

MÉTODOS DE INTELIGENCIA ARTIFICIAL UTILIZADOS EN EL CONTROL TÉRMICO DE EDIFICIOS

Ponce Chaparro, Jesús Omar; Márquez Gutiérrez, Pedro Rafael
Tecnológico Nacional de México – Tecnológico de Chihuahua
División de Estudios de Posgrado
Avenida Tecnológico 2909, Chihuahua, Chih.
Tel. 614-201-200
omar_ponce37@hotmail.com; pmarquez@itch.edu.mx

RESUMEN.

Los métodos clásicos para controlar los sistemas de calefacción, ventilación y aire acondicionado (HVAC) a menudo se utilizan en un rendimiento subóptimo, causando hasta el 40% del consumo total de energía en un edificio. El rendimiento de estos sistemas puede ser optimizado integrando esquemas de control inteligente. La conducción inteligente de los sistemas HVAC tiene el potencial de reducir significativamente el costo de energía. Dado el caso de que el uso de estrategias tradicionales como las basadas en reglas o las basadas en modelos pueden ser un tanto ineficientes en la práctica, se espera que el uso de métodos de aprendizaje de refuerzo tenga altos impactos y contribuciones cuando se orienta a reducir los costos de energía mediante el desarrollo de tecnología de inteligencia artificial para la mejora del rendimiento de HVAC.

Este estudio está orientado en la evaluación de 5 casos de estudio de sistemas de control inteligente de clima y su comparación referente a un proyecto que actualmente se está trabajando en el ITCH, por lo tanto, los resultados mostrados son los correspondientes a los obtenidos en cada caso de estudio

Palabras Clave: HVAC, inteligencia artificial, aprendizaje automático, aprendizaje de refuerzo de aprendizaje profundo, Proceso de decisión de Markov, gestión de energía, control térmico.

ABSTRACT.

Classical methods to control Heating, Ventilation and Air-Conditioned systems (HVAC) are often used in suboptimal performance, causing up to 40% of the total energy consumption in a building. Performance of this kind of systems can be improved by integration intelligent control schemes. Intelligent driving of HVAC systems has the potential to reduce significantly the energy cost. Given the case that traditional control strategies, such as rule-based or model-based, are often inefficient, the use of reinforcement learning methods is expected to have high impacts and contributions when oriented to reduce energy costs by the development of artificial intelligence technology for HVAC performance improvement.

This analysis has the intention to evaluate and compare 5 intelligent climate control systems study cases against a project that is being investigated currently at ITCH, so results showed in this document are related to the behavior and performance of such analysis.

Keywords: HVAC, artificial intelligence, machine learning, deep learning reinforcement learning Markov decision process, energy management, thermal control.

1. INTRODUCCIÓN

El consumo de energía en los espacios de vida y trabajo aumenta cada vez más por el uso ineficiente de los equipos de control climático, y a veces representa más del 40% del consumo total de energía en un edificio, por lo que la posibilidad de reducir el consumo de energía es controlar automáticamente el equipo de HVAC, lo cual significa implementar sistemas inteligentes de control. Los sistemas de inteligencia artificial tratan de imitar el comportamiento humano en las máquinas a través de la codificación matemática de la lógica humana y el análisis de datos.

En este análisis se evaluaron 5 casos de estudio diferentes para implementar sistemas inteligentes de control del clima. Cada método es único, pero todos están orientados a aumentar el rendimiento de los sistemas de control de temperatura. En el ITCH se está desarrollando un sistema inteligente de control de temperatura basado en el paradigma de Aprendizaje Reforzado, cuyo esquema se muestra en la figura 1. La pretensión es contar con un sistema adaptativo que aprenda las preferencias de los usuarios con base en su empleo continuo a lo largo del tiempo. Un sistema de aprendizaje reforzado es un sistema conductual que premia las acciones correctas y penaliza las incorrectas.

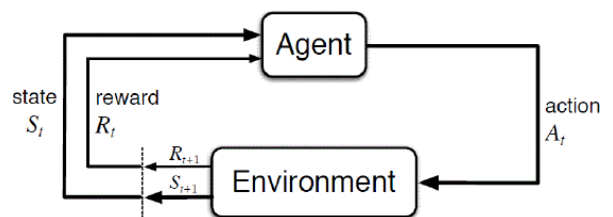


Fig. 1 – Modelo de Aprendizaje Reforzado

2. DISCUSION

A continuación, se presentan diversos trabajos relacionados con el tema de investigación bajo desarrollo y se explica su relevancia, funcionamiento, resultados y comparativos.

