

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 680 769**

21 Número de solicitud: 201731099

51 Int. Cl.:

**G06F 17/00** (2006.01)

**G06Q 50/06** (2012.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

**08.09.2017**

30 Prioridad:

**07.03.2017 ES P201700228**

43 Fecha de publicación de la solicitud:

**10.09.2018**

71 Solicitantes:

**UNIVERSIDAD DE MURCIA (100.0%)**  
**Avda. Teniente Flomesta nº 5**  
**30003 Murcia ES**

72 Inventor/es:

**SKARMETA GÓMEZ, Antonio Fernando;**  
**ZAMORA IZQUIERDO, Miguel Angel;**  
**MORENO CANO, María Victoria y**  
**GONZÁLEZ VIDAL, Aurora**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

54 Título: **Procedimiento para predecir el consumo energético de climatización ambiental en edificios**

57 Resumen:

Procedimiento para predecir el consumo energético de climatización ambiental en edificios.

El procedimiento propuesto consiste en llevar a cabo una serie de pasos consecutivos en los que se aplican diferentes técnicas matemáticas de inteligencia artificial para transformar los datos disponibles sobre: (i) la predicción meteorológica en su zona geográfica; (ii) el consumo energético del sistema de climatización; y, (iii) el nivel de ocupación en el edificio, en información útil para representar la evolución horaria del consumo energético del sistema de climatización ambiental del edificio.

Con el procedimiento de la invención es posible seleccionar de forma automática las características que compondrán el conjunto óptimo de entradas del modelo predictivo, así como la técnica de regresión que estime con mayor precisión la evolución horaria del consumo energético asociado al sistema de climatización ambiental del edificio.

1. Limpiar datos
2. Transformar datos en características y generar conjuntos de entrada
3. Generamos conjuntos de valores para los parámetros de configuración
4. Normalizamos conjuntos de entrada
5. Dividimos conjuntos de entrada y salida en subconjuntos de entrenamiento y subconjuntos de validación
6. Aplicamos a los conjuntos de entradas componentes principales al 95% de la varianza
7. Entrenamos técnicas con diferentes conjuntos de entradas y diferentes valores de configuración
8. Proselección de modelos en función del RMSE y el R Square
9. Selección final del modelo óptimo en función del CVRMSE

FIG. 1

ES 2 680 769 A1

## DESCRIPCIÓN

Procedimiento para predecir el consumo energético de climatización ambiental en edificios

### 5 **Campo técnico de la invención**

La presente invención se refiere al campo del modelado energético de los sistemas de climatización ambiental en edificios residenciales o de oficinas.

Se parte de la definición universal de sistema de climatización ambiental de un edificio, sistema que incluye como mínimo tres parámetros de control: temperatura, humedad relativa  
10 y caudal de renovación ambiental.

### **Antecedentes de la invención**

El coste que actualmente representa la energía es de vital importancia en una especialidad como la climatización que requiere elevados consumos, sea de energía eléctrica, sea de  
15 combustible, por lo que su reducción representa una de las premisas básicas en los criterios de diseño y operación de los sistemas de climatización ambiental en edificios residenciales.

Los recientes avances en Tecnologías de la Información y las comunicaciones (TICs), entre ellos el paradigma del Internet de las Cosas (IoT), han derivado en grandes despliegues de  
20 sensores, infraestructuras de computación y, en consecuencia, a la proliferación de datos proporcionados por sensores. En todos los aspectos de la vida diaria, esto abre grandes oportunidades para analizar y mejorar la eficiencia de los sistemas de gestión de infraestructuras energéticas propuestos hasta ahora, así como para proporcionar nuevos e innovadores servicios.

Existen estudios que demuestran que durante la fase operativa de un edificio se produce el 90% de su consumo energético total alcanzado al final de su ciclo de vida. Del total de este consumo energético asociado a la fase operacional en un edificio residencial o de oficinas, el 76% es debido a la climatización. Por lo tanto, es necesario proponer soluciones que  
30 aseguren una gestión continua y eficiente del uso energético de las infraestructuras de climatización de estos edificios.

