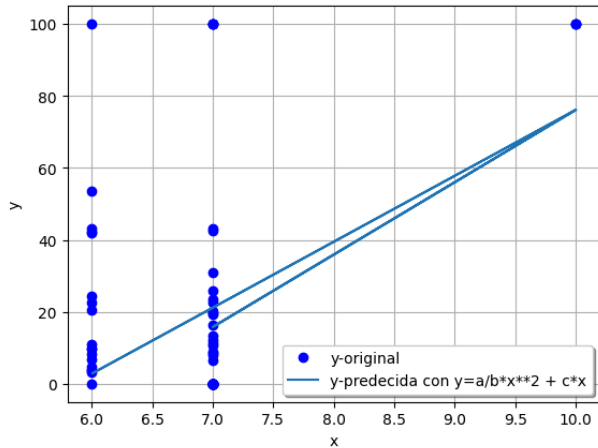


**dificultad (categorizado)**

Modelo lineal: 0.04

Función cuadrática con minijuego\_cat: 0.153

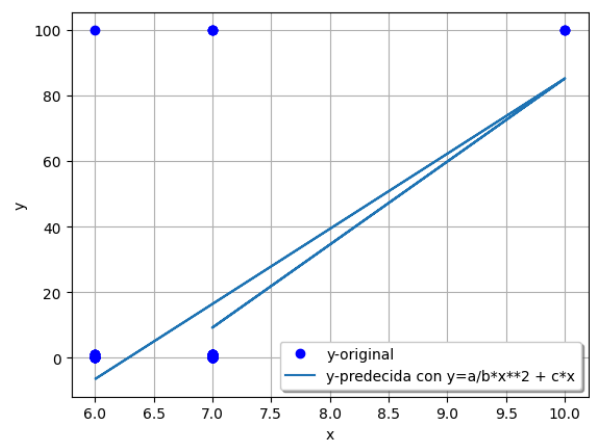
José Ignacio Tadeo:



**tiempo de interacción**

Modelo lineal: 0.34

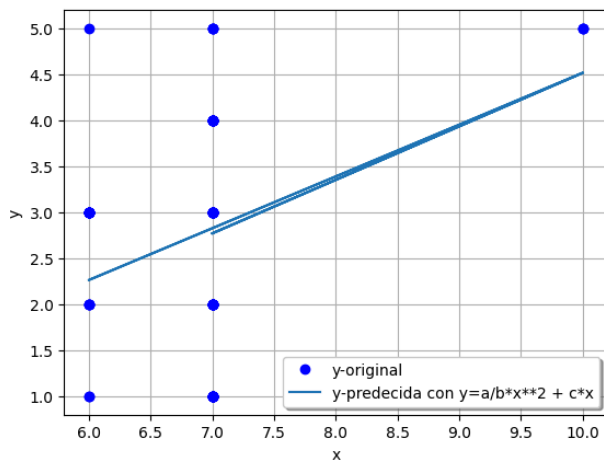
Función cuadrática con minijuego\_cat: 0.41



**botón correcto**

Modelo lineal: 0.49

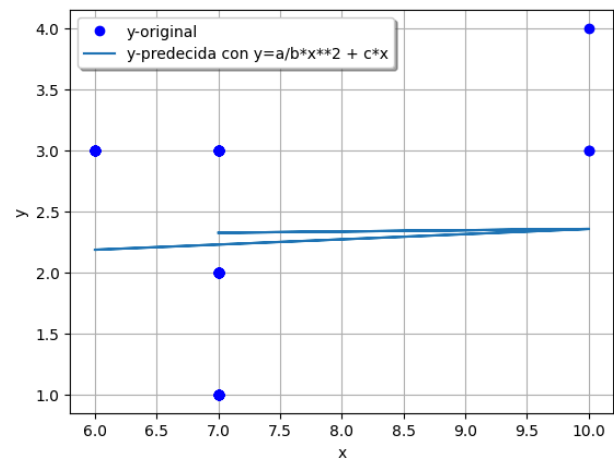
Función cuadrática con minijuego\_cat: 0.54



**color (categorizado)**

Modelo lineal: 0.28

Función cuadrática con minijuego\_cat: 0.28

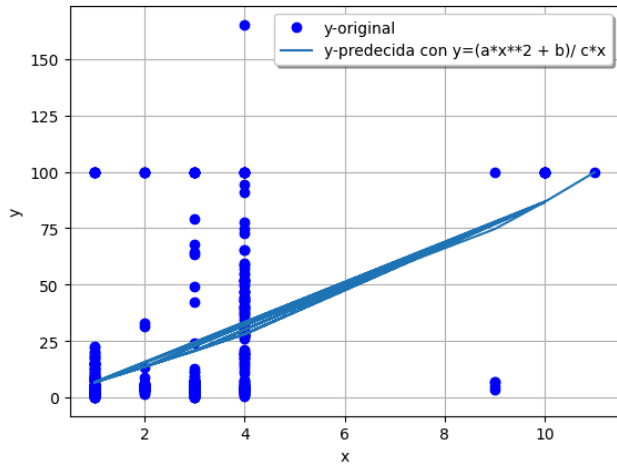


**dificultad (categorizado)**

Modelo lineal: 0.06

Función cuadrática con minijuego\_cat: 0.24

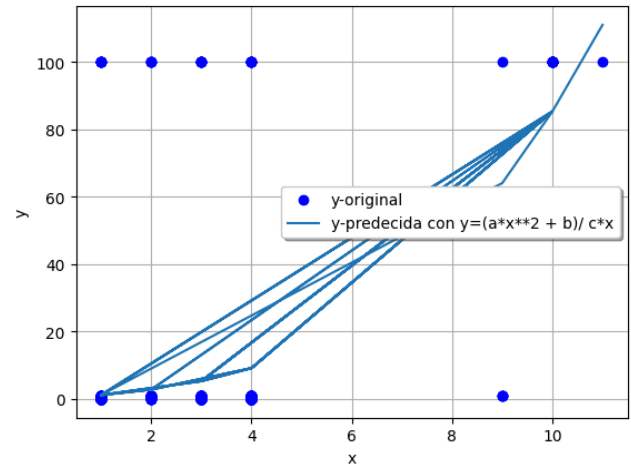
José Javier:



**tiempo de interacción**

Modelo lineal: 0.46

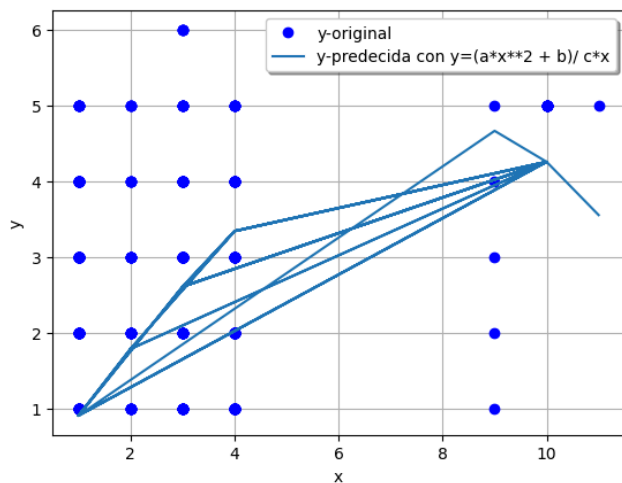
Función cuadrática con minijuego\_cat: 0.46



**botón correcto**

Modelo lineal: 0.41

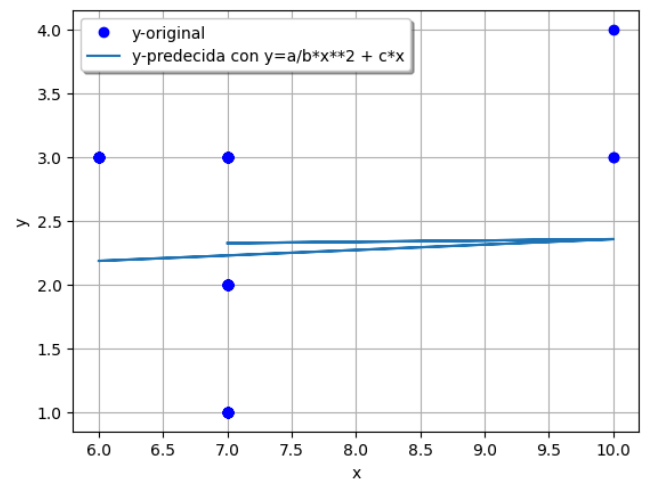
Función cuadrática con minijuego\_cat: 0.52



**color (categorizado)**

Modelo lineal: 0.18

Función cuadrática con minijuego\_cat: 0.86

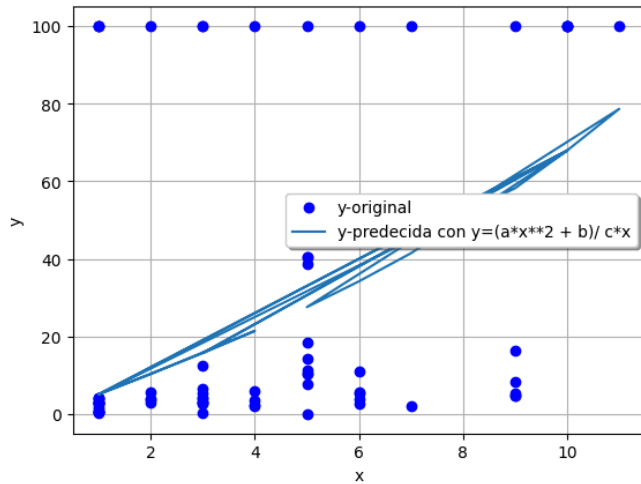


**dificultad (categorizado)**

Modelo lineal: 0.05

Función cuadrática con minijuego\_cat: 0.81

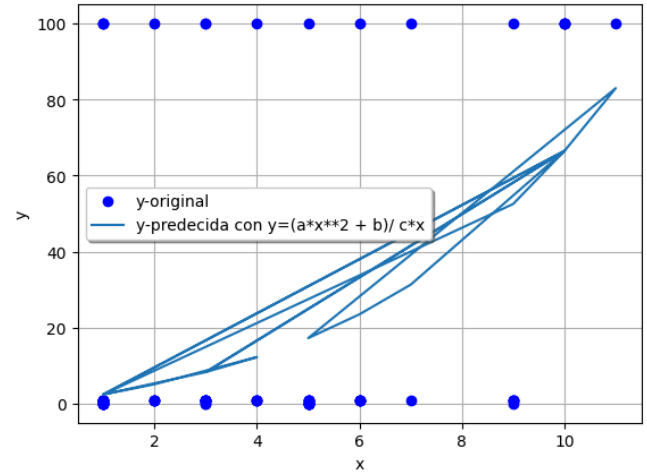
Joshua:



**tiempo de interacción**

Modelo lineal: 0.36

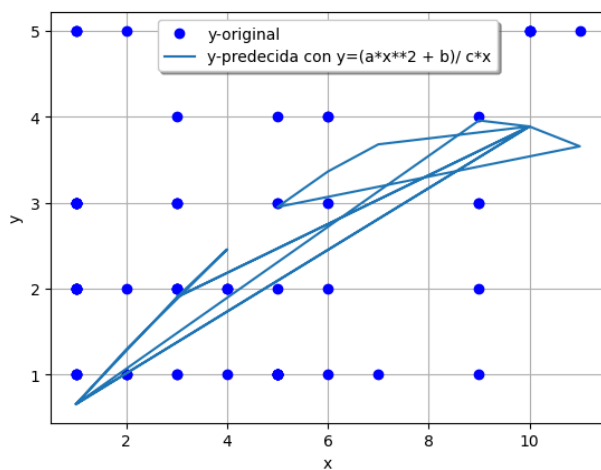
Función cuadrática con minijuego\_cat: 0.34



**botón correcto**

Modelo lineal: 0.31

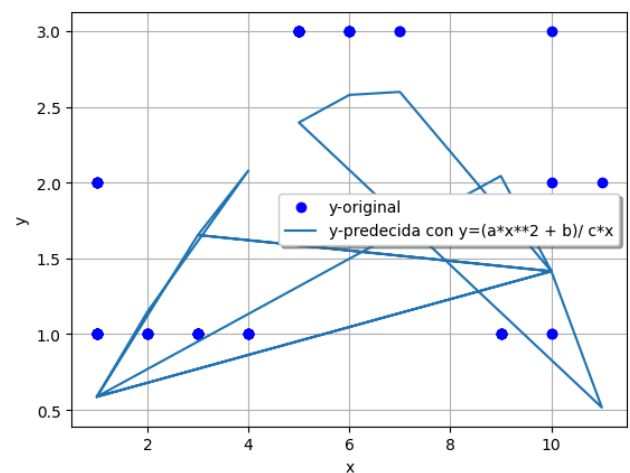
Función cuadrática con minijuego\_cat: 0.35



**color (categorizado)**

Modelo lineal: 0.29

Función cuadrática con minijuego\_cat: 0.67



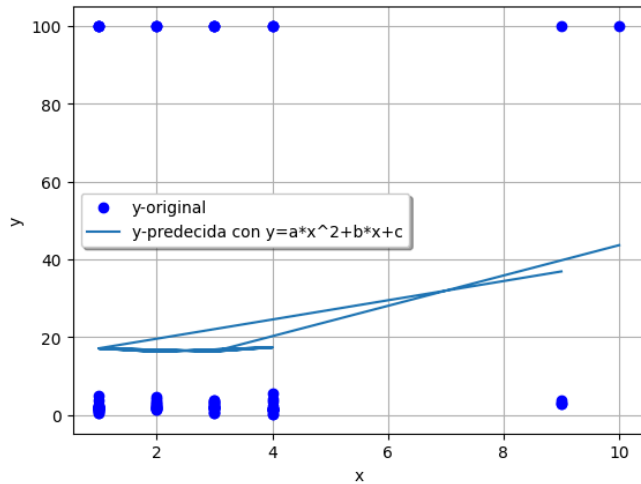
**dificultad (categorizado)**

Modelo lineal: 0.34

Función cuadrática con minijuego\_cat: 0.55



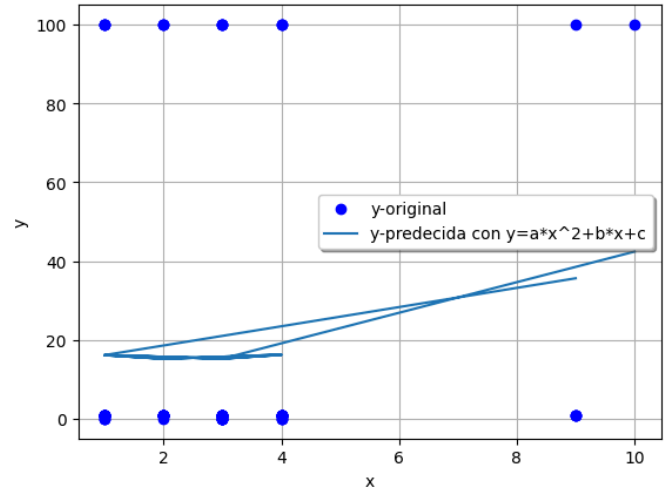
Kytzia:



**tiempo de interacción**

Modelo lineal: 0.12

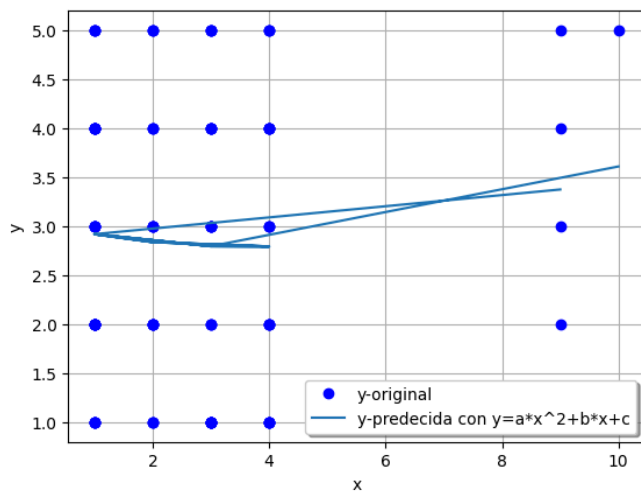
Función cuadrática con minijuego\_cat: 0.14



**botón correcto**

Modelo lineal: 0.11

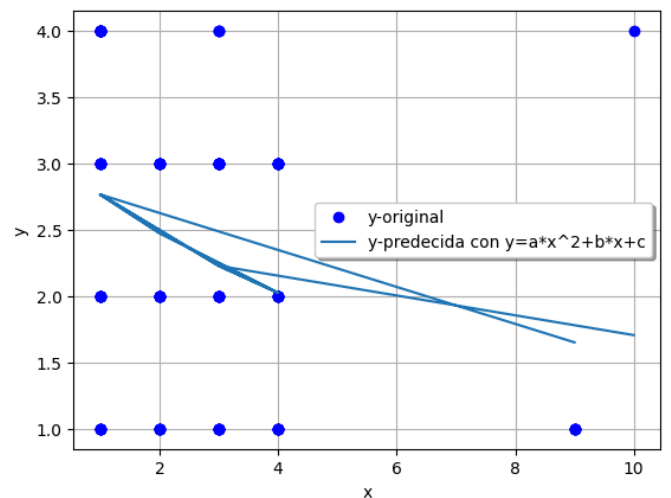
Función cuadrática con minijuego\_cat: 0.13



**color (categorizado)**

Modelo lineal: 0.06

Función cuadrática con minijuego\_cat: 0.10



**dificultad (categorizado)**

Modelo lineal: 0.29

Función cuadrática con minijuego\_cat: 0.31

### **Modelos de regresión no lineal:**

- En términos generales, cualquier modelo de regresión lineal logró superar o alcanzar a los modelos previos de regresión lineal.
- Algunos superaron solo por centésimas, mientras que otros mostraron una mejoría evidente respecto a los coeficientes.
- Los modelos con los que se obtuvieron mejores resultados:
  - Función cociente entre polinomios
  - Función polinomial inversa
  - Función cuadrática