2.1. APRENDIZAJE PROFUNDO DE REFUERZO PARA EL CONTROL ÓPTIMO DE CALENTAMIENTO DE ESPACIOS.

Las ecuaciones, imágenes y resultados de esta sección fueran obtenidos del documento “*Deep reinforcement learning for optimal control of space heating*” [1] enumerado en la lista de referencias.

A) Relación

- Utiliza algoritmos libres de modelo convencional
- Inclinación de refuerzo basada en el proceso de decisión de Markov
- Orientado a maximizar la eficiencia

B) Funcionamiento

El proceso de decisión de Markov se implementa bajo las siguientes consideraciones:

- Espacio de estado continuo, donde:
 $s^t \in S$ 1
- Espacio de acción discreto, donde:
 $a^t \in A$ 2
- Función de transición, donde:
 $T(s^t, a^t, s^{t+1})$ 3
- Función de recompensa, donde:
 $R^t(s^t, a^t, s^{t+1})$ 4

Según las definiciones anteriores, la curva característica del sistema es la siguiente:

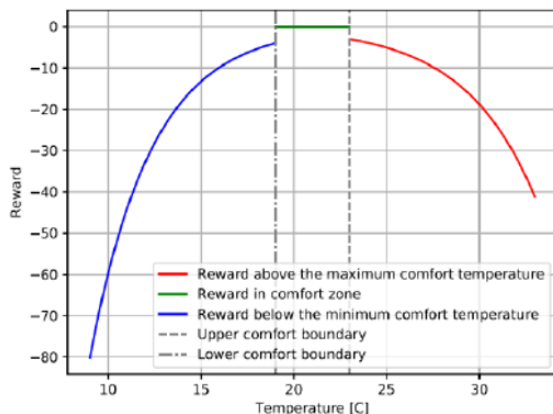


Fig. 2 Función de recompensa por pérdida de confort térmico de ocupación

Se espera que el agente de aprendizaje reforzado equilibre dos corrientes de recompensas. El primero representa la comodidad del ocupante, mientras que el segundo representa el consumo de energía o el dinero gastado al cumplir con el primer objetivo.

En el algoritmo libre de modelo convencional (utilizados en métodos de control clásicos) se utilizan las siguientes aproximaciones:

$$(s, a, w) \approx Q * (s, a) = Q^\pi(s, a) \quad 5$$

El objetivo de esta red neuronal es la minimización del error cuadrático medio con respecto a un valor objetivo, que genera un vector de 6 dimensiones que refleja la "bondad" de las posibles acciones de control dado un par de estado-acción dado.

C) Resultados

La red neuronal muestra un rendimiento aceptable y ha demostrado la dinámica aprendida del sistema casi a la perfección, según se muestra en la figura 3.

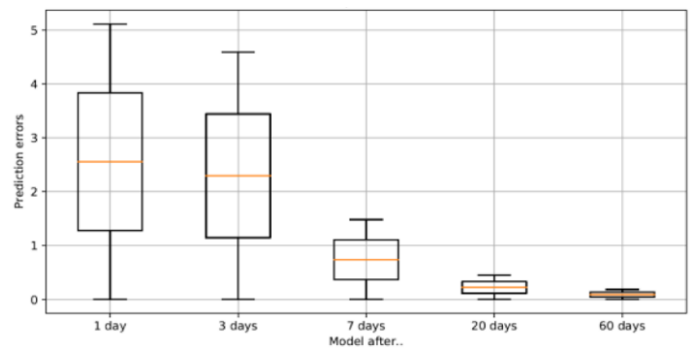


Fig. 3: rendimiento del modelo a lo largo del tiempo

Este algoritmo ofrece tiempo de cómputo más rápido y mayor confiabilidad cuando se producen cambios inesperados en la dinámica ambiental.

D) Comparación con el control convencional.

El control convencional proporciona el rendimiento más bajo. Su funcionamiento se basa en el siguiente modelo:

$$a_t = \begin{cases} P_{max} & \text{si } T_i^t < T_{min} - \Delta T \\ 0 & \text{de otra forma} \end{cases} \quad \text{eq.6}$$

El modelo anterior establece que la bomba de calor comienza a consumir toda la potencia cuando el sistema comienza a funcionar (requiere calor).