Las soluciones propuestas hasta ahora para ahorrar energía durante la fase de operación de edificios residenciales o de oficinas pueden ser agrupadas en dos categorías: (a) aquellas  
35 basadas en herramientas de simulación que requieren del modelo físico del edificio; (b) las soluciones basadas en el modelado de datos e identificación de patrones a través de la

aplicación de técnicas matemáticas de inteligencia artificial. La primera categoría propone soluciones complejas que suelen necesitar de mucha información física del edificio que no siempre está disponible ni es accesible. Además, estas soluciones suelen necesitar mucha información sensorial que permita el calibrado del modelo físico del edificio, lo cual supone una gran carga computacional. La segunda categoría propone soluciones que no requieren disponer de información específica sobre la estructura física y las características del edificio, sino que sólo necesita disponer de información específica medida por sensores instalados en el edificio. Además, la evaluación experimental que los sensores ofrecen en general proporcionará una mejor descripción del sistema puesto que las variables simuladas quedan sujetas a la descripción finita que el investigador proponga, dejando siempre una parte caótica o inexplicada que los modelos matemáticos sí pueden recoger.

A la predicción del consumo energético está enfocada la patente US2012083930, que describe la creación de un modelo que depende del coste y del tiempo. En KR101168153 se utilizan redes neuronales basándose en históricos de consumo y variables meteorológicas. En S2014058572 se utilizan de nuevo variables meteorológicas como entradas de algoritmos matemáticos. Hasta ahora sin embargo no es posible generar de forma automática el modelo predictivo óptimo del consumo energético asociado a la climatización ambiental utilizando las predicciones meteorológicas y el nivel de ocupación del edificio.

20

### **Resumen de la invención**

Con el objetivo último de optimizar el consumo energético a través de estrategias de ahorro energético basadas en controlar el modo de operación de los sistemas de climatización ambiental, la presente invención tiene por objeto generar el modelo predictivo óptimo del consumo energético horario asociado a dichos sistemas. Para generar de forma automática el modelo predictivo óptimo de dicho consumo de climatización ambiental en edificios residenciales o de oficinas se dispone de información horaria sobre: (i) la predicción meteorológica en su zona geográfica; (ii) el consumo energético del sistema de climatización; y, (iii) el nivel de ocupación en el edificio. El consumo energético al que nos referimos en esta invención es independiente al tipo de energía, es decir, ya sea ésta eléctrica o por combustible.

30

Como predicción meteorológica nos estaremos refiriendo como mínimo a la predicción de la temperatura ambiental, pero ésta podría abarcar - sin alterar el método propuesto - a la predicción de la humedad ambiental, la radiación solar y la velocidad del viento.

35

Diferentes técnicas de regresión se han venido aplicando con notable éxito hasta ahora al problema del modelado energético en este tipo de edificios. De entre todas las técnicas aplicadas a este problema, se han identificado cinco de ellas capaces de proporcionar predicciones del consumo energético con un bajo y similar error de predicción. Las cinco técnicas a destacar son: (1) Red neuronal perceptrón multicapa (MLP); (2) Máquina de vectores soporte (SVM); (3) Función Gauss con núcleo de base radial (Gauss); (4) Red neuronal con función bayesiana regularizada (BRNN); y (5) Random Forest (RF).

De entre estas cinco técnicas, seleccionar aquella que proporciona los resultados más precisos en un determinado problema no es trivial. Cada técnica de regresión tendrá una configuración óptima de sus parámetros de ajuste, diferente según las características del problema. Además, la solución óptima a un edificio concreto también dependerá del conjunto de entradas considerado para generar el modelo.

Aplicando el procedimiento propuesto en la presente invención es posible analizar de forma automática cuál es el comportamiento de cada una de las técnicas de regresión propuestas, con diferentes valores para sus parámetros de configuración y diferentes conjuntos de entradas.

El comportamiento de cada combinación “técnica-configuración-conjunto de entradas” se evaluará en función de dos métricas calculadas durante la fase de entrenamiento del modelo predictivo, utilizando el conjunto de datos de entrenamiento (*training*). Estas métricas son: el error cuadrático medio de la estimación (RMSE, E.1) y el coeficiente de determinación (R Square, E.2).

$$(E.1) \quad RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Siendo  $y_i$  el valor puntual del consumo (KW/h),  $\hat{y}_i$  la estimación de éste y  $n$  el número de estimaciones.

$$(E.2) \quad \rho^2 = 1 - \frac{\sigma_r^2}{\sigma^2}$$

Siendo  $\rho^2$  el coeficiente de determinación,  $\sigma^2$  la varianza del conjunto de valores del consumo energético (KW/h) y  $\sigma_r^2$  la varianza residual.

