

Modelo inteligente para la predicción de la demanda doméstica de energía en clientes sin telemedida

Óscar Poncelas Villabrille
Empleado en Indra Sistemas

Cátedra TotalEnergies de Analítica de Datos e Inteligencia Artificial



Acceso al documento



Introducción

En el sistema eléctrico español, las comercializadoras ofertan un precio de la energía eléctrica a los clientes domésticos que debe ser competitivo en el mercado y a la vez rentable para la empresa.

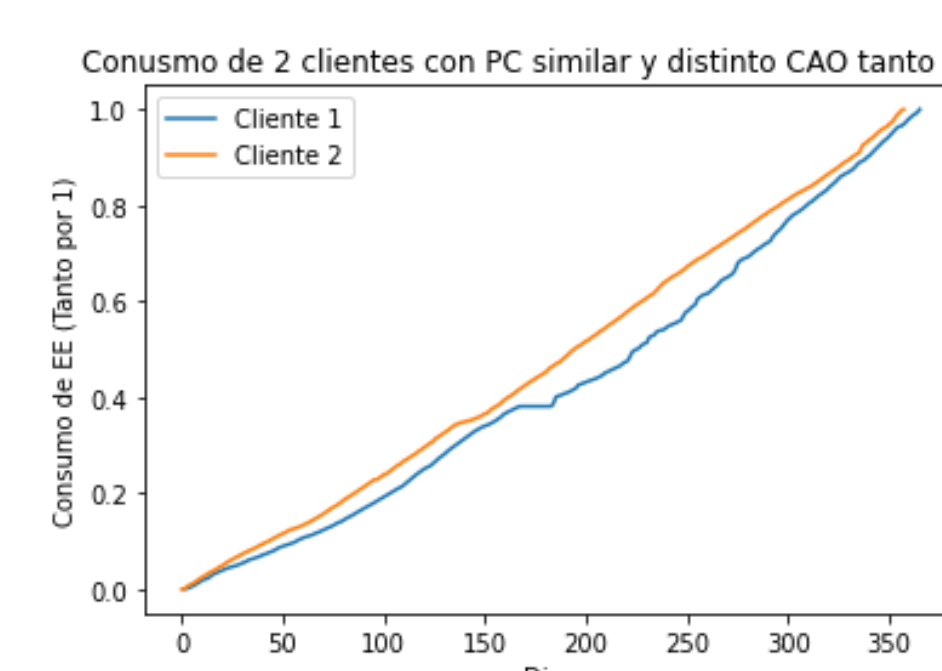
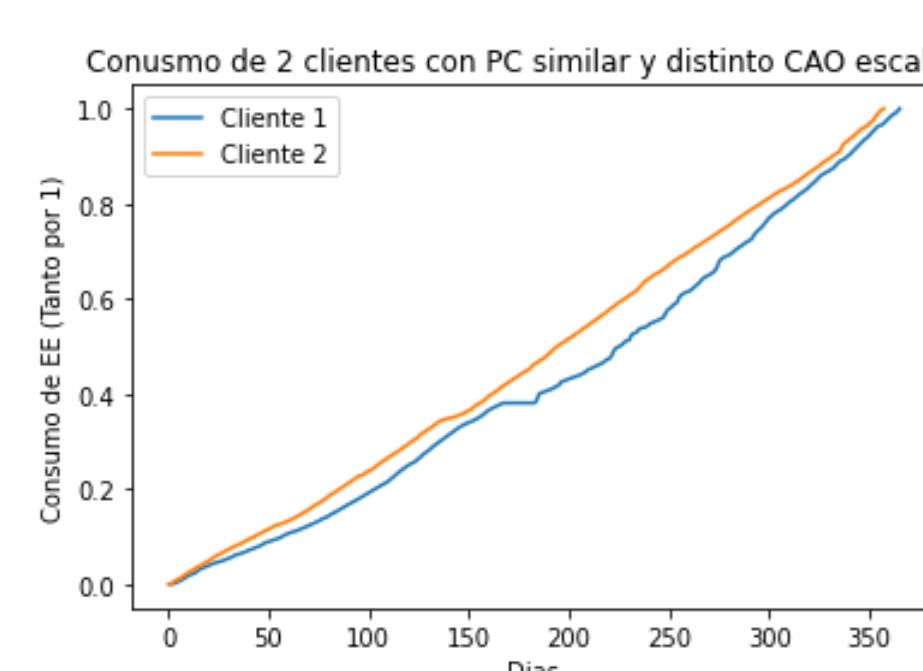
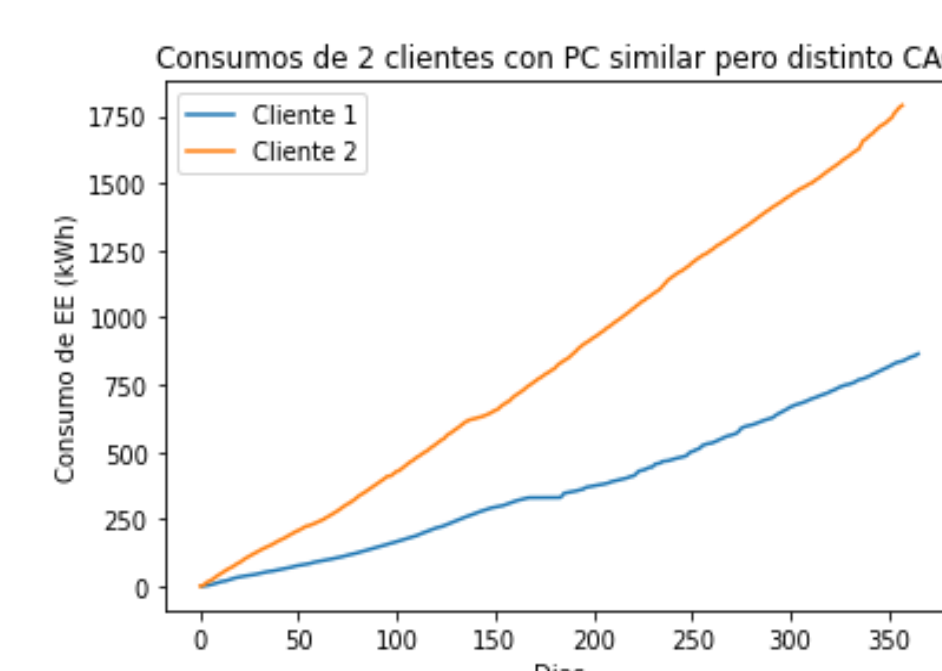
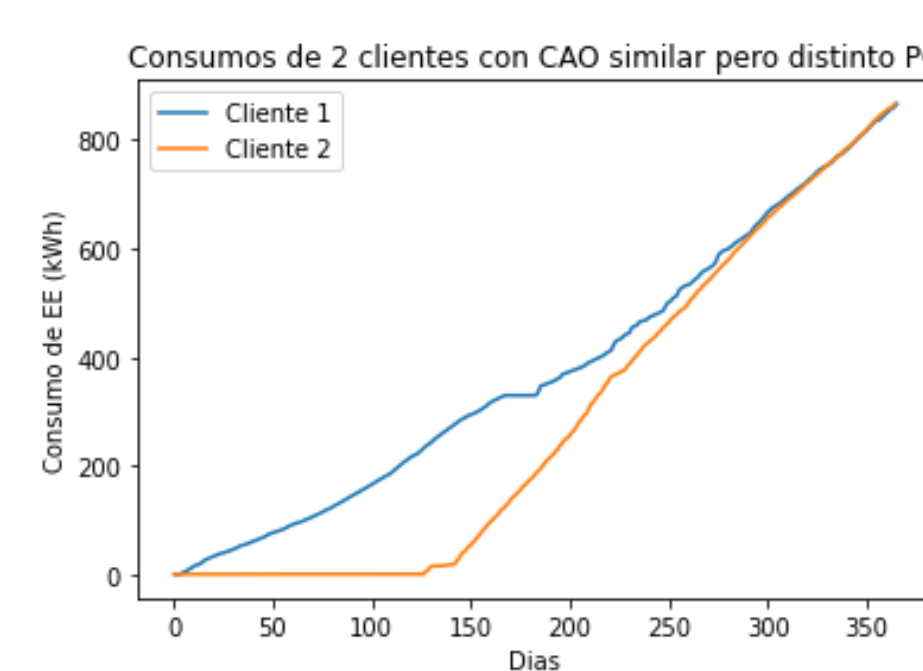
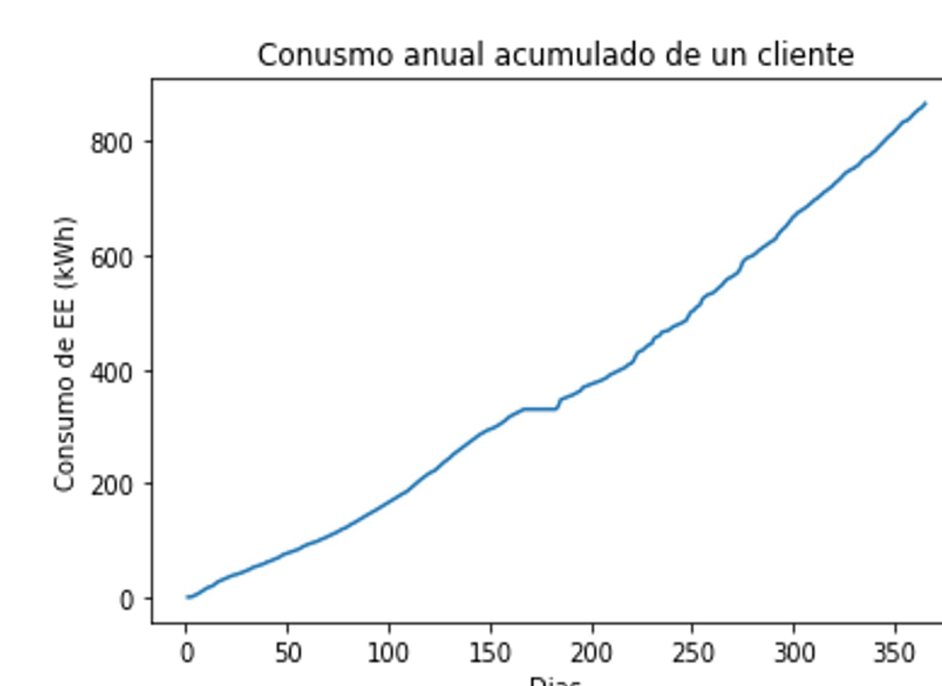
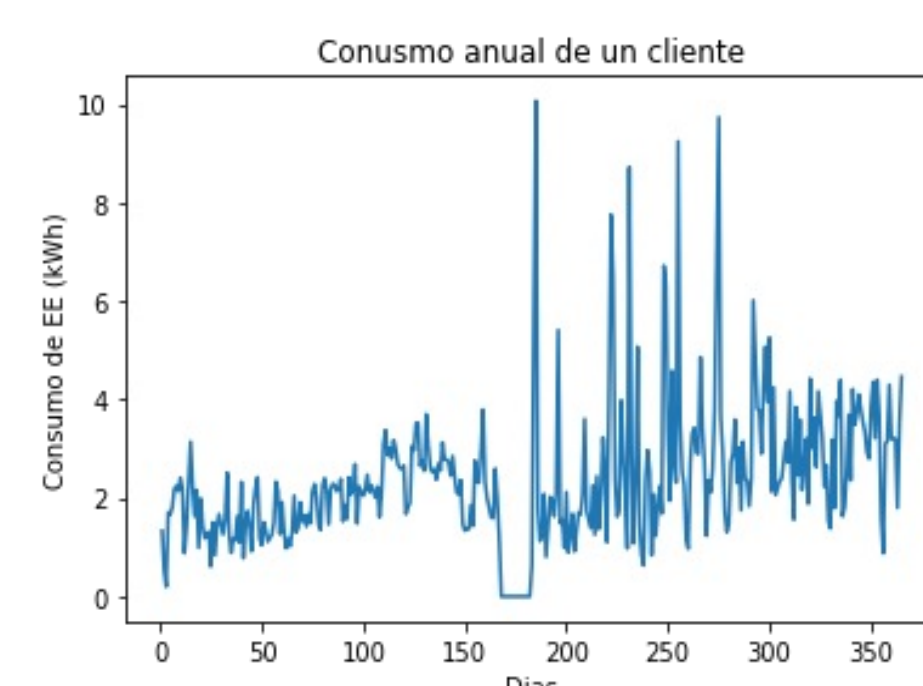
El precio final ofertado depende de diversos factores que, además del coste de la propia energía consumida, incluyen peajes y otros cargos que emite el distribuidor.