Algorithm	Consumption change (%)	Cost change (%)	Loss of comfort (EUR)
Rule based	0	0	0
Model-free RL	-6.4	-6.4	5.23

Tabla 1. Resultados experimentales

Si se produce una alteración en la temperatura ambiente como resultado de un cambio en la ocupación de la habitación, según la regla (modelo convencional), solo el agente de aprendizaje reforzado puede adaptar su comportamiento a la nueva condición, pero esta acción requerirá muchos datos de entrenamiento.

2.2 CONTROL DE CONFORT TÉRMICO ENERGÉTICAMENTE EFICIENTE EN EDIFICIOS INTELIGENTES A TRAVÉS DEL APRENDIZAJE PROFUNDO DE REFUERZO.

Las ecuaciones, imágenes y resultados de esta sección fueran obtenidos del documento “Energy efficient thermal comfort control in smart buildings via deep reinforcement learning” [2] enumerado en la lista de referencias.

A) Relación

- Orientado a reducir el consumo de energía.
- Predice el punto de confort.
- Utiliza redes neuronales y aprendizaje reforzado como método de control.

B) Funcionamiento

El proceso de decisión de Markov se utiliza para formular la optimización energética y el control del confort térmico en el edificio con los siguientes detalles:

- Estados del sistema representados de la siguiente manera:

$$S_t = (T_t^{in}, H_t^{in}, T_t^{out}, H_t^{out}) \quad 7$$

- Acciones es el punto de ajuste de temperatura y humedad, representado como:

$$A_t = (T_t^{set}, H_t^{set}) \quad 8$$

La recompensa se define de la siguiente manera:

$$R_t(S_t, A_t) = \begin{cases} 0 & -D < M_t < D \\ (M_t - D) & M_t > D \\ (-D - M_t) & M_t < -D \end{cases} \quad 9$$

El siguiente modelo ANN se utiliza para predecir la comodidad:

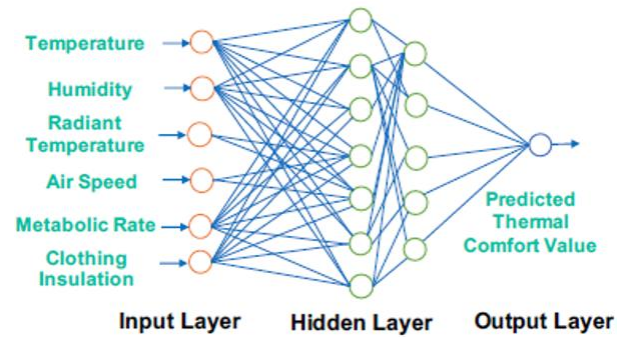


Fig. 4 - Estructura de la red neuronal profunda para predecir el control de confort.

Se utiliza un método de regularización Bayesiano para minimizar el error al entrenar los pesos de la red neuronal, representado como:

$$Cost\ function = \alpha_1 \sum_{i=1}^N (Y_i - Y'_i)^2 + \alpha_2 \sum_{j=1}^m \omega_j^2 \quad 10$$

La acción de control especificada por la red de actores se utiliza para seleccionar acciones durante el entrenamiento. Después del entrenamiento, solo se requiere la red de actores para realizar acciones de control, por lo que los puntos de ajuste del sistema HVAC se establecen como las acciones de control especificadas por la red de actores.

C) Resultados

El error de entrenamiento puede converger después de varias épocas, lo que significa que el error se minimiza:

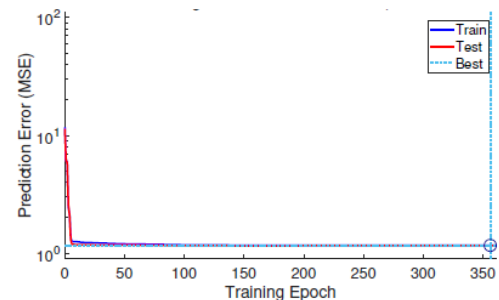


Fig. 5 - Rendimiento de entrenamiento

Los resultados del análisis muestran que el algoritmo puede lograr altos rendimientos para la predicción de confort térmico.