Tras el cálculo automático de estas dos métricas, la selección de las combinaciones pre-definitivas de “técnica-configuración-conjunto de entradas” se hará seleccionando - por cada

conjunto de entradas - la combinación “técnica-configuración” con el menor RMSE y la combinación con el mayor R Square.

5 De entre las dos combinaciones finalistas, la combinación definitiva y óptima será la que tenga el mínimo valor obtenido del coeficiente de variación del RMSE (CVRMSE, E.3), calculado - en todos los casos – a partir del conjunto de datos de validación (*test*). El CVRMSE es una medida que representa el porcentaje de error de forma absoluta en la predicción.

(E.3) 
$$CVRMSE = \frac{RMSE}{\hat{y}}$$

10 Siendo  $\hat{y}$  el valor medio de la muestra del consumo energético (KW/h).

El procedimiento de la invención sigue por tanto un enfoque de “modelo libre” basado en la selección automática - para un edificio específico - de la óptima combinación “técnica-configuración-conjunto de entradas”, es decir, aquella combinación que proporciona la  
15 estimación más precisa (de acuerdo al conjunto de datos de entrenamiento) y estable (de acuerdo al conjunto de datos de validación).

Para implementar este enfoque de modelo libre, se puede utilizar por ejemplo el paquete CARET de la herramienta software R. Este paquete permite implementar el entrenamiento  
20 automático de algoritmos predictivos, buscando la configuración óptima de los mismos a partir de un conjunto de valores asignados a sus parámetros de ajuste.

### **Descripción de los dibujos o figuras**

25 Para complementar la descripción y con el objetivo de ayudar a una mejor comprensión de las características de la invención, se acompañan como parte integrante de esta memoria dos diagramas con carácter ilustrativo y no limitativo:

En la figura 1 se ha representado cada uno de los pasos del procedimiento propuesto de acuerdo a una realización de la invención.

30 La figura 2 representa de forma visual el uso de la metodología propuesta en la invención.

### **Ejemplo de realización de la invención**

35 Se parte de la información sobre la evolución horaria de los siguientes parámetros: (i) la predicción meteorológica en la zona geográfica del edificio; (ii) el consumo energético del sistema de climatización; y, (iii) el nivel de ocupación en el edificio, durante un periodo de

tiempo. De esta forma tendremos información que represente la variabilidad de circunstancias que ocurren durante un periodo de tiempo en la climatización de un edificio: frío, calor, calefacción, refrigeración, edificio ocupado, vacío, etc.

5 Los pasos que componen el método propuesto en la presente invención son los que aparecen en la FIG.1, y que describimos a continuación:

1. Limpiar datos. Se limpia el conjunto de datos asociados a los parámetros de consumo, predicción meteorológica y nivel de ocupación. Para esto filtraremos el conjunto de datos  
10 para quedarnos sólo con aquellos que tienen sentido dada la naturaleza del problema: eliminamos combinaciones con valores negativos y nulos para el consumo energético, y eliminamos combinaciones con valores de predicciones meteorológicas y niveles de ocupación que no sean realistas.

15 2. Transformar datos en características - desde un punto de vista estadístico - y generar conjuntos de entrada. Se generan los conjuntos de características candidatas a ser entradas del modelo predictivo en el instante  $t$ . El conjunto de características base partirá de dos entradas como mínimo: (i) la predicción meteorológica en el instante  $t$  y el nivel de ocupación en el instante  $t$ . Los siguientes conjuntos de características estarán compuestos  
20 por históricos de la predicción meteorológica, tales como: (ii) la predicción meteorológica en el instante  $t-1$ , la predicción meteorológica en el instante  $t$  y el nivel de ocupación en el instante  $t$ ; (iii) la predicción meteorológica en el instante  $t-2$ , la predicción meteorológica en el instante  $t-1$ , la predicción meteorológica en el instante  $t$  y el nivel de ocupación en el instante  $t$ ; y así sucesivamente hasta el histórico  $t-x$ , siendo  $x$  el número de horas a partir del  
25 cual deja de haber correlación entre el histórico de los datos meteorológicos (en el instante  $t-x$ ) y el consumo energético de climatización (en el instante  $t$ ).