Objetivos del estudio

En este contexto, la predicción del consumo de los clientes en cada franja horaria es esencial, pues permite anticiparse a los costes asociados a cada cliente y ajustar de forma eficaz el precio ofertado por la comercializadora

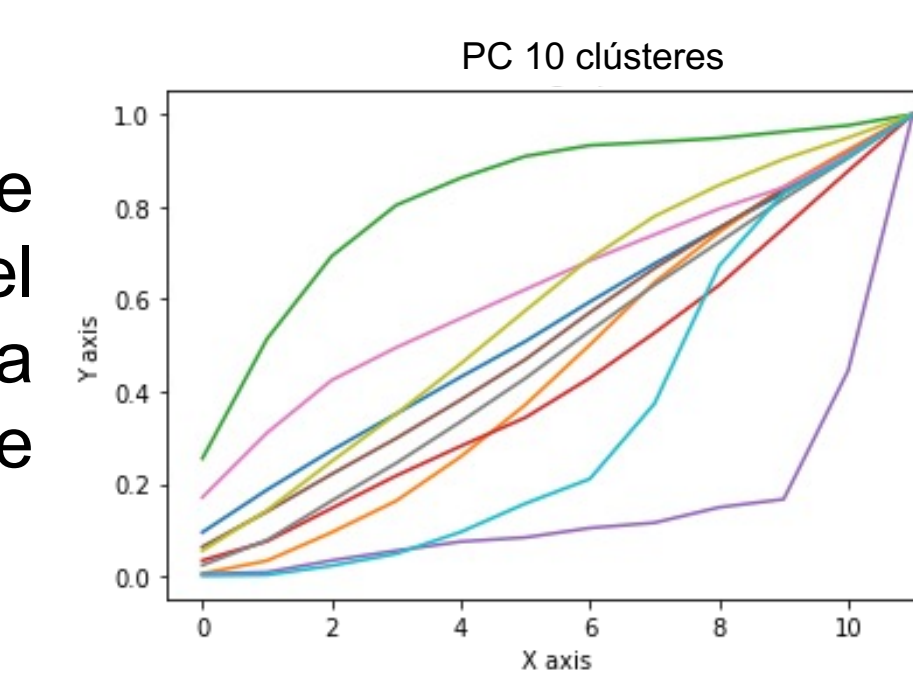
Métodos

En este proyecto se desea desarrollar un módulo predictivo con técnicas de inteligencia artificial explicable (xAI) para aquellos clientes de los que no se disponga de información completa acerca de sus consumos horarios (clientes cuyo contador no disponga de telemedida).