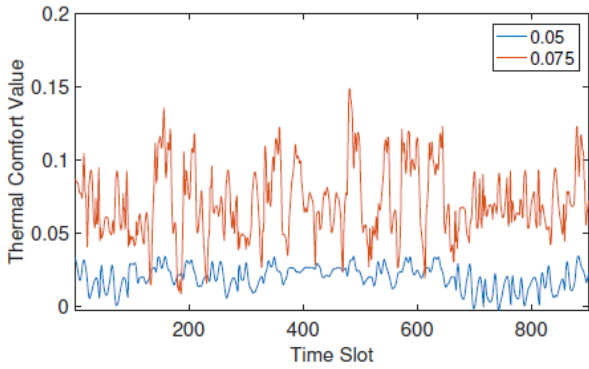


Fig. 6 - Valores de confort térmico bajo diferentes pesos de costo de energía a lo largo del tiempo

D) Comparación con el control convencional.

Este método puede lograr un mayor desempeño, comodidad y eficiencia energética en comparación con los métodos clásicos de control. Esto debido a que está actualizando su parámetro en cada iteración, mientras que los métodos convencionales están fijos durante todo el tiempo.

2.3 APRENDIZAJE PROFUNDO DE REFUERZO PARA CONSTRUIR CONTROL DE HVAC.

Las ecuaciones, imágenes y resultados de esta sección fueran obtenidos del documento “Deep reinforcement learning for building HVAC control” [3] enumerado en la lista de referencias.

A) Relación

- Reduce el consumo de energía mediante la programación automatizada.
- Aprende estrategias efectivas para la operación de HVAC
- Usa enfoque basado en datos para el aprendizaje por refuerzo.

B) Funcionamiento

La operación de control de HVAC está formulada como proceso de decisión de Markov de la siguiente manera:

- Los estados del sistema se determinan con base en la observación de un patrón meteorológico, considerando una secuencia corta de pronóstico del tiempo.
- Las acciones de control se definen de la siguiente manera:

$$A(A^1, A^2, \dots, A^n) \quad 11$$

El objetivo de la función de recompensa es minimizar el costo total de energía mientras se mantiene la temperatura de cada zona dentro de un rango deseado, por lo que la función de recompensa se calcula de la siguiente manera:

$$r_t = -cost(a_{t-1}, s_{t-1}) - \lambda \sum_{i=1}^z ([T_t^i - \bar{T}_t^i] + [T_t^i - T_t^i]) \quad 12$$

Las recompensas negativas se utilizan para que el algoritmo de aprendizaje de refuerzo profundo maximice las recompensas totales.

Para el marco DRL, se definió una aproximación de red neuronal:

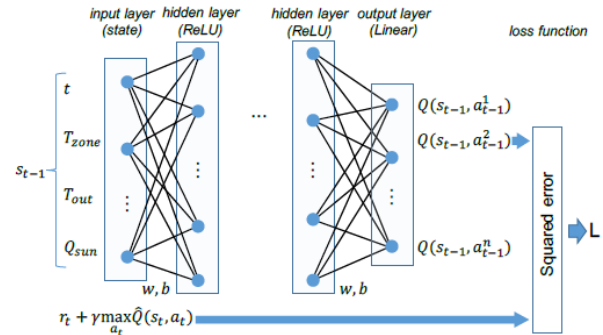


Fig. 7 - Estructura propuesta de la red neuronal

La salida de la red neuronal anterior se define mediante la siguiente función:

$$L = \frac{1}{2n} \sum_{i=1}^n [Q * (s_t, a_t^i) - Q(s_t, a_t^i)]^2 \quad 13$$

Con esta topología, la eficiencia puede mejorarse enormemente al seleccionar acciones con la política E-codiciosa.

RELU se usa como la función de activación para capas ocultas y la capa lineal es utilizada para inferir el valor de acción en la salida.

C) Resultados

La red neuronal necesita ser entrenada para obtener el mejor rendimiento:



Fig. 8 - Efectividad del algoritmo DRL

Se puede observar que el algoritmo es efectivo para mantener la temperatura de la zona en un rango deseado, y por lo tanto la temperatura de confort.

D) Comparación con los métodos convencionales.

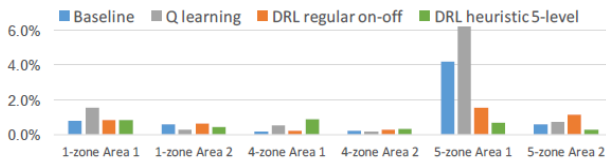


Fig. 9 - Comparación de la tasa de violación de temperatura entre Algoritmo DRL, enfoque de línea de base y aprendizaje Q

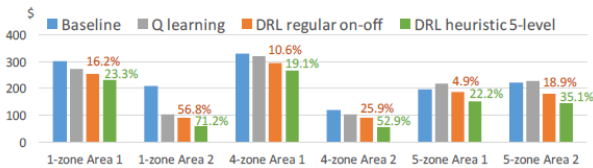


Fig. 10 - Comparación del costo de energía entre DRL algoritmos, enfoque de línea base y aprendizaje de Q

2.4 Un enfoque de aprendizaje reforzado profundo para usar el modelo de energía de todo el edificio para un control óptimo de HVAC.