3. Generar los conjuntos de valores a evaluar para configurar cada uno de los parámetros de configuración de cada una de las técnicas bajo estudio. Cuanta mayor granularidad en  
30 los valores, más ajuste se conseguirá en la configuración, pero a costa de un mayor tiempo de computación. Los parámetros a configurar para cada una de las técnicas evaluadas y los conjuntos de valores a probar son:

a. MLP: Red neuronal perceptrón multicapa

35

Un perceptrón es una neurona (o nodo) que aplica una función a la combinación lineal de los datos de entrada y sus pesos y produce un resultado predictivo.

5 Si la predicción es incorrecta se procede a recalcular los pesos de forma que el error se reduzca. En el caso de la invención, se usa el algoritmo que añade una capa intermedia al paradigma: capa oculta, cuya salida será la entrada de la capa final y los errores y pesos se van re-calculando de atrás hacia delante.

- i. Nombre de la función en R: *mlp*
- 10 ii. Parámetro de ajuste *Size*, que hace referencia al número de neuronas en la capa oculta de la red neuronal.
- iii. Valores de configuración durante el entrenamiento: 5, 10, 20, 30, 40, 50

b. SVM: Máquina de vectores soporte

15 Sea  $x_n$  el conjunto de entrada e  $y_n$  el vector de salida, SVM busca la función  $f(x) = x'\beta + b$  que sea lo más plana posible, esto es  $\min \frac{1}{2} \|\beta\|^2$  sujeto a que la diferencia entre la función y el valor a predecir sea mínima, o sea,  $|y_n - f(x_n)| < \epsilon$ . Cuando dicha función no exista, se añaden variables positivas auxiliares  $\alpha$  de forma que el objetivo sea  $\min \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^l (\alpha_i + \alpha_i^p)$  y las restricciones queden como  $y_i - f(x_i) \leq \epsilon + \alpha_i$  y  $f(x_i) - y_i \leq \epsilon + \alpha_i$

- 20 i. Nombre de la función en R: *svmRadialCost*
- ii. Parámetro de ajuste *Cost (C)*, el cual controla el compromiso entre maximizar la generalidad del modelo y minimizar su error de estimación.
- iii. Valores de configuración durante el entrenamiento: 1, 15, 18, 20, 22, 25

c. GAUSS: Función Gauss con núcleo de base radial

25 Se trata de una alternativa al Perceptrón multicapa y es reconocida por su aplicabilidad en problemas de predicción y clasificación. Este método está diseñado con neuronas en la capa oculta, las cuales se activan por medio de funciones radiales de carácter no lineal (Gaussianas) y en cuanto a la capa de salida son funciones lineales.

- 30 i. Nombre de la función en R: *gaussprRadial*
- ii. Parámetro de ajuste *Sigma*, el cual indica el radio de la función de base radial.
- iii. Valores de configuración durante el entrenamiento: 0.1, 0.5, 1, 1.5, 1.8
- d. BRNN: Red neuronal con función bayesiana regularizada

En una red neuronal, debido a la gran cantidad de conexiones, puede ocurrir sobreentrenamiento. Para evitarlo hemos usado este algoritmo que incluye técnicas de regularización que controlan esto.

- i. Nombre de la función en R: *brnn*
- 5 ii. Parámetro de ajuste *neurons*, el cual indica el número de neuronas de la red neuronal.
- iii. Valores de configuración durante el entrenamiento: 5, 10, 20, 30, 40, 50
- e. RF: Ramdon Forest

Este algoritmo crea árboles que hacen predicciones. La media de todas ellas se toma como valor final. Los árboles se construyen buscando para cada nodo la partición de datos con menor suma de cuadrados de los residuos.

Nombre de la función en R: *rf*

- i. Parámetro de ajuste *mtry*, el cual nos indica el número de variables que se van a escoger aleatoriamente para cada separación de un nodo.
- 15 ii. Valores de configuración durante el entrenamiento: 2, 3, 4, 5, 6, 7

15

Normalizar conjuntos de entrada. La normalización de un conjunto de datos significa ajustar los valores medidos en diferentes escalas respecto a una escala común, con el objetivo de restringir la gama de valores en un conjunto de datos dado y haciéndolas así comparables entre ellas. En este caso, se aísla la salida del modelo (el consumo) y se normalizan los diferentes conjuntos de entradas compuestos por valores de la predicción meteorológica y el nivel de ocupación. La normalización la haremos en el intervalo [0, 1], esto es, la variable  $x_i$  será transformada a  $z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$ , quedando todos los valores comprendidos entre los valores 0 y el 1. Este paso es importante para evitar problemas durante el aprendizaje, como por ejemplo la saturación de las neuronas.