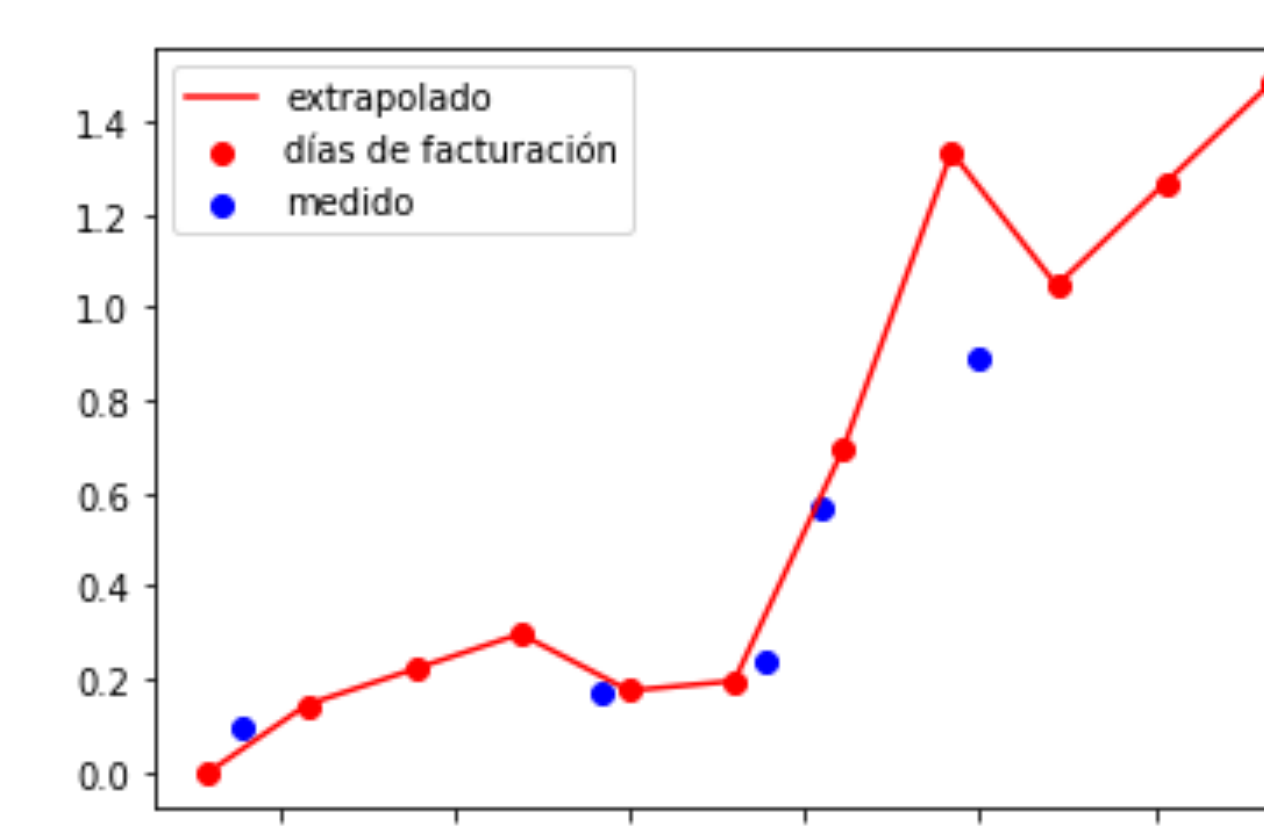
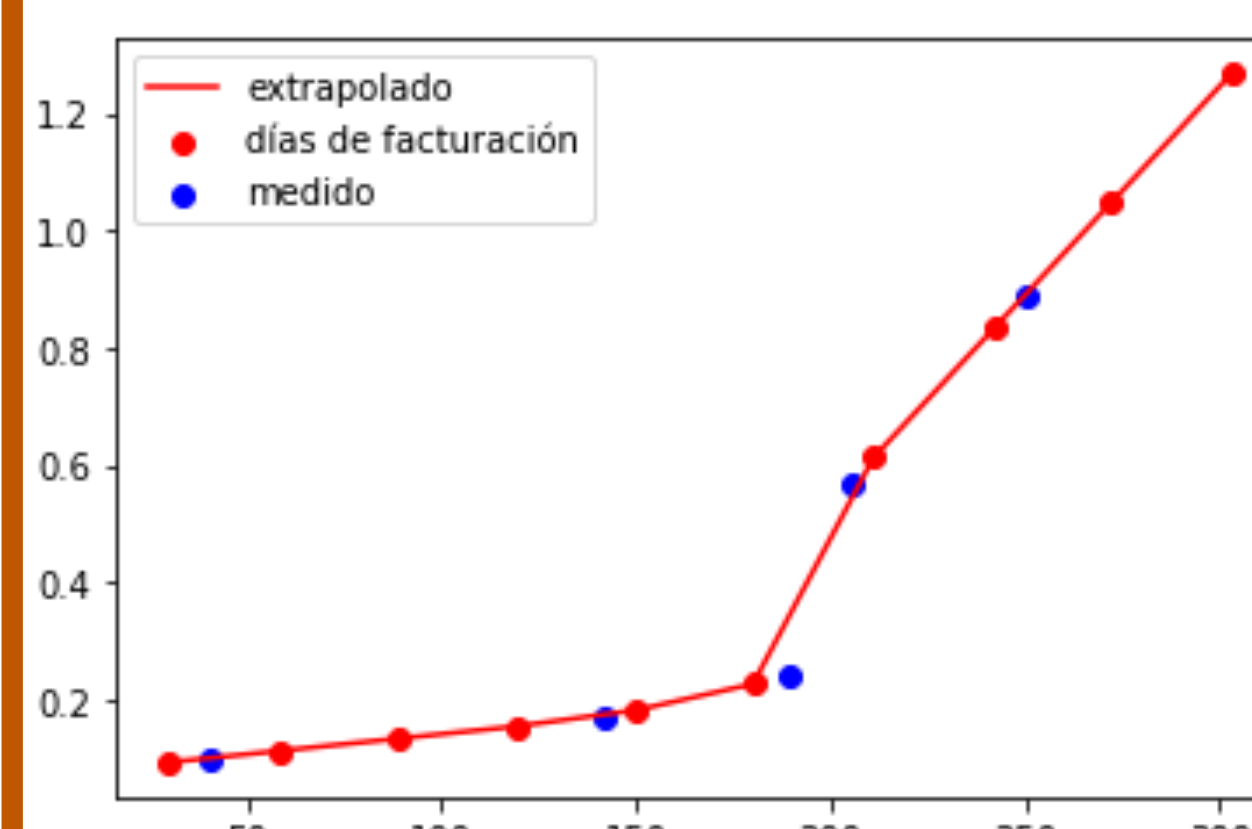