Las ecuaciones, imágenes y resultados de esta sección fueran obtenidos del documento “A deep reinforcement learning approach to using whole building energy model for HVAC optimal control” [4] enumerado en la lista de referencias.

A) Relación

- Orientado a reducir el consumo de energía.
- Aprendizaje de refuerzo aplicado a sistemas multi-agente.
- Control óptimo para maximizar la eficiencia.

B) Funcionamiento

El trabajo propuesto consiste en las siguientes etapas:

- El modelado BEM es necesario para comprender las características de variación de todo el entorno.
- Calibración del modelo. Una vez que se ha adquirido el modelo BEM, debe calibrarse y clasificarse de acuerdo con el comportamiento ambiental. Esto se realiza en función de la calibración Bayesiana con la formulación:

$$y(x) = \zeta(x) + \epsilon(x) = \eta(x, c') + \epsilon(x) \quad 14$$

- Se requiere capacitación de agentes DRL para una política de control óptima. El siguiente modelo se utiliza para este propósito:

$$state = \{\{SiteOb, BldOb\}_t, \dots, \{SiteOb, BldOb\}_{t-n}\} \quad 15$$

La ecuación de recompensa está definida por siguiente relación:

$$reward = - \begin{cases} [\tau * ((PPD - 0.1) * \rho)^2 + \beta * E_{HVAC}]_0^1 |_{occu=1} \\ [\tau * ((PPD - 0.1) * \rho)^2 + \beta * E_{HVAC}]_0^1 |_{occu=0} \end{cases} \quad 16$$

El agente DRL capacitado controlará el entorno a través del sistema HVAC. Durante esta tarea, el agente refina continuamente su política de control al aprender del comportamiento real del sistema HVAC.

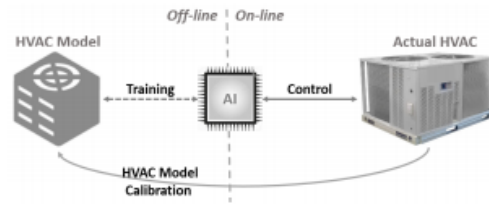


Fig.11 - Marco de control BEM-DRL

C) Resultados

El desempeño de DRL ha sido analizado por medio de su rendimiento de optimización y rendimiento de control. El rendimiento de optimización se evalúa en función de la calidad de la convergencia del entrenamiento DRL y el rendimiento de control se evalúa por medio del consumo de energía de calefacción y el confort térmico.

Tabla 2 - Control de rendimiento

#	Hyperparameters		Training Model			Testing Model		
	Action Repeat ^a	$\tau, \beta, \rho^{\dagger}$	Heating Energy (kWh)	PPD _{mean} (%)	PPD _{std} (%)	Heating Energy (kWh)	PPD _{mean} (%)	PPD _{std} (%)
Basecase	N/A	N/A	45302	10.48	4.48	43709	9.46	5.59
1	1	1.0, 1.5, 20	52806	8.72	4.31	47522	8.23	2.46
2	1	1.0, 2.5, 20	44549	11.58	5.35	39484	11.11	4.53
3	1	1.0, 2.5, 10	40101	16.09	9.46	37238	14.2	8.65
4	3	1.0, 1.5, 20	42255	11.46	4.26	38550	10.63	3.34
5	3	1.0, 2.5, 20	43532	10.63	4.23	39109	10.44	3.75
6	3	1.0, 2.5, 10	42104	11.49	4.24	37131	11.71	3.76

D) Comparación con los métodos convencionales.

Este algoritmo tiene capacidades de aprendizaje basado en recompensas del sistema. Este tipo de comportamiento permite que el sistema corrija su funcionalidad si su rendimiento está fuera de alcance. Este tipo de comportamiento no está incluido en los sistemas de control convencional, por lo que tendrán el mismo comportamiento sin importar los cambios en el entorno.