25

4. Dividir el conjunto de muestras de entrada y salida en subconjuntos de entrenamiento y subconjuntos de validación. Para esto se toma (i) un subconjunto del 75% del conjunto total de muestras como datos para entrenamiento, y (ii) el 25% restante como subconjunto de validación -(proporciones generalmente aplicadas en la literatura). El objetivo de calcular estos subconjuntos es el de utilizar el subconjunto de entrenamiento para generar el modelo predictivo, y utilizar el subconjunto de validación para verificar y comprobar luego el comportamiento del modelo generado.

30

5. Aplicar a los conjuntos de entrada un análisis de componentes principales (ACP) al 95% de la varianza. Para ello, se aísla de nuevo la salida de los conjuntos de entrenamiento

35

y validación, y se busca la proyección según la cual los datos queden mejor representados en términos de mínimos cuadrados. Esto crea las componentes principales, que son combinaciones lineales de las variables de entrada, o características, que no están correlacionadas entre sí. A continuación, se seleccionan las características que juntas supongan el 95 % de la varianza del conjunto inicial, reduciendo su dimensionalidad.

6. Entrenar las técnicas con diferentes conjuntos de entradas y diferentes valores de configuración (detallados para cada algoritmo en la Etapa 3). Asociar de nuevo la salida a los nuevos conjuntos de entradas, y entrenamos cada uno de los algoritmos de regresión seleccionados utilizando: (i) cada uno de los conjuntos de entradas; (ii) los diferentes valores asignados a sus parámetros de configuración; (iii) la métrica RMSE; y, (iv) el método de validación cruzada con 10 iteraciones.

7. Preseleccionar los modelos en función del RMSE y el R Square. Para cada uno de los modelos obtenidos en el paso anterior, se seleccionan para cada conjunto de entradas, las dos mejores soluciones atendiendo al menor valor del RMSE y el mayor R Square.

8. Selección final del modelo óptimo en función del CVRMSE. Con cada uno de los modelos obtenidos en el paso anterior, calculamos el valor del CVRMSE utilizando para ello el conjunto de validación y, finalmente, escogemos el modelo con el menor CVRMSE asociado.

Este modelo finalmente seleccionado es el modelo óptimo capaz de estimar con la mayor precisión la evolución horaria del consumo energético del sistema de climatización ambiental, dada información horaria sobre la predicción meteorológica y el nivel de ocupación del edificio.

En la FIG.2 se muestra un ejemplo de generación, uso y evaluación del modelo predictivo óptimo generado siguiendo el procedimiento propuesto. En la parte izquierda del esquema se representa la fase de generación y selección - entre diferentes alternativas - del modelo predictivo óptimo, y en la parte derecha se muestra un ejemplo de aplicación de dicho modelo para la predicción del consumo energético de climatización y su evaluación final.

## REIVINDICACIONES

1. Procedimiento para obtener de forma automática el modelo óptimo predictivo del consumo energético horario asociado al sistema de climatización ambiental en edificios residenciales o de oficinas, utilizando como parámetros de entrada conocidos: (i) la predicción meteorológica en su zona geográfica; (ii) el consumo energético del sistema de climatización; y, (iii) el nivel de ocupación en el edificio; caracterizado porque comprende los siguientes pasos:
  - a. limpiar datos asociados a los parámetros de consumo, predicción meteorológica y nivel de ocupación;
  - b. transformar datos en características representativas y generar conjuntos de entrada;
  - c. generar los conjuntos de valores a evaluar para configurar cada uno de los parámetros de configuración de cada una de las técnicas bajo estudio (MLP, BRNN, SVM, GAUSS y RF);
  - d. normalizar conjuntos de entrada en el intervalo [0, 1];
  - e. dividir conjuntos de entrada y salida en subconjuntos de entrenamiento y subconjuntos de validación;
  - f. aplicar a los conjuntos de entradas componentes principales al 95% de la varianza;
  - g. entrenar las técnicas MLP, BRNN, SVM, GAUSS y RF con diferentes conjuntos de entradas y diferentes valores de configuración;
  - h. preseleccionar los dos mejores modelos, con sus respectivos mejores conjuntos de entrada y valores para sus configuraciones, en función del menor RMSE y el mayor R Square;
  - i. selección final del modelo óptimo en función del menor CVRMSE.