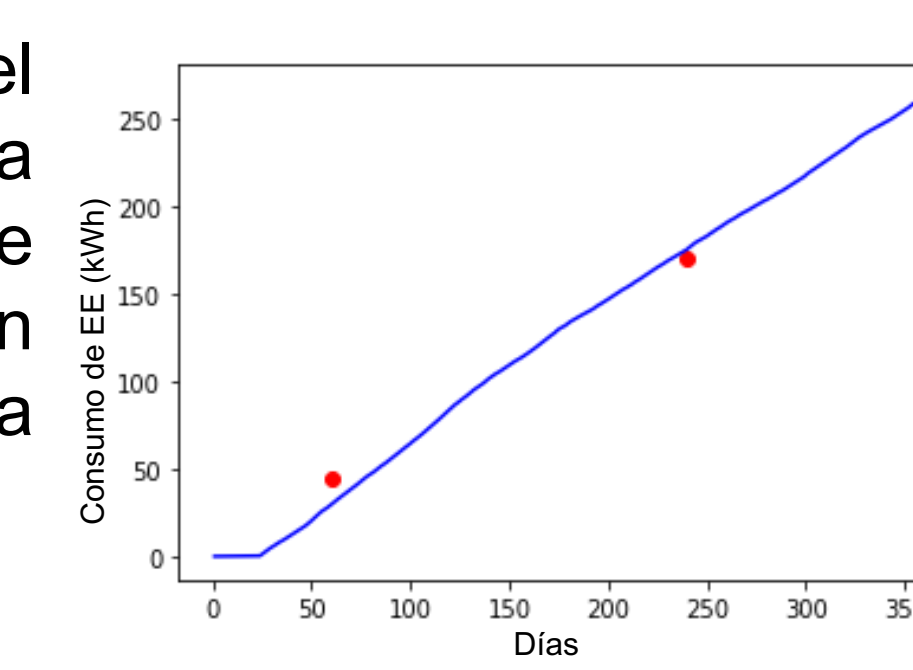


Figuras y Resultados

Tras aplicar el algoritmo de clustering se obtienen N perfiles de consumo, siendo N el número de clústeres parametrizable, y a cada cliente se le asigna el que mejor represente su propio PC.



Generando un conjunto de datos a partir del original en el que en lugar de una medida diaria para cada cliente, se tienen únicamente unas pocas, se puede estimar con gran precisión su consumo cualquier día del año ya que se conoce a que cluster pertenece.



Esto supone una gran ventaja frente al modelo lineal utilizado anteriormente. El modelo lineal funciona extrapolando las medidas realizadas manualmente por un técnico cada cierto tiempo (puntos azules) al resto de días. En la gráfica de la izquierda la predicción (línea roja) se ajusta muy bien a las medidas ya que se utilizan todas, siendo así una interpolación. Esto en un escenario real no sucede ya que a la hora de emitir una factura únicamente se tienen las medidas realizadas con anterioridad, teniendo que extrapolar por necesidad y cometiendo un error más significativo

Errores modelo lineal

N.º de medidas	2	3	4	5	6	7	8	9
Error	0.17486	0.09914	0.06758	0.05137	0.04420	0.03684	0.03394	0.03245

Errores modelo clustering

N.º clústeres	10	20	30	40	50	60	70
Error	0.02250	0.01606	0.01359	0.01142	0.00984	0.00873	0.00772

En la primera tabla se muestra el error obtenido dependiendo del número de medidas utilizadas, que como cabía esperar se reduce cuantas más se tienen. En la segunda todos los errores son los obtenidos con una única medida, y no mejora cuantas más se tengan, sino cuanto mayor sea el número de clústeres que se utilicen en el algoritmo.

Conclusiones

Se obtienen unos resultados realmente interesantes. Como cabía esperar, cuantos más clústeres se utilizan para categorizar a los clientes, mejores resultados se obtienen. Pero, a diferencia del modelo lineal, es capaz de mantener el error que comete cuando cuenta con un gran número de muestras a las situaciones en las que solo cuenta con una única medida.

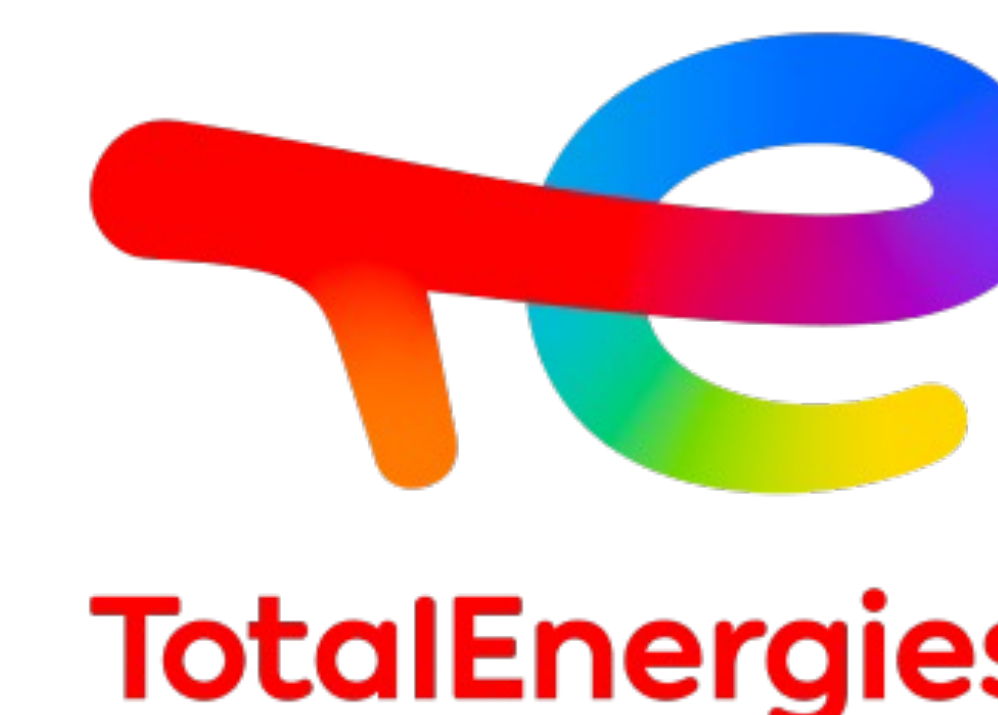
Esta estabilidad puede deberse a que la mayoría de los clientes tiene uno de los perfiles más lineales, y los ajusta de forma muy exacta. El número de clientes cuyo PC puede ser más problemático a la hora de estimarlo son escasos por lo que el error cometido con estos puede verse atenuado al aplicar la media aritmética al conjunto de errores. Pero el modelo lineal utiliza el mismo conjunto de datos y las diferencias en los resultados son notables, especialmente cuando se tienen pocas medidas. En conversaciones con TE, la empresa interesada en el presente estudio, han recalcado el interés en que el modelo sea preciso con clientes con PC no lineales, así como en escenarios en los que se cuenta con pocas medidas, ya que son los que más dificultades presentan utilizando su modelo de estimación. Y esos son precisamente los casos en los que el modelo obtenido genera los mejores resultados.

El modelo basado en clústeres es incluso capaz de obtener tan buenos resultados con una única medida, que en muchos escenarios reales es de lo único que se dispone. El modelo lineal, a parte de ser peor, necesita al menos dos medidas para poder aplicar el algoritmo de extrapolación, ya que con un único punto no es posible

Trabajo futuro

El estudio se ha realizado con datos públicos de consumidores en América del Sur. Aunque estos datos permiten validar el algoritmo de forma reproducible, el siguiente paso es adaptar esta técnica a los datos reales de demanda doméstica de la comercializadora, y realizar un análisis cluster de los mismos.

El siguiente paso es realizar un predictor del cluster en función de variables medibles del cliente, como la superficie de su vivienda, la potencia contratada, la temperatura de la zona, el que sea una vivienda habitual o una segunda vivienda, etc.



Universidad de Oviedo