2.5 Control de ventilación y calefacción de calefacción asistida por inteligencia artificial y la demanda insatisfecha de sensores

Las ecuaciones, imágenes y resultados de esta sección fueran obtenidos del documento “Intelligent model based fault detection and diagnosis for HVAC system using statistical

machine learning methods” [5] enumerado en la lista de referencias.

A) Relación

- Inteligencia Artificial orientada al control del sistema HVAC
- Minimización del consumo de energía.
- Algoritmo basado en el modelo de Markov.

B) Funcionamiento

En el proceso de detección y diagnóstico de fallas se reciben datos de los componentes del HVAC. Estos datos se utilizan como información estadística para implementar un modelo oculto de Markov. Este tipo de modelos necesitan capacitación para aprender patrones de condiciones normales y defectuosas. El modelo que define esta acción es el siguiente:

- Vector que contiene probabilidad previa de estados ocultos

$$\pi_i = p\{q_0 = i\} \quad 17$$

- Conjunto de probabilidades de transición de estado

$$a_{ij} = p\{q_{t+1} = j | q_t = i\} \quad 18$$

- Probabilidad de observación

$$b_j(k) = p\{O_t = V_k | q_t = j\} \quad 19$$

La siguiente imagen muestra la interacción entre diferentes sensores así como estados de observación y estados ocultos del modelo:

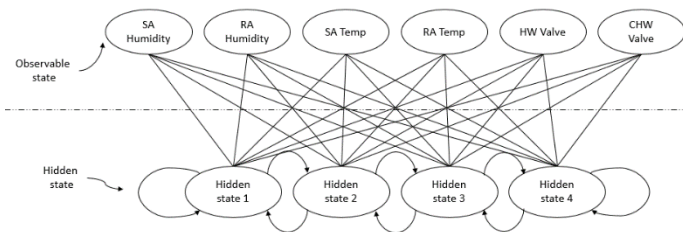


Fig.12 Modelo HMM

C) Resultados

El sistema se dejó en funcionamiento durante una semana con dispositivos con una falla conocida y el algoritmo puede detectar las fallas.

Fault description	Date
EA Damper Stuck (Fully Open)	8/20/2007
EA Damper Stuck (Fully Close)	8/21/2007
Return Fan at fixed speed (30%spd)	8/22/2007
Return Fan complete failure	8/23/2007
Cooling Coil Valve Control unstable (Reduce PID PB by half)	8/24/2007

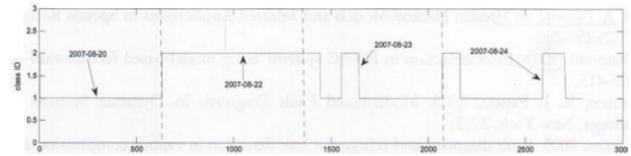


Fig.13 Algoritmo FDD

D) Comparación con los métodos convencionales.

Los métodos convencionales no pueden predecir fallas. Pueden detectar si ya ocurrió una falla colocando un sensor para una condición específica, pero no pueden pronosticar una condición de falla, ese es el propósito del algoritmo diseñado.

2.6 CONCLUSIONES.

El control clásico es comúnmente sencillo de implementar y suele encontrarse comúnmente en los sistemas HVAC, pero no proveen características de aprendizaje y auto ajuste de sus parámetros de operación.

En este estudio hemos investigado el método de control de aprendizaje por reforzamiento. Los esfuerzos realizados en la construcción de sistemas de control ambiental son cruciales para hacer que los edificios sean más inteligentes en términos de reducción de costos al mejorar la eficiencia en sistemas como el aire acondicionado. Como método basado en datos, RL sin modelo convencional puede funcionar en un paradigma de simulación y ofrece novedades en aplicaciones de inteligencia artificial, obteniendo resultados satisfactorios al evaluar el desempeño del sistema en terminaos de utilización de energía, tiempos de respuesta y confort

3 REFERENCIAS.

[1] Nagy, Adam, “Deep reinforcement learning for optimal control of space heating”, September 2018.
 [2] Gao, Guanyu, “Energy efficient thermal comfort control in smart buildings via deep reinforcement learning”, January 2019.
 [3] Tianshu, Wei, “Deep reinforcement learning for building HVAC control”, June 2017.
 [4] Zhang, Zhiang, “A deep reinforcement learning approach to using whole building energy model for HVAC optimal control”, September 2018.
 [5] Guo, Ying, “Intelligent model based fault detection and diagnosis for HVAC system using statistical machine learning methods”, 20