25

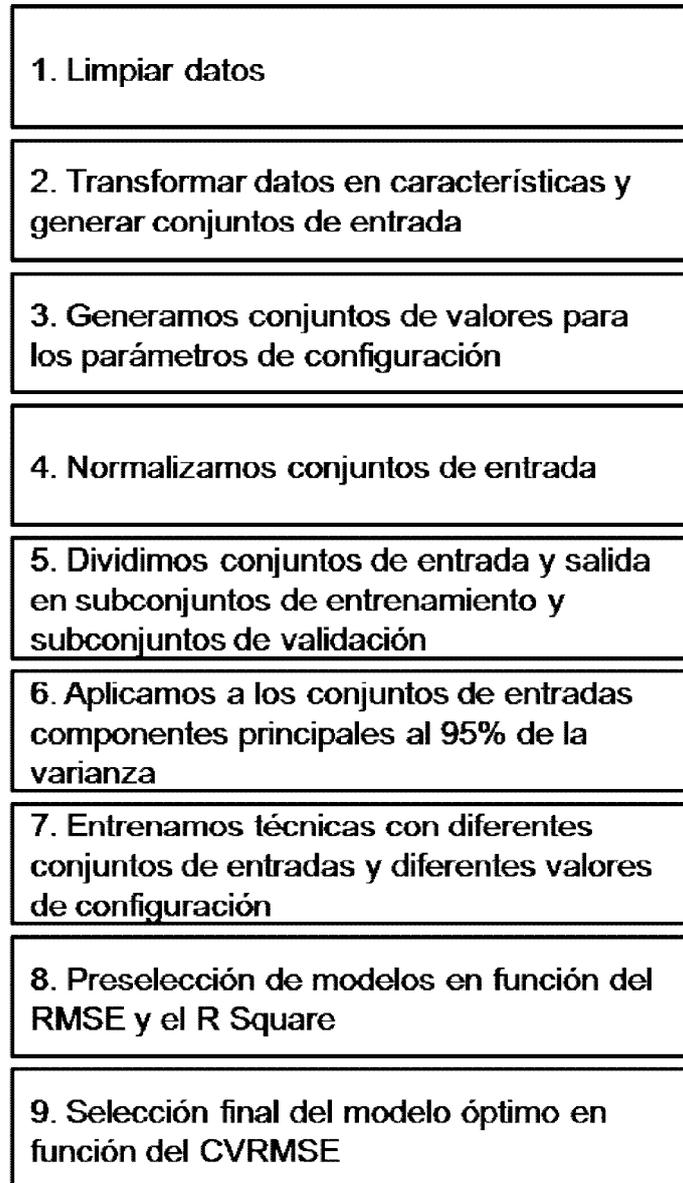


FIG. 1

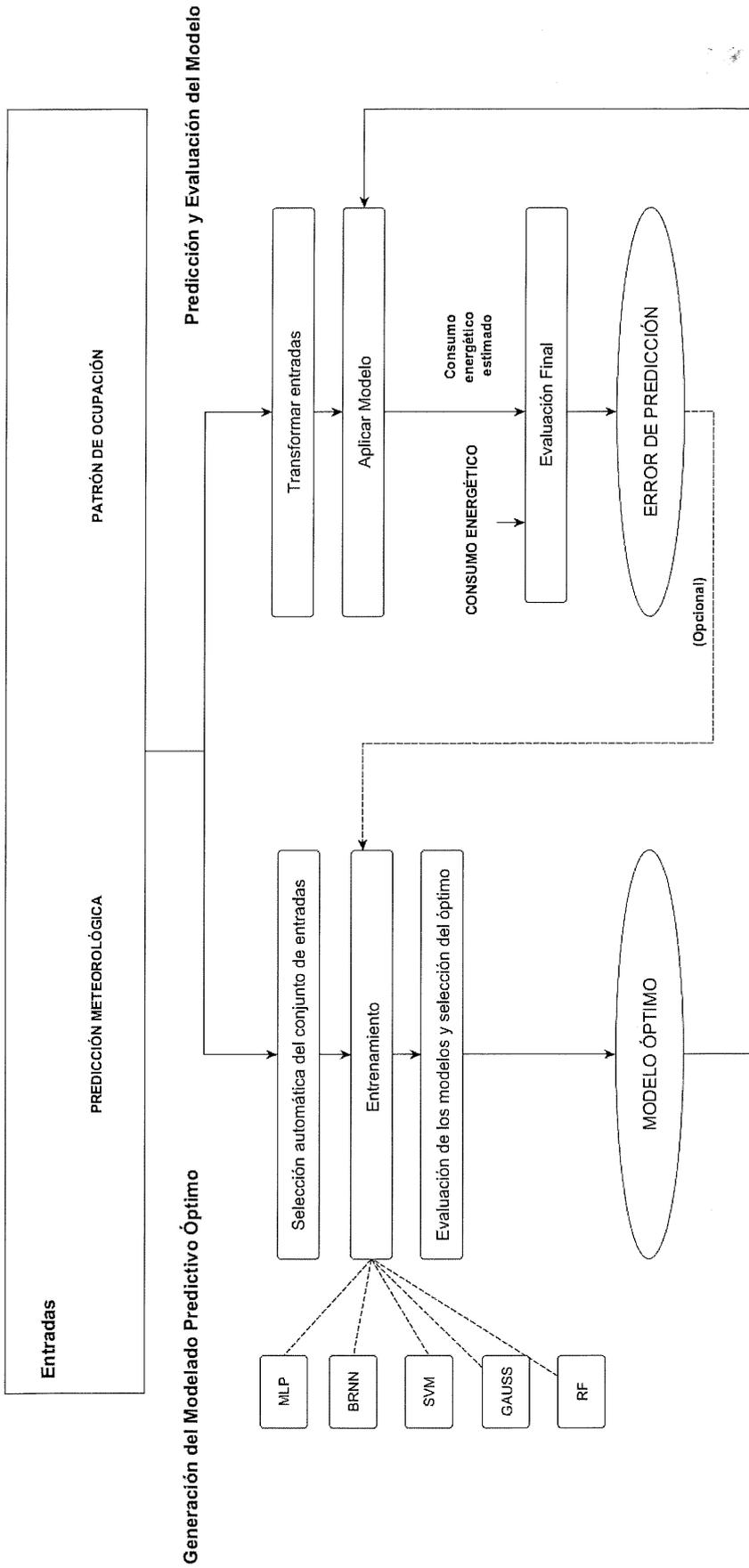


FIG. 2



- ②① N.º solicitud: 201731099  
②② Fecha de presentación de la solicitud: 08.09.2017  
③② Fecha de prioridad: **07-03-2017**

INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤① Int. Cl.: **G06F17/00** (2006.01)  
**G06Q50/06** (2012.01)

DOCUMENTOS RELEVANTES

Categoría	⑤⑥ Documentos citados	Reivindicaciones afectadas
X	GONZÁLEZ-VIDAL, A., MORENO-CANO, V., TERROSO-SÁENZ, A., SKARMETA, A. F. Towards Energy Efficiency Smart Buildings Models Based on Intelligent Data Analytics. Procedia Computer Sciences, 12/05/2016, Vol. 83, Páginas 994-999 Recuperado de Internet <URL: <a href="https://www.sciencedirect.com/science/article/pii/S1877050916302460">https://www.sciencedirect.com/science/article/pii/S1877050916302460</a> >, <DOI: <a href="https://doi.org/10.1016/j.procs.2016.04.213">https://doi.org/10.1016/j.procs.2016.04.213</a> >. Abstract, Secciones 2-4, Fig. 1	1
A	HEDÉN, WILLIAM. Predicting Hourly Residential Energy Consumption using Random Forest and Support Vector Regression: An Analysis of the Impact of Household Clustering on the Performance Accuracy. KTH School of Engineering Sciences, 2016	1
A	US 2016018835 A1 (GAASCH WILLIAM HUGH et al.) 21/01/2016	1
A	US 2012240072 A1 (ALTAMURA FRANK et al.) 20/09/2012	1
A	WO 2014075108 A2 (UNIV COLUMBIA) 15/05/2014	1

Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

**El presente informe ha sido realizado**

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe  
16.04.2018

Examinador  
F. Díaz Madrigal

Página  
1/2

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

G06F, G06Q

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC