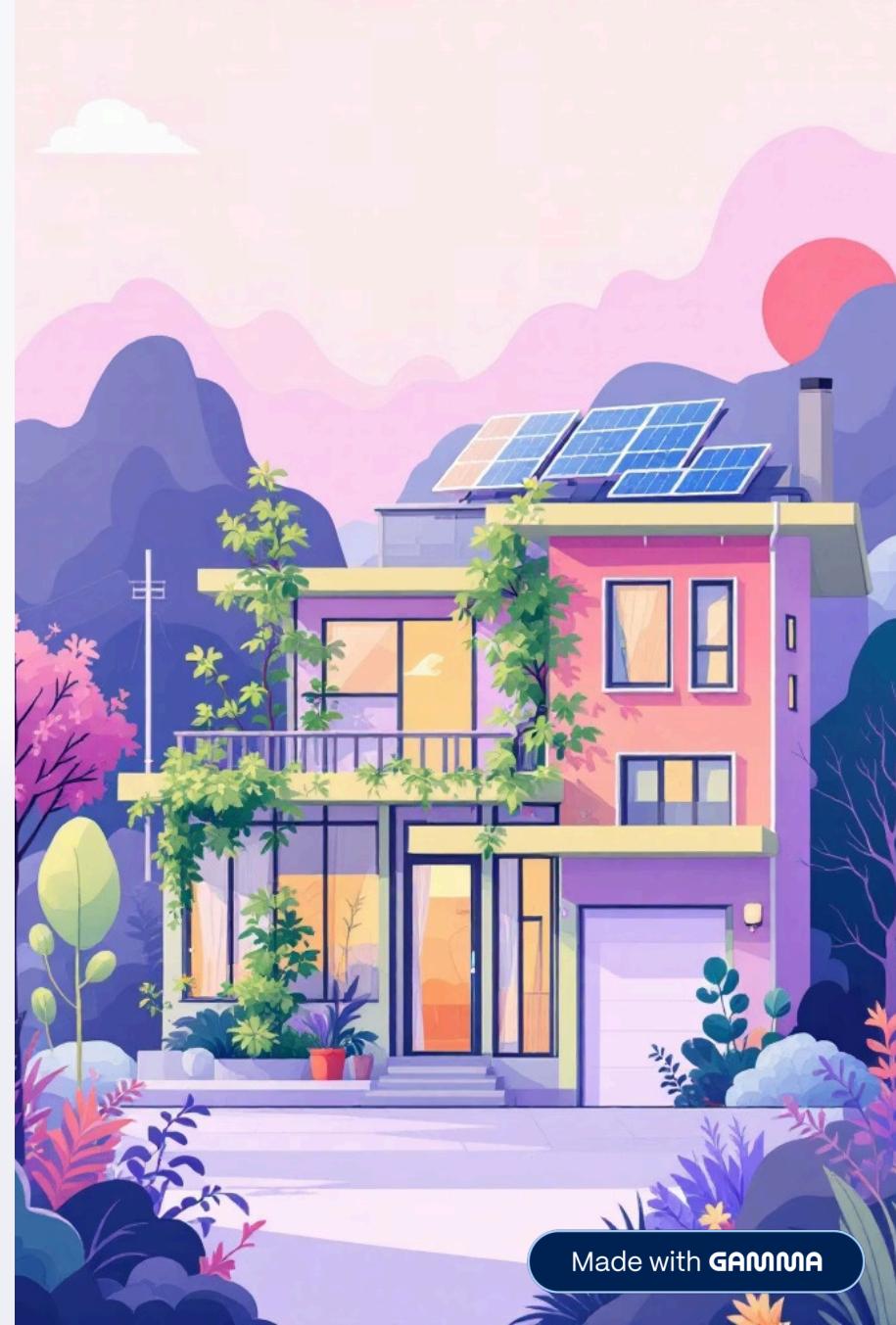


# **Predicción de Eficiencia Energética en Edificios Residenciales**

Presentación del Proyecto Final. Integración de Ingeniería de Datos,  
Regresión y Clasificación para optimizar el diseño arquitectónico sostenible.

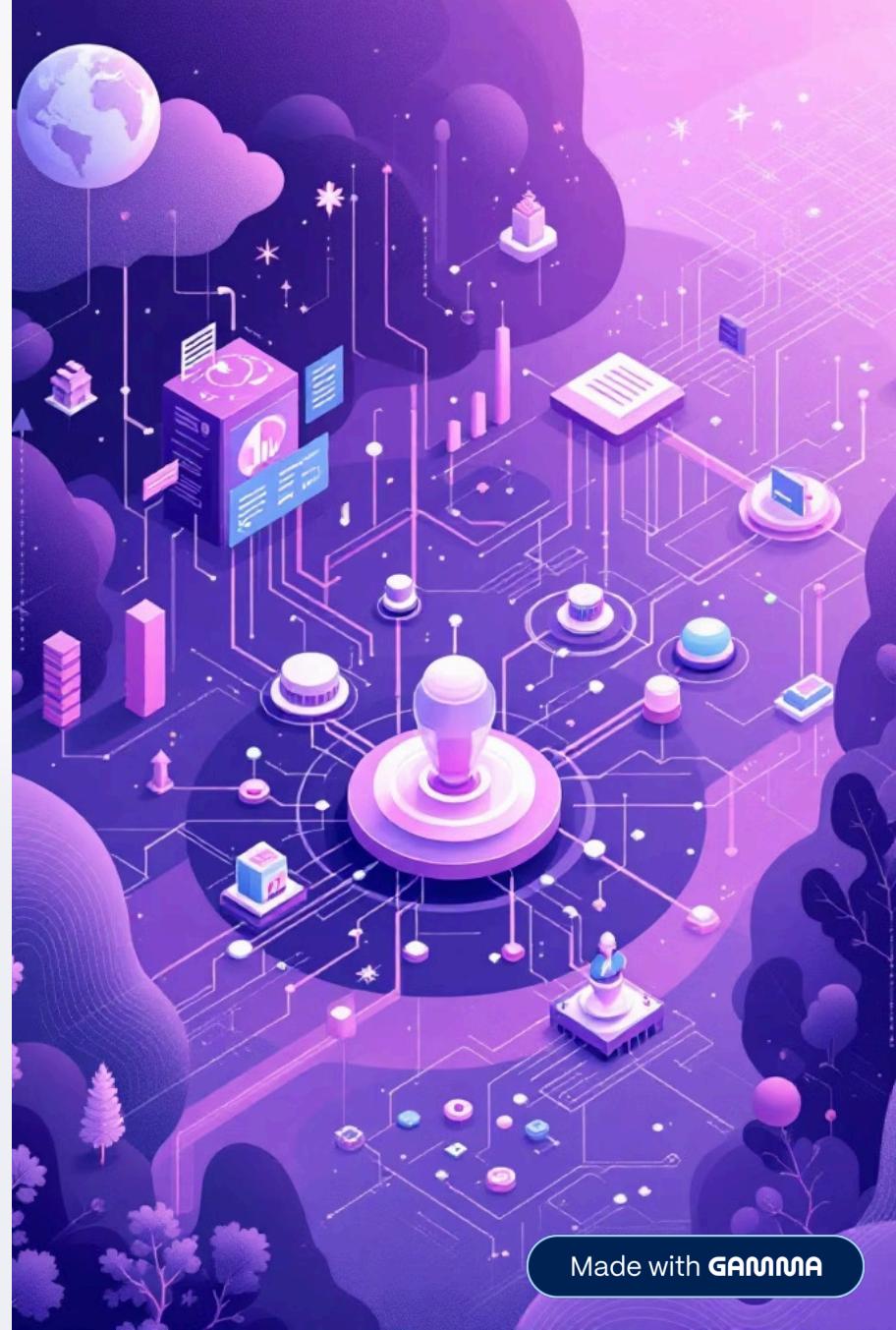


# Contexto del Proyecto

- El consumo de energía de los edificios ha aumentado de forma constante a nivel mundial.
- Los sistemas de calefacción, ventilación y aire acondicionado (HVAC) representan la mayor parte del uso de energía para regular el clima en estos espacios.
- Para diseñar edificios eficientes, los arquitectos necesitan calcular la carga de calefacción (HL) y la carga de refrigeración (CL) para dimensionar correctamente los equipos.
- Tradicionalmente, esto se hace usando software de simulación de energía que consume mucho tiempo y requiere conocimientos técnicos altamente especializados.

