



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería en Ciencias y Sistemas

**OPTIMIZACIÓN DE ENERGÍA AL CONTROLAR VENTILACIÓN Y AIRE ACONDICIONADO
POR MEDIO DE UN ALGORITMO DE APRENDIZAJE POR REFUERZO PROFUNDO,
IMPLEMENTADO EN EL DEPARTAMENTO DE INGENIERÍA MECÁNICA, UNIVERSIDAD
NACIONAL CHIAO TUNG, HSINCHU, TAIWÁN**

Jorge Mario Gutiérrez Ovando
William Fernando Valladares Muñoz
Asesorado por el Ing. Marlon Antonio Pérez Türk

Guatemala, septiembre de 2019

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**OPTIMIZACIÓN DE ENERGÍA AL CONTROLAR VENTILACIÓN Y AIRE ACONDICIONADO
POR MEDIO DE UN ALGORITMO DE APRENDIZAJE POR REFUERZO PROFUNDO,
IMPLEMENTADO EN EL DEPARTAMENTO DE INGENIERÍA MECÁNICA, UNIVERSIDAD
NACIONAL CHIAO TUNG, HSINCHU, TAIWÁN**

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

JORGE MARIO GUTIÉRREZ OVANDO

WILLIAM FERNANDO VALLADARES MUÑOZ

ASESORADOS POR EL ING. MARLON ANTONIO PÉREZ TÜRK

AL CONFERÍRSELES EL TÍTULO DE

INGENIEROS EN CIENCIAS Y SISTEMAS

GUATEMALA, SEPTIEMBRE DE 2019

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANA	Inga. Aurelia Anabela Córdova Estrada
VOCAL I	Ing. José Francisco Gómez Rivera
VOCAL II	Ing. Mario Renato Escobedo Martínez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Br. Luis Diego Aguilar Ralón
VOCAL V	Br. Christian Daniel Estrada Santizo
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANO	Ing. Pedro Antonio Aguilar Polanco
EXAMINADOR	Ing. José Alfredo González Díaz
EXAMINADOR	Ing. Edgar Estuardo Santos Sutuj
EXAMINADOR	Ing. Pedro Pablo Hernández Ramírez
SECRETARIA	Inga. Lesbia Magalí Herrera López

Jorge Mario Gutiérrez Ovando

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANA	Inga. Aurelia Anabela Córdova Estrada
VOCAL I	Ing. José Francisco Gómez Rivera
VOCAL II	Ing. Mario Renato Escobedo Martínez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Br. Luis Diego Aguilar Ralón
VOCAL V	Br. Christian Daniel Estrada Santizo
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANO	Ing. Pedro Antonio Aguilar Polanco
EXAMINADOR	Ing. Marlon Francisco Orellana López
EXAMINADOR	Ing. César Augusto Fernández Cáceres
EXAMINADOR	Ing. Herman Igor Véliz Linares
SECRETARIA	Inga. Lesbia Magalí Herrera López

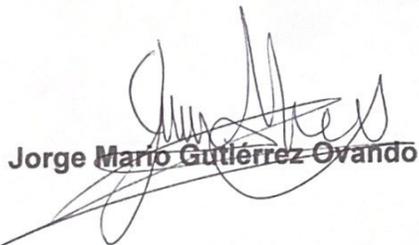
William Fernando Valladares Muñoz

HONORABLE TRIBUNAL EXAMINADOR

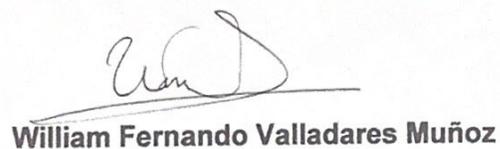
En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presentamos a su consideración nuestro trabajo de graduación titulado:

**OPTIMIZACIÓN DE ENERGÍA AL CONTROLAR VENTILACIÓN Y AIRE ACONDICIONADO
POR MEDIO DE UN ALGORITMO DE APRENDIZAJE POR REFUERZO PROFUNDO,
IMPLEMENTADO EN EL DEPARTAMENTO DE INGENIERÍA MECÁNICA, UNIVERSIDAD
NACIONAL CHIAO TUNG, HSINCHU, TAIWÁN**

Tema que nos fue asignado por la Dirección de la Escuela de Ingeniería en Ciencias y Sistemas, con fecha abril 2018.



Jorge Mario Gutiérrez Ovando



William Fernando Valladares Muñoz

UNIVERSIDAD DE SAN CARLOS
DE GUATEMALA



FACULTAD DE INGENIERIA
ESCUELA DE CIENCIAS Y SISTEMAS
TEL. 24439500 EXT. 1534

Guatemala, 25/abril/2019

Ingeniero Carlos Azurdia
Coordinador Trabajos de Tesis
Escuela de Ingeniería en Ciencias y Sistemas
Facultad de Ingeniería, USAC
Presente

Estimado Ing. Azurdia:

Por este medio informo que he revisado y aprobado el Trabajo de Tesis titulado: "OPTIMIZACIÓN DE ENERGÍA AL CONTROLAR VENTILACIÓN Y AIRE ACONDICIONADO POR MEDIO DE UN ALGORITMO DE APRENDIZAJE POR REFUERZO PROFUNDO, IMPLEMENTADO EN EL DEPARTAMENTO DE INGENIERÍA MÉCANICA, UNIVERSIDAD NACIONAL CHIAO TUNG, HSINCHU, TAIWÁN", de los estudiantes Jorge Mario Gutiérrez Ovando y William Fernando Valladares Muñoz.

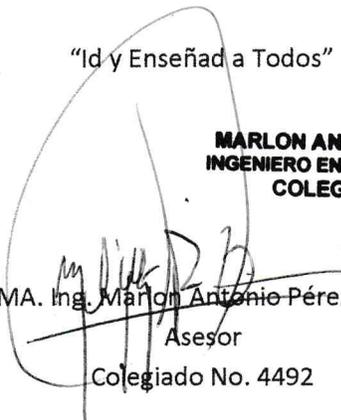
Con base en la evaluación realizada hago constar que he evaluado la calidad, validez, pertinencia y coherencia de los resultados obtenidos en el trabajo presentado, por lo cual el trabajo evaluado cuenta con mi aprobación.

Agradeciendo su atención y deseándole éxitos en sus actividades profesionales me suscribo.

Atentamente,

"Id y Enseñad a Todos"

MARLON ANTONIO PEREZ TURK
INGENIERO EN CIENCIAS Y SISTEMAS
COLEGIADO No. 4492


MA. Ing. Marlon Antonio Pérez Türk
Asesor
Colegiado No. 4492



Universidad San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería en Ciencias y Sistemas

Guatemala, 16 de mayo del 2019

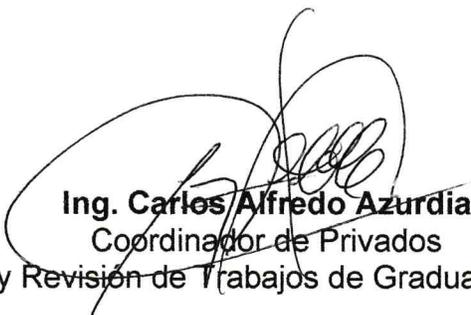
Ingeniero
Marlon Antonio Pérez Türk
Director de la Escuela de Ingeniería
En Ciencias y Sistemas

Respetable Ingeniero Pérez:

Por este medio hago de su conocimiento que he revisado el trabajo de graduación de los estudiantes **JORGE MARIO GUTIÉRREZ OVANDO** con carné **201213124** y **CUI 2340 46309 0114**, y **WILLIAM FERNANDO VALLADARES MUÑOZ** con carné **201314510** y **CUI 2446 39647 0101**, titulado: **OPTIMIZACIÓN DE ENERGÍA AL CONTROLAR VENTILACIÓN Y AIRE ACONDICIONADO POR MEDIO DE UN ALGORITMO DE APRENDIZAJE POR REFUERZO PROFUNDO, IMPLEMENTADO EN EL DEPARTAMENTO DE INGENIERÍA MÉCANICA, UNIVERSIDAD NACIONAL CHIAO TUNG, HSINCHU, TAIWÁN**, y a mi criterio el mismo cumple con los objetivos propuestos para su desarrollo, según el protocolo.

Al agradecer su atención a la presente, aprovecho la oportunidad para suscribirme,

Atentamente,


Ing. Carlos Alfredo Azurdia
Coordinador de Privados
y Revisión de Trabajos de Graduación



UNIVERSIDAD DE SAN CARLOS
DE GUATEMALA



FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA EN
CIENCIAS Y SISTEMAS
TEL: 24767644

*El Director de la Escuela de Ingeniería en Ciencias y Sistemas de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del asesor con el visto bueno del revisor y del Licenciado en Letras, del trabajo de graduación **“OPTIMIZACIÓN DE ENERGÍA AL CONTROLAR VENTILACIÓN Y AIRE ACONDICIONADO POR MEDIO DE UN ALGORITMO DE APRENDIZAJE POR REFUERZO PROFUNDO, IMPLEMENTADO EN EL DEPARTAMENTO DE INGENIERÍA MECÁNICA, UNIVERSIDAD NACIONAL CHIAO TUNG, HSINCHU, TAIWÁN”**, realizado por los estudiantes, JORGE MARIO GUTIÉRREZ OVANDO y WILLIAM FERNANDO VALLADARES MUÑOZ aprueba el presente trabajo y solicita la autorización del mismo.*

“ID Y ENSEÑAD A TODOS”

Ing. Carlos Gustavo Alonzo

Director

Escuela de Ingeniería en Ciencias y Sistemas

Guatemala, 03 de septiembre de 2019



DTG. 329.2019

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería en Ciencias y Sistemas, al Trabajo de Graduación titulado: **OPTIMIZACIÓN DE ENERGÍA AL CONTROLAR VENTILACIÓN Y AIRE ACONDICIONADO POR MEDIO DE UN ALGORITMO DE APRENDIZAJE POR REFUERZO PROFUNDO, IMPLEMENTADO EN EL DEPARTAMENTO DE INGENIERÍA MECÁNICA, UNIVERSIDAD NACIONAL CHIAO TUNG, HSINCHU, TAIWÁN,** presentado por los estudiantes universitarios: **Jorge Mario Gutiérrez Ovando y William Fernando Valladares Muñoz,** y después de haber culminado las revisiones previas bajo la responsabilidad de las instancias correspondientes, autoriza la impresión del mismo.

IMPRÍMASE:

Inga. Anabela Cordova Estrada
Decana

Guatemala, septiembre de 2019

/gdech



ACTO QUE DEDICO A:

- Dios** Por regalarme la vida y permitir que el cauce natural de las cosas desembocara en este momento.
- Mis padres** Roberto Gutiérrez y Miriam Ovando, gracias por haber creído en mí, brindándome su amor y apoyo a lo largo de toda mi vida.
- Mi hermano** Pablo Gutiérrez, mi único hermano, gracias por sus consejos y ser un ejemplo en mi vida.
- Mi demás familia** Por siempre brindarnos su ayuda, aun en los momentos más difíciles. Gracias de todo corazón.
- Mis amigos** Por todos los momentos de alegría, tristeza y risa que hemos compartido. Gracias, sin duda no hubiera sido igual sin ustedes.

Jorge Mario Gutiérrez Ovando

ACTO QUE DEDICO A:

Dios	Por haber permitido que todo sucediera.
Mis padres	Armando Dagoberto Valladares Morales y Vilma Leticia Muñoz Valenzuela, por haberme apoyado y amado siempre.
Mis hermanos	Edward Vinicio y Keneth Armando, por el cariño y apoyo constante que me dan.
Mi demás familia	Por ser una importante influencia en mi vida.
Mis amigos	Por haber tenido la fortuna de conocerlos, y de compartir muchas experiencias a lo largo de la carrera.

William Fernando Valladares Muñoz

AGRADECIMIENTOS A:

Universidad de San Carlos de Guatemala	Por abrirme las puertas y darme los recursos necesarios para poder crecer como profesional y poder aportar al país.
Facultad de Ingeniería	Por el conocimiento que adquirí dentro de sus instalaciones.
Departamento de Matemática	Por brindarme la oportunidad de trabajar y de ser parte de la formación de los futuros ingenieros.
Prof. Chi-Chuan Wang	Por permitir la investigación y elaboración de este trabajo de graduación dentro de sus laboratorios.
Mi asesor	Por tomarse el tiempo de leer y colaborar con el desarrollo de este trabajo, gracias ingeniero.
Mis amigos de Taiwán	Gracias por haber hecho inolvidable la experiencia de elaborar esta investigación en Taiwán.

Jorge Mario Gutiérrez Ovando

AGRADECIMIENTOS A:

Universidad de San Carlos de Guatemala	Por acogerme y brindarme el privilegio de acceder a estudios superiores.
Facultad de Ingeniería	Por el conocimiento que adquirí y las experiencias que viví dentro de sus instalaciones.
Departamento de Matemática	Por brindarme mi primera experiencia laboral y la oportunidad de ser parte de la formación de futuros ingenieros.
Prof. Chi-Chuan Wang	Por su cálido recibimiento, consentimiento y apoyo en la investigación y elaboración de este trabajo de graduación.
Mi asesor	Por su constante apoyo y colaboración en el desarrollo de este trabajo.
Mis amigos de Taiwán	Por haber hecho inolvidable la experiencia de elaborar esta investigación en Taiwán.

William Fernando Valladares Muñoz

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	V
LISTA DE SÍMBOLOS	IX
GLOSARIO	XI
RESUMEN	XV
OBJETIVOS.....	XVII
Hipótesis.....	XVIII
INTRODUCCIÓN	XIX
1. IDENTIFICACIÓN DEL PROBLEMA.....	1
1.1. Climatización de ambientes.....	1
1.2. Sistema de ventilación.....	2
1.3. Sistema de aire acondicionado.....	2
1.4. Confort.....	3
1.4.1. Confort térmico	4
1.4.2. Calidad del aire interior	7
2. PLANTEAMIENTO DE LA SOLUCIÓN	11
2.1. Agente inteligente	11
2.2. Aprendizaje profundo.....	12
2.2.1. Neurona.....	13
2.2.2. Red neuronal	14
2.3. Aprendizaje por refuerzo	16
2.3.1. Aprendizaje-Q.....	17
2.4. Aprendizaje por refuerzo profundo	18
2.4.1. Aprendizaje-Q profundo.....	18

2.4.2.	Aprendizaje-Q doble.....	18
3.	ELEMENTOS DE LA SOLUCIÓN.....	21
3.1.	Simulación.....	24
3.1.1.	SketchUp.....	24
3.1.2.	OpenStudio	25
3.1.3.	EnergyPlus	26
3.2.	Co-simulación.....	27
3.2.1.	BCVTB	28
3.3.	Agente DRL.....	29
3.3.1.	TensorFlow.....	29
3.3.2.	TensorBoard.....	29
3.4.	Internet de las cosas	30
3.4.1.	Capa de aplicación.....	31
3.4.2.	Capa de red.....	32
3.4.3.	Capa de percepción	32
4.	SISTEMA DRL PARA CONTROL DE VENTILACIÓN Y AIRE ACONDICIONADO	37
4.1.	Estado del sistema.....	37
4.2.	Acciones de control.....	39
4.3.	Función de recompensa.....	40
4.3.1.	Penalización de índice PMV	42
4.3.2.	Penalización del nivel de CO ₂	43
4.3.3.	Penalización del sistema de AC	44
4.3.4.	Penalización del sistema de ventilación	48
4.4.	Arquitectura de la implementación	49
4.4.1.	Proceso de pre-entrenamiento	51
4.4.2.	Proceso de control	53

4.4.3.	Proceso de aprendizaje	56
4.5.	Hiperparámetros	58
5.	ENTRENAMIENTO DEL SISTEMA.....	61
5.1.	Desarrollo de la simulación.....	61
5.1.1.	Creación del modelo.....	61
5.1.2.	Adaptación del modelo	63
5.1.3.	Modelación de los sistemas.....	64
5.1.4.	Comunicación	64
5.2.	Entrenamiento	64
5.3.	Pruebas y validación.....	67
6.	RESULTADOS EXPERIMENTALES Y DISCUSIÓN.....	71
	CONCLUSIONES	79
	RECOMENDACIONES	81
	BIBLIOGRAFÍA.....	83
	APÉNDICES	89

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Diagrama de refrigeración por compresión	3
2.	Gráfica de PPD respecto al PMV	5
3.	Estructura básica de un agente.....	12
4.	Estructura de una célula nerviosa o neurona	13
5.	Modelo matemático de una neurona simple.....	14
6.	Estructura básica de una red neuronal.....	15
7.	Diagrama del proceso de aprendizaje por refuerzo.....	17
8.	Diagrama de despliegue para el proceso de entrenamiento	22
9.	Diagrama de despliegue en los ambientes reales.....	23
10.	Entorno de trabajo de SketchUp	24
11.	Entorno de trabajo de OpenStudio	25
12.	Entorno de ejecución de EnergyPlus	26
13.	Arquitectura de comunicación por co-simulación	27
14.	Entorno de trabajo de BCVTB.....	28
15.	Plano del laboratorio 509	34
16.	Plano del salón 132.....	35
17.	Secuencia de control de estados S y acciones A	39
18.	Penalización del índice PMV	42
19.	Penalización del nivel de CO_2	44
20.	Penalización de la energía del sistema de AC	46
21.	Diagrama de flujo principal del algoritmo	50
22.	Arquitectura del algoritmo	51
23.	Diagrama de flujo del proceso de preentrenamiento	52

24.	Valor del factor codicioso.....	55
25.	Diagrama de flujo del proceso de control.....	56
26.	Diagrama de flujo del proceso de aprendizaje.....	57
27.	Modelo 3D del laboratorio 509.....	62
28.	Modelo 3D del salón 132.....	62
29.	Interfaz de IDFEditor.....	63
30.	Entrenamiento del algoritmo.....	66
31.	Octavo periodo de entrenamiento.....	67
32.	PMV promedio en ambientes simulados.....	68
33.	Nivel de CO ₂ promedio en ambientes simulados.....	69
34.	Valor PMV y temperatura exterior en el laboratorio 509.....	72
35.	Consumo energético y niveles de CO ₂ en el laboratorio 509.....	74
36.	Consumo energético de AC para diferentes casos.....	76

TABLAS

I.	Valores asumidos para el cálculo del PMV.....	7
II.	Ambiente térmico recomendado para un confort general.....	7
III.	Concentraciones de CO ₂ y sus efectos.....	8
IV.	Herramientas empleadas para implementación del algoritmo.....	21
V.	Elementos IoT utilizados.....	30
VI.	Configuración de unidades AC.....	31
VII.	Configuración de ventiladores.....	31
VIII.	Configuración de sensores (temperatura, CO ₂ , humedad).....	33
IX.	Configuración térmica de globo negro.....	33
X.	Configuración del enchufe inteligente.....	34
XI.	Valores que conforman el estado del sistema.....	38
XII.	Valores de los pesos de la función de recompensa.....	41
XIII.	Valores de hiperparámetros.....	59

XIV.	Detalles de configuración de las simulaciones	65
XV.	Casos experimentales en el laboratorio 509 con y sin agente	72
XVI.	Resultados resumidos obtenidos en el laboratorio 509	73
XVII.	Comparación de pesos entre Agente-1 y Agente-2	75
XVIII.	Configuración de los casos de prueba simulados	75

LISTA DE SÍMBOLOS

Símbolo	Significado
A	Acción
L	Corte de recompensa
CO₂	Dióxido de carbono
S	Estado
<i>e</i>	Exponencial
°C	Grados centígrados
Hz	Hertz
PMV	Índice de confort térmico
\int	Integral
J	Joule
Kg	Kilogramo
kWh	Kilowatt-hora
\pm	Más/menos
<i>Max</i>	Máximo
M	Memoria
m	Metro
mm	Milímetro
<i>Min</i>	Mínimo
ppm	Parte por millón
Pa	Pascal
w	Peso de penalización
%	Porcentaje
<i>ϵ_{decay}</i>	Probabilidad de decaimiento de exploración

ε	Probabilidad de exploración
R	Recompensa
\hat{Q}	Red neuronal auxiliar
Q	Red neuronal principal
s	Segundo
Σ	Sumatoria
α	Tasa de aprendizaje
τ	Tasa de copia de pesos entre Q y \hat{Q}
γ	Tasa de descuento
Δt_c	Tiempo de control
W	Watt

GLOSARIO

AC	Aire acondicionado.
Climatización	Proceso de creación controlada de condiciones de temperatura, humedad y calidad del aire en interiores.
Confort	Bienestar o comodidad material.
Confort térmico	Sensación neutra de un individuo respecto a un ambiente térmico determinado.
COP	Coeficiente de rendimiento entre el enfriamiento realizado y la electricidad consumida de unidades AC.
Co-simulación	Proceso de sincronización entre dos o más programas de simulación independientes.
Deep learning	Permite aprender la representación de datos con múltiples niveles de abstracción, mediante modelos computacionales compuestos de múltiples capas de procesamiento.
DQL	Aprendizaje-Q Profundo

DQN	Red-Q Profunda
DRL	Aprendizaje por refuerzo profundo.
Hiperparámetro	Parámetro arbitrario que afecta la efectividad y eficiencia del entrenamiento de un agente inteligente.
Humedad relativa	Cantidad de agua, vapor de agua o cualquier otro líquido que esté presente en un cuerpo de aire.
HVAC	Sistema de calefacción, ventilación y aire acondicionado.
IAQ	Calidad del aire interior.
Inteligencia artificial	Ciencia que estudia la simulación de los procesos de inteligencia humana por parte de las máquinas.
Machine learning	Estudio de los algoritmos y modelos estadísticos que usan las máquinas para mejorar su rendimiento por sí mismas. También se denomina aprendizaje automático.
PMV	Índice del valor medio de los votos emitidos por un grupo de personas respecto a su sensación térmica.
Potencia nominal	Potencia máxima que demanda una máquina o aparato en condiciones de uso normales.

PPD	Porcentaje de personas insatisfechas en unas condiciones térmicas determinadas.
Proceso de Markov	Proceso cuya probabilidad del comportamiento futuro únicamente depende del estado actual del proceso.
Python	Lenguaje de programación multiparadigma y muy utilizado en el ámbito matemático y estadístico.
Red neuronal	Modelo computacional basado en un conjunto de neuronas simples.
RL	Aprendizaje por reforzamiento.
Sinapsis	Mecanismo de comunicación que existe entre dos o más neuronas.
Temperatura radiante	Medición del calor emitido por radiación de los elementos del entorno.

RESUMEN

Este trabajo de graduación propone una implementación de inteligencia artificial para mantener el confort térmico y la calidad del aire dentro de niveles óptimos, mientras se consume la menor cantidad de energía proveniente de sistemas de ventilación y aire acondicionado. El algoritmo propuesto se implementa como un agente inteligente que se entrena con diez años de información generada por simulación, y se destina a un ambiente subtropical en Taiwán con sistemas de aire acondicionado en modalidad de solo enfriamiento. Las simulaciones consisten en un laboratorio de investigación con aproximadamente 2 a 10 ocupantes, así como en un salón de clases de hasta, aproximadamente, 60 ocupantes.

El agente propuesto se comprueba con el décimo conjunto de datos de entrenamiento, seguido por el control y la comparación entre las simulaciones y los ambientes reales de prueba. Se encontró que dicho agente puede controlar y equilibrar las necesidades de confort térmico, la calidad del aire interior y el consumo de energía de sistemas de ventilación y aire acondicionado. En ambos ambientes, un laboratorio y un salón de clases, el agente puede manipular bien el ambiente interior, alcanzando valores PMV satisfactorios que van desde $-0,1$ a $+0,07$ durante todo el tiempo de operación. Para la calidad del aire interior, en términos del nivel de dióxido de carbono, los resultados también son satisfactorios con niveles promedio que caen por debajo de 800 ppm. Además, el consumo de energía al emplear el agente propuesto es aproximadamente de 4 % a 5 % más bajo que otros métodos tradicionales, y todavía puede ofrecer un PMV superior y un nivel 10 % más bajo de concentraciones de dióxido de carbono.

OBJETIVOS

General

Implementar un algoritmo de aprendizaje por refuerzo profundo para el control inteligente de los sistemas de ventilación y refrigeración de un ambiente real en la Universidad Nacional Chiao Tung en Hsinchu, Taiwán.

Específicos

1. Optimizar el consumo de energía eléctrica de los sistemas de ventilación y refrigeración del ambiente.
2. Mantener un confort térmico adecuado para las personas, según los estándares ANSI/ASHRAE 55 e ISO 7730.
3. Reducir la concentración de dióxido de carbono dentro del ambiente, según el límite estándar del gobierno de Taiwán.
4. Comparar el rendimiento de la implementación del algoritmo de aprendizaje profundo con otro tipo de sistema tradicional.
5. Utilizar ambientes simulados para optimizar el proceso de entrenamiento del algoritmo.

Hipótesis

Cada vez que la humanidad desarrolla nuevos conocimientos y ve en ellos un gran potencial de aplicación, uno por uno los elementos de su alrededor van evolucionando conforme se van aplicando. Tal es el caso de la electricidad que, hace más de 100 años se descubrió y, luego, se consolidó en el estilo de vida de la humanidad, hasta llegar a la actualidad, donde es imprescindible en el diario vivir. De la misma forma, se espera que la inteligencia artificial se propague a cada aspecto de la vida cotidiana del ser humano. Siguiendo los pasos de dicha evolución, se pondrá en evidencia la capacidad de nuestra implementación de un algoritmo de inteligencia artificial para mejorar el control de un aspecto tan cotidiano como la ventilación y la climatización de ambientes.

Hipótesis nula

Es viable la utilización de un algoritmo de inteligencia artificial propuesto para controlar la ventilación y el aire acondicionado, ya que logra mantener un confort térmico aceptable, bajos niveles de dióxido de carbono, y un consumo optimizado de energía.

Hipótesis alternativa

La implementación del algoritmo de inteligencia artificial propuesto no es capaz de controlar simultáneamente la ventilación y el aire acondicionado, ya que obtiene valores fuera de los rangos recomendados, o los controla utilizando una cantidad excesiva de energía.

INTRODUCCIÓN

En los países del sur de Asia prevalece el clima tropical y subtropical, que normalmente presentan una humedad muy alta durante todo el año¹. Los países en esta área también son propensos a terremotos de magnitudes variables; por consiguiente, para evitar daños significativos, se deben seguir lineamientos de construcción muy estrictos donde el concreto a menudo se emplea como componente principal para construir casas modernas y edificios de oficinas. No obstante, el concreto no es un buen aislante térmico y durante el verano, el interior de estos edificios puede superar fácilmente los 30 °C sin la presencia de una unidad de aire acondicionado². Por tanto, el aire acondicionado se ha convertido en un producto indispensable en tales regiones.

Considerando a Taiwán como ejemplo, para mantener un buen nivel de confort térmico durante el verano, más del 80 % de las residencias y la mayoría de los edificios comerciales taiwaneses poseen unidades de aire acondicionado o sistemas híbridos con ventilación natural³. Además de las necesidades de refrigeración para dichos ambientes, también se debe considerar el confort del usuario en sistemas típicos de aire acondicionado. Así, por ejemplo, el confort térmico y la calidad del aire interior.

Según la Sociedad Americana de Ingenieros de Calefacción, Refrigeración y Aire Acondicionado (ASHRAE, siglas del nombre en inglés), el confort térmico

¹ KELLY. *Taiwan Weather, Taiwan Climate in Spring, Summer, Autumn and Winter*. <https://www.chinahighlights.com/taiwan/weather.htm>. Consulta: 18 octubre de 2018.

² KUO-LIANG, L. *Energy Consumption Analysis for Concrete Residences—A Baseline Study in Taiwan*. p. 1-13.

³ HWANG, R. *Comfort Temperature and Preferred Temperature in Taiwan*. p. 155-163.

se define como “el estado de condición mental que expresa satisfacción con el ambiente térmico y que se evalúa mediante un criterio subjetivo”⁴. Aunque un reciente estudio concluye que los residentes taiwaneses ven el uso de una unidad de aire acondicionado como último recurso, afirmando que prefieren métodos más baratos para lidiar con el calor, como abrir ventanas, usar ventiladores eléctricos o cambiarse de ropa⁵. Taiwán ha experimentado un alto aumento en la demanda de energía durante los veranos, y se están tomando medidas especiales⁶. Entre ellas, un aumento del 3 % en las facturas eléctricas, la reactivación de un reactor nuclear y la construcción de una central eléctrica de carbón para reemplazar una central retirada en 2016⁷.

Sin embargo, para lograr una mejor situación general para el usuario, también debe considerarse la calidad del aire en el interior del ambiente, donde el nivel de dióxido de carbono o CO₂, se considera un índice importante para medirla. Los altos niveles de CO₂ producen una serie de efectos perjudiciales para las personas que lo inhalan. Varios estudios concluyen que existe una correlación entre los niveles de CO₂ en el ambiente y la salud de las personas, afectando principalmente sus capacidades de concentración, toma de decisiones y desempeño en general^{8,9}; por consiguiente, se requiere resolver esta problemática sujeta al consumo total de energía, confort térmico y ventilación adecuados para las personas.

⁴ ASHRAE. *Standard 55-2010 Thermal environmental conditions for human occupancy*. p. 4.

⁵ HWANG, R. *Occupants' Behavior in Taiwan*. p. 247-251.

⁶ LIAO, S. *The Non-Linear Relationship between Electricity Consumption and Temperature in Taiwan: An Application for STR (Smooth Transition Regression) Model*. p. 587-605.

⁷ FERRY, T. *The Number-one Concern: Electric Power*. p. 16-19.

⁸ MYHRVOLD, A. N. *Indoor environment in schools—pupils health and performance in regard to CO₂ concentrations*. p. 369-371.

⁹ SATISH, U. *Is CO₂ an indoor pollutant? Direct effects of low-to-moderate CO₂ concentrations on human decision-making performance*. p. 1671-1677.

Debido a las complejas interacciones entre los ambientes térmicos y las variables a considerar, es sumamente difícil manipular el ambiente térmico de manera efectiva; pero un posible medio para lograrlo es la Inteligencia Artificial. Es posible referirse a la inteligencia artificial como la nueva electricidad, comparando su impacto cuando fue descubierta hace más un siglo, y cambiando la vida cotidiana de las personas en todos los sentidos¹⁰. Al hacer uso de ella, problemas complejos en aplicaciones del mundo real se pueden resolver de manera eficiente. En 2013, la compañía DeepMind propuso una nueva familia de algoritmos de inteligencia artificial llamada Aprendizaje por Refuerzo Profundo y también el algoritmo de Aprendizaje-Q Profundo, los cuales hacen uso de una Red-Q Profunda que consiste en una red neuronal con muchas capas para definir una estrategia que puede seleccionar acciones óptimas en busca de mejores recompensas. Mediante el uso de estos algoritmos, DeepMind demostró que un agente inteligente ya entrenado puede jugar videojuegos de Atari incluso mejor que un humano¹¹. Estos avances y aplicaciones motivan el presente trabajo de graduación para solventar la problemática antes mencionada, donde se implementa el algoritmo de aprendizaje por refuerzo profundo para optimizar el consumo de energía de un sistema de aire acondicionado en asociación con el confort térmico y la calidad del aire.

En resumen, la principal contribución de este trabajo es la implementación de un controlador basado en el aprendizaje por refuerzo profundo para optimizar el consumo de energía de los sistemas de ventilación y aire acondicionado en asociación con el confort térmico y la calidad del aire interior. Según el conocimiento de los autores, este estudio es el primero en aplicar una

¹⁰ NG, Andrew. *Neural Networks and Deep Learning, Deep Learning Specialization*. <https://www.coursera.org/learn/neural-networks-deep-learning>. Consulta: 22 de noviembre de 2018.

¹¹ MNIH, V. *Playing atari with deep reinforcement learning*. p. 8.

técnica mejorada de DRL para resolver las complejas interacciones entre el confort térmico, el consumo de energía y la calidad del aire interior al realizar experimentos en entornos simulados y reales mediante el uso de una plataforma de Internet de las cosas. Los experimentos se realizaron en dos ambientes diferentes: un laboratorio y un salón de clases, ambos ubicados en el campus de la Universidad Nacional Chiao Tung en Hsinchu, Taiwán. Así mismo, se demuestra que la implementación del algoritmo de inteligencia artificial propuesto es capaz de equilibrar el consumo de energía, el confort térmico y la calidad del aire interior.

Este trabajo de graduación está organizado de la siguiente manera: la primera sección ofrece una descripción de la literatura y el trabajo relacionado al identificar el problema a resolver; la segunda sección explica el trasfondo y planteamiento de la solución propuesta; la tercera sección detalla los elementos de la solución; la cuarta sección presenta la implementación del algoritmo de aprendizaje por refuerzo profundo propuesto; la quinta sección define la metodología de aprendizaje y entrenamiento; la sexta sección discute los resultados experimentales obtenidos y, finalmente, se desarrollan nuestras conclusiones y recomendaciones.

1. IDENTIFICACIÓN DEL PROBLEMA

La optimización de la energía es una problemática que ha estado adquiriendo importancia desde hace décadas debido a que su demanda se ha incrementado exponencialmente. Prueba de ello es que cada vez hay más dispositivos electrónicos, automóviles, residencias, edificios, entre otros. Este incremento ha generado serios problemas con el medioambiente y el cambio climático, en especial porque las principales fuentes de energía eléctrica siguen siendo los recursos no renovables.

En los países con climas cálidos, la ventilación y el aire acondicionado se han convertido en bienes de primera necesidad, por ejemplo, en Taiwán el 80 % de residencias y edificios comerciales poseen por lo menos una de estas unidades¹²; por ende, un uso adecuado de estos sistemas puede generar un ahorro de energía significativo, y para solventarlo, es importante comprender el proceso de climatización de ambientes, el funcionamiento de los sistemas de ventilación y aire acondicionado, y además, considerar el confort del usuario.

1.1. Climatización de ambientes

La climatización consiste en “dar a un espacio cerrado las condiciones de temperatura, humedad del aire y, a veces también, de presión necesarias para la salud o la comodidad de quienes lo ocupan”¹³. Para llevarlo a cabo se hacen uso de sistemas de calefacción, ventilación y aire acondicionado (HVAC, siglas de su nombre en inglés) y, dependiendo de las condiciones climáticas externas,

¹² HWANG, R. *Comfort Temperature and Preferred Temperature in Taiwan*. p. 155-163.

¹³ DE LA LENGUA ESPAÑOLA, R. A. *Diccionario de la lengua española*. p. 35.

se puede prescindir de la calefacción, tal es el caso en Taiwán donde solo se implementan sistema de ventilación y aire acondicionado, debido a las condiciones climáticas tropicales y subtropicales de la región.

1.2. Sistema de ventilación

Un sistema de ventilación típico consiste en una serie de ventiladores, entradas y salidas donde se conducen flujos de aire en espacios cerrados. Los sistemas de ventilación más complejos incluyen conductos de aire para extraer e inyectar flujos mayores y más precisos. Generalmente, están asociados a otros sistemas de calefacción, sistemas de aire acondicionado o a ambos, como en los sistemas HVAC. Así mismo, una fuente natural de ventilación son las ventanas, resultando en una alternativa más económica.

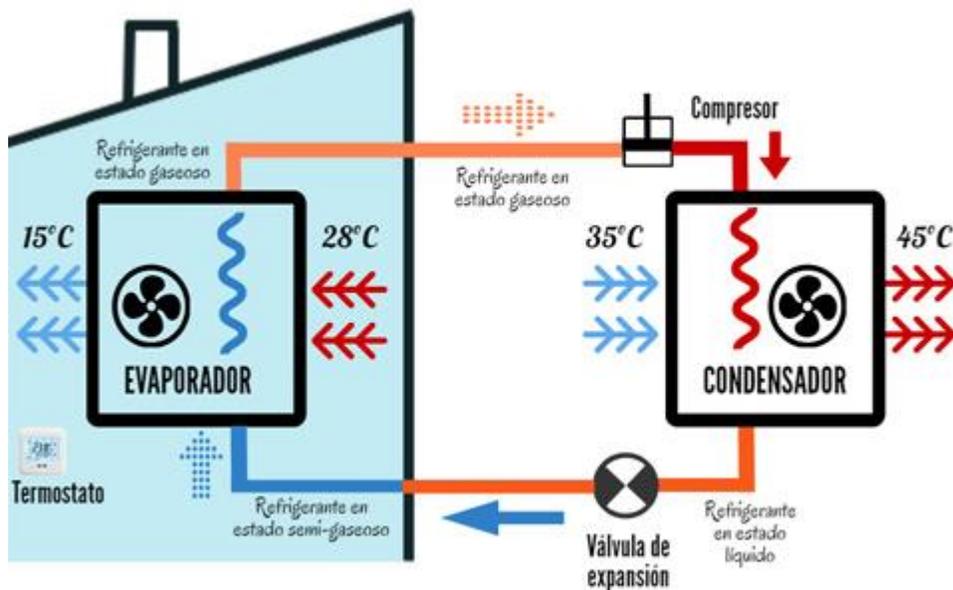
1.3. Sistema de aire acondicionado

Los sistemas de aire acondicionado («*air conditioning*» o AC) basan su funcionamiento en las leyes de la termodinámica, transportando energía en forma de calor de un ambiente a otro. Un sistema de AC típico se compone de las unidades interior (compresor) y exterior (condensador), la válvula de expansión, el líquido refrigerante y el termostato.

El ciclo de refrigeración tradicional es, generalmente, por compresión, el cual consiste en cambiar el estado de un refrigerante, de líquido a gaseoso y viceversa. Inicialmente, se evapora el líquido refrigerante en un serpentín denominado evaporador. Seguidamente, en el compresor se succiona y comprime el refrigerante ya en estado gaseoso, lo cual aumenta su presión y temperatura. Luego, el gas refrigerante se transporta al condensador donde vuelve a su estado líquido, desechando al exterior el aire caliente resultante.

Posteriormente, en la válvula de expansión se libera la presión del refrigerante, convirtiéndolo a un estado semigaseoso y de baja temperatura. Finalmente, utilizando el refrigerante, se suministra aire a menor temperatura al ambiente a ser climatizado¹⁴. Una representación gráfica de este proceso se puede observar en la figura 1.

Figura 1. Diagrama de refrigeración por compresión



Fuente: ARNABAT, I. *¿Cómo funciona el aire acondicionado? Infografía.* <https://www.caloryfrio.com/aire-acondicionado/como funciona el aire acondicionado-infografia.html>. Consulta: 12 de enero de 2019.

1.4. Confort

Los sentidos humanos (el tacto, la vista, el oído, el olfato y el gusto) son los medios por los cuales el mundo es percibido por nosotros y cuando detectan alguna perturbación, nuestro estado de comodidad se altera, lo que afecta la

¹⁴ ARNABAT, I. *¿Cómo funciona el aire acondicionado? Infografía.* <https://www.caloryfrio.com/aire-acondicionado/como funciona el aire acondicionado-infografia.html>. Consulta: 12 de enero de 2019.

concentración, la productividad y el rendimiento. Es posible referirse a la naturaleza del confort como un concepto fenomenológico, es decir, que está sujeto a la conciencia y experiencia de los individuos. Además, se clasifica en cinco tipos: térmico, visual, auditivo, olfativo e higiénico¹⁵. El estándar EN 15251¹⁶ proporciona una guía detallada para el diseño y la evaluación del rendimiento energético en edificios al considerar la calidad del aire interior (confort olfativo e higiénico), el ambiente térmico (confort térmico), la iluminación (confort visual) y la acústica (confort auditivo).

Si bien, el confort visual y auditivo, forman parte importante del confort del usuario, no están incluidos en el alcance de este trabajo de graduación. Esto se debe al aumento de la complejidad para desarrollar escenarios experimentales apropiados y para capacitar adecuadamente el algoritmo de inteligencia artificial propuesto.

1.4.1. Confort térmico

El confort térmico es un estado de satisfacción con el ambiente térmico, es subjetivo y está condicionado por la psicología y la fisiología de las personas, lo que dificulta el definir las condiciones ideales para satisfacer a todos los ocupantes de un ambiente. Los estándares actuales con respecto al confort térmico se basan en el amplio trabajo hecho por el danés Povl Ole Fanger¹⁷, cuyos estudios se documentan en los estándares ANSI/ASHRAE 55¹⁸ e ISO 7730¹⁹.

¹⁵ BODUCH, M., FINCHER, W. *Standards of human comfort: relative and absolute*. p. 1.

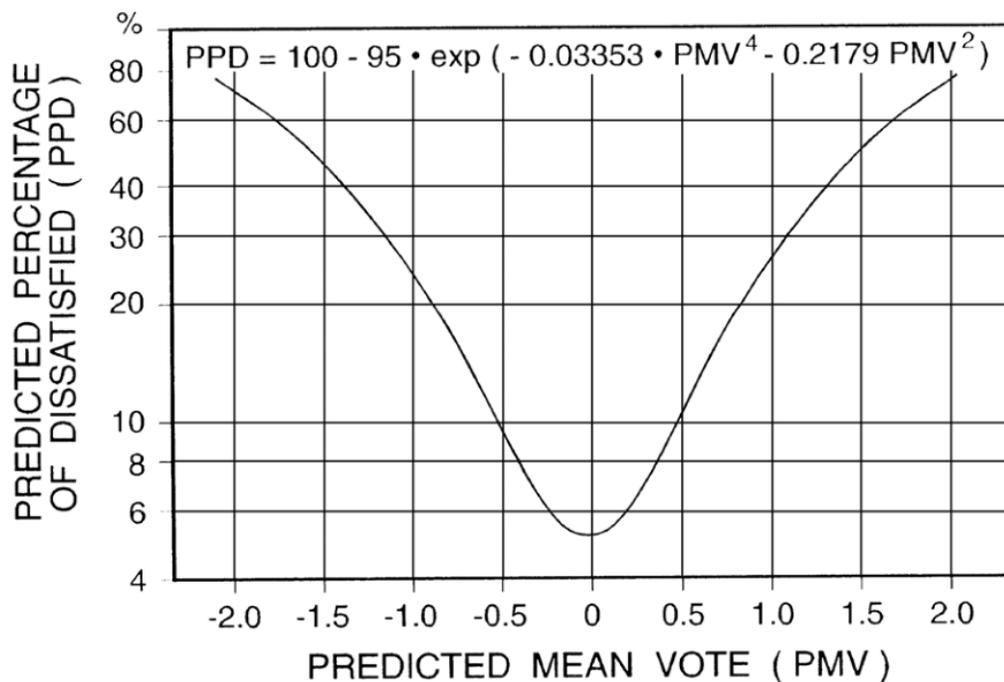
¹⁶ CEN. 15251: 2007 *Indoor environmental input parameters for design assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics*. p. 14-24.

¹⁷ FANGER, P. O. *Thermal comfort. Analysis and applications in environmental engineering*. p. 244.

¹⁸ ASHRAE. 55-2010: *Thermal environmental conditions for human occupancy*. p. 4-13.

Fanger definió el índice de opinión medio estimado («*Predicted Mean Vote*» o PMV) como la predicción cuantitativa para el voto promedio de los individuos en una escala de sensación térmica que varía de -3 a +3; donde -3 es muy frío, 0 es neutral y +3 es muy caliente. Con base en el PMV, también definió el porcentaje previsible de insatisfechos («*Predicted Percentage of Dissatisfied*» o PPD) como la predicción sobre el porcentaje de individuos que no están cómodos con las condiciones térmicas. Esta relación se puede observar en la figura 2.

Figura 2. **Gráfica de PPD respecto al PMV**



Fuente: ASHRAE. 55-2010: *Thermal environmental conditions for human occupancy*. p. 7.

¹⁹ ISO. *Standard 7730:2005 Ergonomics of the thermal environment*. p. 2-4.

En la práctica, la función PMV está determinada por seis parámetros principales:

$$PMV = f(M, I_{cl}, v, t_r, t_{db}, P_s)$$

Donde:

- M es la tasa metabólica (met)
- I_{cl} es el índice de la tela (clo)
- v es la velocidad del aire (m/s)
- t_r es la temperatura radiante media (°C)
- t_{db} es la temperatura de bulbo seco (°C)
- P_s es la presión de vapor de agua (kPa)

En cambio, el PPD solo está sujeto a un parámetro:

$$PPD = 100 - 95 \cdot e^{-(0,03353 \cdot pmv^4 + 0,2179 \cdot pmv^2)}$$

Donde:

- pmv es el índice de opinión medio estimado.

En este trabajo de graduación, para calcular el PMV, debido a las condiciones estables en los ambientes, se asumieron los valores constantes para el índice de la tela, la tasa metabólica y la velocidad del aire, y los otros parámetros se midieron empleando sensores. Consulte la tabla I para obtener más detalles.

Tabla I. **Valores asumidos para el cálculo del PMV**

Parámetro	Valor	Razón
Índice de la tela	0,5 clo	Los ocupantes vestían ropa veraniega durante las mediciones.
Ritmo metabólico	1 met	Los ocupantes se encontraban sentados en los ambientes.
Velocidad del aire	0,1 m/s	Ventanas y puertas estaban cerradas y únicamente las unidades de AC y ventiladores proveían aire a los ambientes.

Fuente: elaboración propia.

Así mismo, en este estudio se siguieron las recomendaciones de los estándares ANSI/ASHRAE 55 e ISO 7730 en mantener un PMV entre $-0,5$ y $+0,5$, con lo cual el PPD es menor al 10 %. Estos detalles se presentan en la tabla II.

Tabla II. **Ambiente térmico recomendado para un confort general**

PPD	Rango PMV
< 10	$- 0,5 < PMV < +0,5$

Fuente: ASHRAE. 55-2010: *Thermal environmental conditions for human occupancy*. p. 9.

1.4.2. Calidad del aire interior

La calidad del aire interior («*Internal Air Quality*» o IAQ) se refiere a la concentración de contaminantes, humedad y temperatura del aire en un ambiente cerrado. Entre estos contaminantes hay algunos como dióxido de carbono CO_2 , olores, polvo y partículas de humo. Aunque el CO_2 no es el único contaminante presente en el aire, cuando alcanza niveles altos de concentración, causa problemas potenciales de salud a las personas que lo inhalan y, por tanto, su concentración promedio en partes por millón (ppm) se

utiliza generalmente para evaluar la IAQ y los niveles de ventilación en los edificios²⁰. Por eso, un sistema de ventilación adecuado puede mantener una IAQ aceptable, lo cual no solo beneficia la salud de las personas, sino también ofrece un ambiente más cómodo y sin olores. En la tabla III se detallan los efectos potenciales en las personas al estar expuestas a determinadas concentraciones de CO₂.

