

Energía y Edificación

Volumen 49, junio de 2012

Estimación cuantitativa precisa del rendimiento energético de edificios residenciales utilizando herramientas estadísticas de aprendizaje automático.

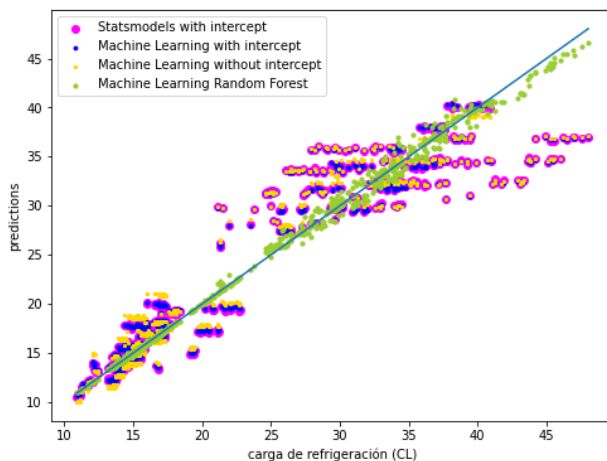
Athanasios Tsanas, Angeliki Xifara

Yuleidis Mesa, Ciencias básicas, Universidad Tecnológica de Bolívar, Colombia, e-mail: ymesa@utb.edu.co

RESUMEN

Objetivo: Esta investigación desarrolla un marco de aprendizaje automático estadístico para estudiar el efecto de ocho variables independientes: compacidad relativa, área de superficie, área de pared, área de techo, altura total, orientación, áreas de acristalamiento y la distribución del área de acristalamiento, sobre dos variables a saber, la carga energética de calefacción (HL) y carga de refrigeración (CL), de edificios residenciales.

Material y método: El conjunto de datos comprende 8 características para 768 muestras, con el objetivo de predecir dos respuestas de valor real, sin embargo, para efectos de este trabajo solo se trabajará la carga energética de refrigeración como variable a predecir. Se usaron diferentes técnicas estadísticas; desde la visualización de diagramas de correlación hasta regresiones por métodos de Mínimos Cuadrados Ordinarios, Machine learning y Random Forest; con el fin de medir la fuerza de asociación de cada una de las variables independientes con la variable carga de refrigeración (CL), y así identificar las variables más fuertemente relacionadas.



Resultado: Las simulaciones realizadas para los datos de los 768 edificios residenciales muestran que podemos predecir la carga energética de refrigeración (CL) con un error medio bajo. Los resultados de este estudio respaldan la viabilidad de usar herramientas de aprendizaje automático como Machine learning y Random Forest para estimar los parámetros de la regresión, donde el Score pasó de 0.887 usando metodología OLS a 0.99 usando esta última herramienta de aprendizaje automático (Random Forest) y

disminuyó el error de predicción en un 82% al estimarlo en 0.57. El diagrama, resume visualmente, cada una de las metodologías de regresión utilizadas y el ajuste de los datos.

Conclusiones: para este conjunto de datos, mecanismos no paramétricos como el Random Forest permiten una mejor estimación de los parámetros, asimismo se identificaron que las variables con mayor grado de asociación con la carga de refrigeración (CL), son la compacidad relativa (de forma inversa), que mide proximidad de los componentes que configuran el inmueble y la altura total del mismo (en forma directa).