# Objetivos

- Desarrollar un marco de aprendizaje automático estadístico para estudiar el efecto de ocho características de diseño de un edificio (variables de entrada) sobre las cargas de calefacción y refrigeración (variables de salida).
- Investigar sistemáticamente qué tan fuerte es la asociación de cada variable de entrada con los resultados de consumo de energía.
- Comparar dos modelos para predecir nuestras variables para de esta manera hallar el mejor modelo para nuestros datos.



# Datos

## 768 edificios residenciales simulados

Se emplearon **768 edificios residenciales simulados** generados mediante el software Ecotect, el cual se estableció como la "verdad base" para el estudio. Para esta simulacion se tomo en cuenta como locacion la ciudad de Atenas (Grecia)

## Algunos datos de la simulacion

Todos los edificios partieron de cubos elementales y mantuvieron el mismo volumen exacto ( $771.75 \text{ m}^3$ ), pero se configuraron en 12 formas y dimensiones diferentes. Donde se considero a que los edificios estaban habitados por 7 personas

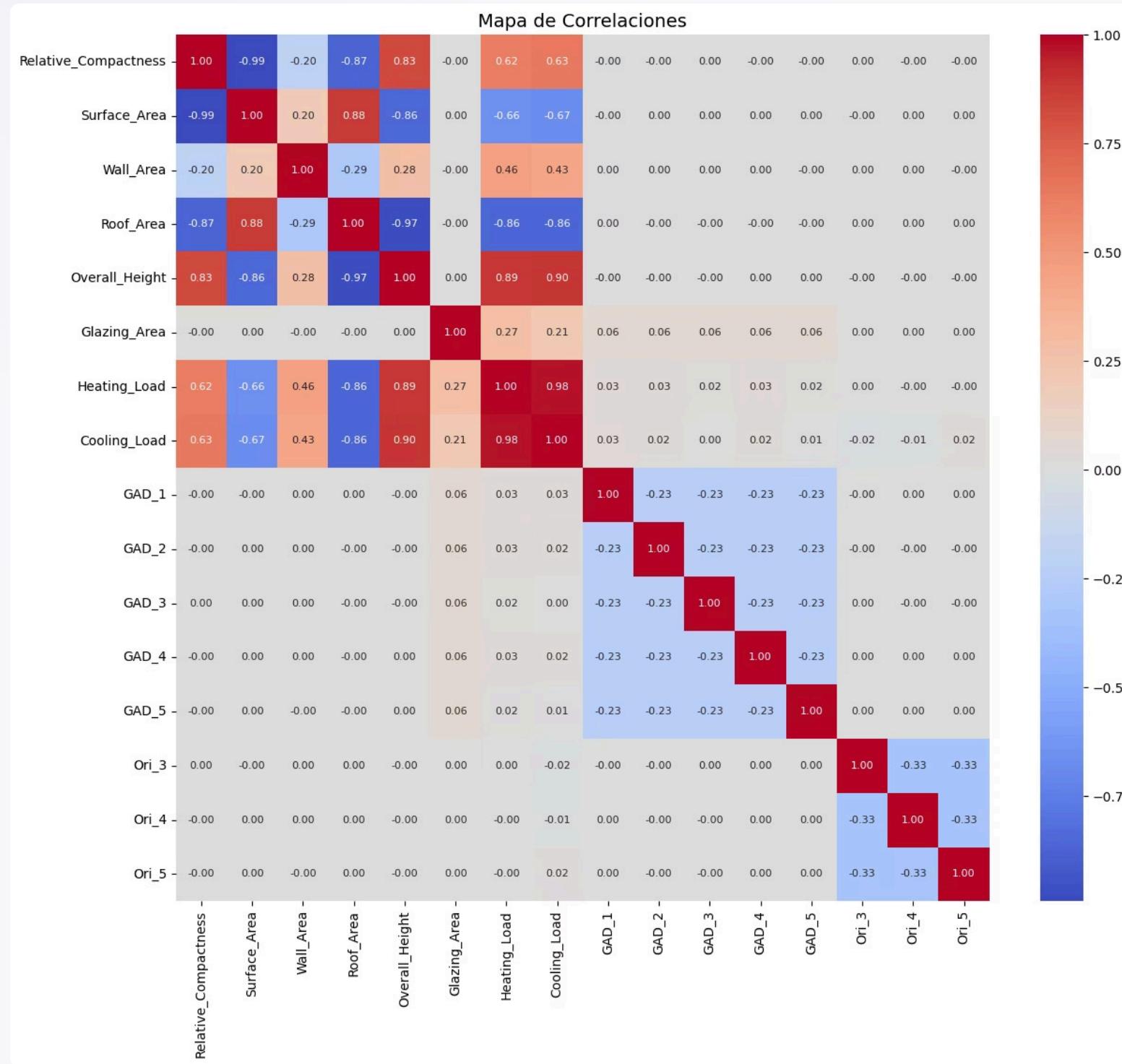
## 8 Variables de Entrada

Compacidad relativa, área de superficie, área de pared, área de techo, altura total, orientación, área de acristalamiento (ventanas) y distribución del acristalamiento.

## 2 variables de salida (y)

Carga de calefacción (HL) y carga de refrigeración (CL).

# Metodología



**Exploración inicial:** Se graficaron histogramas y diagramas de dispersión para explorar las propiedades estadísticas, comprobando que las variables no siguen una distribución normal (Gaussiana) y que su relación con el consumo de energía no es trivial.

# Metodología

- **Análisis estadístico:** Se aplicaron herramientas estadísticas como el coeficiente de correlación de rangos de Spearman y la Información Mutua (MI) para cuantificar la fuerza de la relación entre cada variable de diseño del edificio y su carga térmica.
- **Modelado (Machine Learning):** Se entrenaron y compararon dos enfoques de regresión para mapear las entradas con las salidas. El primero fue Mínimos Cuadrados Iterativamente Re-ponderados (IRLS), una técnica robusta que ajusta pesos para disminuir el efecto de los valores atípicos (outliers). El segundo modelo empleado fue la **Regresión Ridge**, una técnica de regularización lineal que añade una penalización a los coeficientes grandes; esto es particularmente útil en este contexto para manejar la alta correlación (multicolinealidad) que existe entre algunas variables de entrada, como el área del techo y la altura.



# Metodología

- **Validación del modelo:** Para comprobar el rendimiento de generalización de los modelos frente a datos desconocidos, se utilizó la técnica de validación cruzada (cross validation) de 10 pliegues, repitiendo el proceso de entrenamiento y prueba 100 veces con los datos permutados aleatoriamente.
- **Evaluación de errores:** El desempeño final de las predicciones fuera de la muestra se evaluó registrando el Error Absoluto Medio (MAE), el Error Cuadrático Medio (MSE) y el Error Relativo Medio (MRE).



# Resultados

Tras el preprocesamiento de los datos y el entrenamiento de los modelos de Regresión Lineal, Regresión Ridge y Regresión Logística, se identificaron los siguientes resultados clave sobre el comportamiento térmico de las edificaciones:

- **Impacto crítico de la geometría:** La Compacidad Relativa y la Altura Total son los predictores dominantes del modelo. Los edificios más altos (por ejemplo, de dos niveles o 7 metros) presentan una carga de calefacción significativamente superior frente a los de menor altura.
- **El efecto del acristalamiento:** La variable del área real de vidrio (*Glazing Area*) mostró una correlación positiva directa con el consumo de energía. El modelo confirma que, térmicamente, el vidrio actúa como un puente menos eficiente que una pared sólida.
- **Orientación y distribución:** Gracias al tratamiento de las variables categóricas (*One-Hot Encoding*), el modelo detectó que las fachadas orientadas al Norte incrementan la carga de calefacción. Además, una distribución uniforme de las ventanas equilibra mejor el requerimiento térmico frente a diseños que concentran todo el vidrio en una sola fachada.
- **Solución de multicolinealidad:** La implementación de la Regresión Ridge estabilizó los coeficientes del modelo, controlando con éxito la alta correlación encontrada entre variables físicas inversamente proporcionales, como el área de techo y la altura.

# Metricas

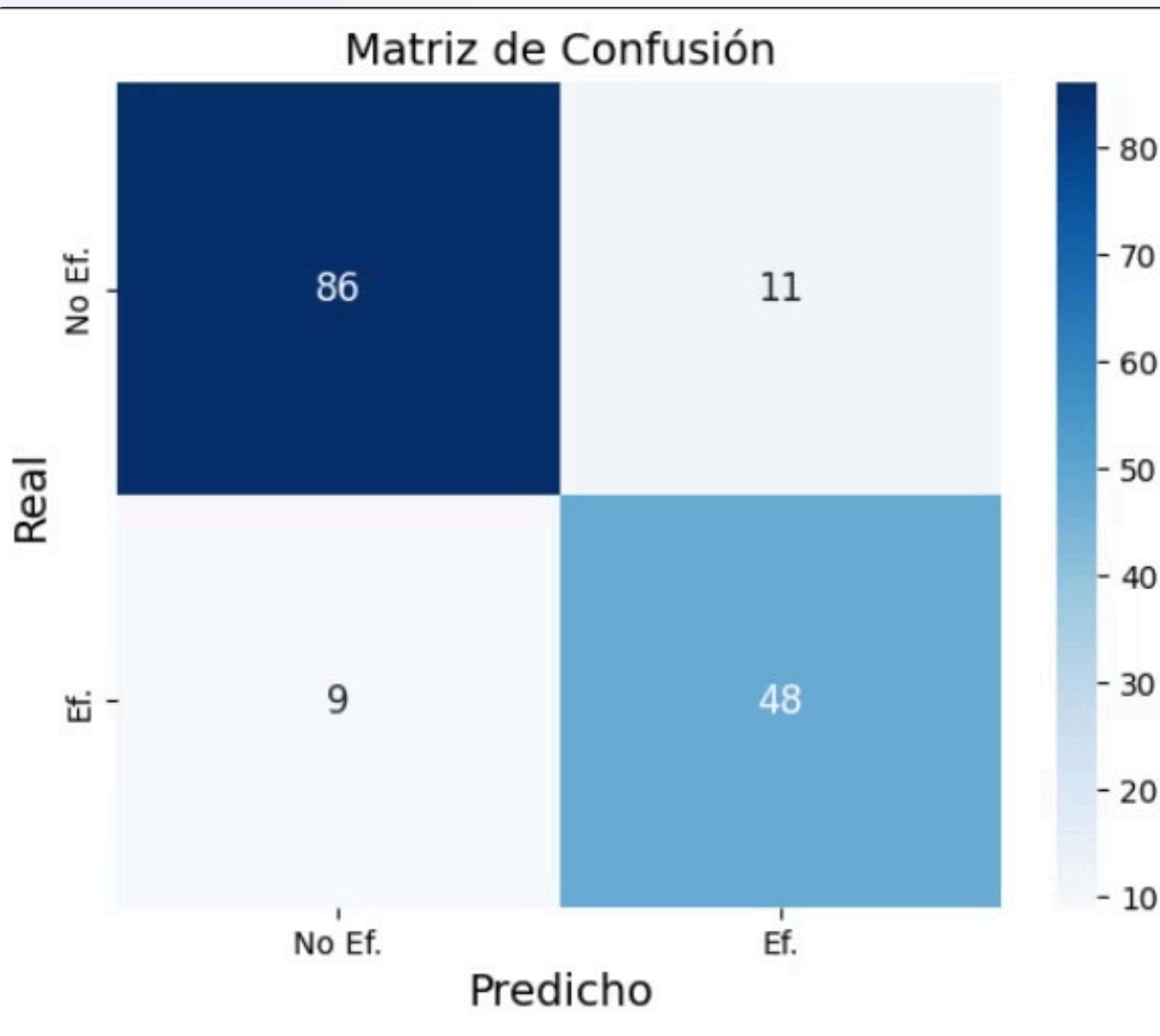
Para garantizar la fiabilidad de las predicciones en este conjunto de 768 muestras, se evaluaron los modelos con las siguientes métricas estadísticas rigurosas:

Se utilizó el Error Cuadrático Medio de la Raíz (RMSE) mediante una técnica de Validación Cruzada (K-Fold) con 5 iteraciones.

Esto permitió obtener una estimación del error promedio en las mismas unidades que la variable objetivo (Heating Load), asegurando que el modelo generalice bien y no dependa de una división aleatoria de los datos.

Para evaluar la capacidad del modelo de Regresión Logística al separar edificios "Eficientes" (1) y "No Eficientes" (0), se utilizó el Classification Report, obteniendo métricas clave de Precision, Recall y F1-Score.

# Metricas



Adicionalmente, se generó una Matriz de Confusión para visualizar los verdaderos positivos frente a los falsos negativos en la clasificación de los diseños.



# Conclusiones

01

---

## Prediccion mediante machine learning

Las herramientas de aprendizaje automático (machine learning) pueden predecir con gran precisión la carga de calefacción y refrigeración de los edificios residenciales basándose únicamente en ocho variables de diseño

02

---

## Detección Sobresaliente

Si bien una evaluación superficial de las métricas muestra un RMSE casi idéntico entre la Regresión Lineal Múltiple (3.0333 con desviación de 0.4647) y la Regresión Ridge (3.0392), el modelo Ridge superó decisivamente al método clásico en la **calidad estructural de sus predicciones**. La regresión lineal tradicional fallaba al lidiar con variables altamente correlacionadas (como el área de techo y la altura). Al introducir la penalización matemática (regularización L2), el modelo Ridge estabilizó los pesos de estas variables, garantizando que el algoritmo sea confiable y generalizable al aplicarlo en futuros diseños arquitectónicos que el modelo nunca antes ha visto.

03

---

## Variables mas correlacionadas

La Compacidad Relativa y la Altura Total resultaron ser predictores dominantes. Los edificios con mayor altura (7 metros o dos niveles en este estudio) presentan una carga de calefacción significativamente superior. Esto indica que, para el mismo volumen, las construcciones más esbeltas pierden calor con mayor facilidad que aquellas más compactas.

Asi tambien la variable de ingeniería Glazing Area (superficie real de vidrio) mostró una correlación positiva directa con el consumo. Si bien la luz natural es deseable, el modelo confirma que el vidrio actúa como un puente térmico menos eficiente que la pared sólida.

# Recomendaciones

## 1 Ratio Ventana-Pared (WWR)

Limitar las proporciones de acristalamiento superiores al 25% en fachadas con alta exposición solar, a menos que se implementen medidas de mitigación estrictas

## 2 Controlar el impacto de la altura

Puesto que añadir niveles aumenta drásticamente la demanda de calefacción, es estrictamente necesario reforzar el aislamiento térmico en techos (*Roof Area*) para compensar la pérdida de calor en construcciones de múltiples pisos.

## 3 Priorizar la compacidad del diseño

En zonas climáticas frías, se deben favorecer diseños con alta compacidad relativa (estructuras más cercanas a un cubo o esfera) para reducir el área de superficie expuesta a la intemperie.