Tabla III. **Concentraciones de CO₂ y sus efectos**

Concentración CO₂ (ppm)	Efectos
250 – 450	Niveles normales en la atmósfera.
450 – 1 000	Niveles típicos en ambientes con ventilación adecuada.
1 000 – 2 000	Somnolencia y falta de aire.
2 000 – 5 000	Dolor de cabeza, pérdida de concentración y atención, aumento de la frecuencia cardiaca y náuseas leves.
5 000 – 40 000	Dificultad para respirar, náuseas, vómito, pérdida de la conciencia.
> 40 000	Pérdida de conciencia repentina, muerte.

Fuente: BONINO, S. *Carbon Dioxide Detection and Indoor Air Quality Control*.

<https://ohsonline.com/articles/2016/04/01/carbon-dioxide-detection-and-indoor-air-quality-control.aspx>. Consulta: 17 de enero de 2019.

El estándar ANSI/ASHRAE 62²¹ recomienda tasas de ventilación mínimas (velocidad del aire) y otras medidas destinadas a proporcionar una IAQ aceptable, y que minimice los efectos adversos para la salud, lo cual ha influenciado el surgimiento de regulaciones internacionales con respecto a los niveles de CO₂ y sus límites. Por ejemplo, en Estados Unidos, la Administración de Salud y Seguridad Ocupacional (OSHA, por sus siglas en inglés) estableció

²⁰ ASHRAE. *Standard 62.1-2010 Ventilation for acceptable indoor air quality*. p. 5.

²¹ Ídem.

un límite de 5 000 ppm durante una jornada laboral de ocho horas promedio ponderado en el tiempo («*Time Weighted Average*» o TWA)²². Australia estableció un límite de 5 000 ppm para una semana de 40 horas de trabajo TWA²³. Taiwán emitió la Ley de Calidad del Aire Interior que establece un límite de 1 000 ppm para todo edificio en general²⁴, y el Reino Unido definió un límite promedio de 1 000 ppm que puede llegar hasta 1 500 ppm, pero no por más de 20 minutos²⁵. En este estudio, se siguió el límite de 1 000 ppm establecido en Taiwán.

²² NIOSH. *Carbon Dioxide*. <https://www.cdc.gov/niosh/pel88/124-38.html>. Consulta: 12 de septiembre de 2018.

²³ AUSTRALIA, W. *Adopted National Exposure Standards for Atmospheric Contaminants in the Occupational Environment*. <https://www.safeworkaustralia.gov.au/doc/adopted-national-exposure-standards-atmospheric-contaminants-occupational-environment-nohsc-1003>. Consulta: 23 de noviembre de 2018.

²⁴ YUAN, E. *Environmental Protection Agency of the Executive Yuan*. <https://oaout.epa.gov.tw/law/LawContent.aspx?id=FL068252>. Consulta: 26 de noviembre de 2018.

²⁵ EDUCATION, D. *Building Bulletin 101 Guidelines on ventilation, thermal comfort and indoor air quality in schools*. <https://www.gov.uk/government/publications/building-bulletin-101-ventilation-for-school-buildings>. Consulta: 28 de noviembre de 2018.

2. PLANTEAMIENTO DE LA SOLUCIÓN

Considerando la salud y comodidad de las personas al hacer uso de sistemas de AC y ventilación, es decir, confort térmico y calidad del aire interior, resulta muy complejo optimizar la energía de manera equilibrada. Por tal razón se recurrió a la Inteligencia Artificial (IA), una rama de las ciencias de la computación que se enfoca en proporcionar mecanismos para resolver problemas de una manera automática y eficiente al combinar la informática, la matemática y la estadística.

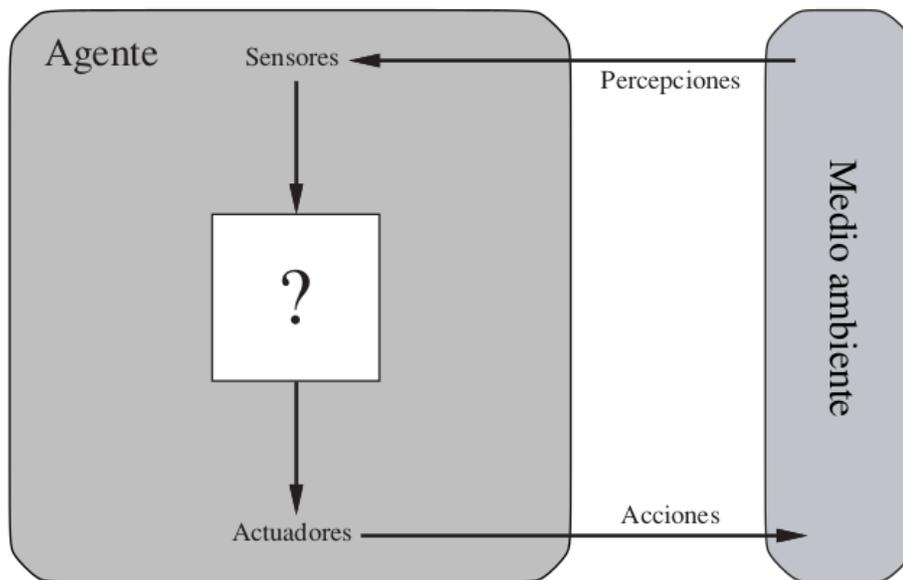
Entre sus múltiples ramas cabe destacar el Aprendizaje Automático o «*Machine Learning*», donde la idea básica del aprendizaje consiste en que la máquina analiza una gran cantidad de información para poder detectar patrones y predecir comportamientos futuros. La solución que se plantea en este trabajo de graduación involucra el desarrollo de un agente inteligente que implementa un algoritmo que resulta de combinar dos subdisciplinas del aprendizaje automático: el aprendizaje profundo y el aprendizaje por refuerzo.

2.1. Agente inteligente

En el ámbito de la IA, un agente inteligente es esencialmente un ente que es capaz de percibir su ambiente y de responder ante ello realizando una acción o serie de acciones. Según Russell y Norvig, un agente es cualquier cosa capaz de percibir su ambiente con la ayuda de sensores y actuar en ese medio utilizando actuadores; es decir, elementos que reaccionan a un estímulo

realizando una acción²⁶. Además, la estructura de un agente permite asociarlo a tareas o procesos más complejos e independientes que recibirán la información de los sensores y decidirán la acción a realizar. La estructura básica de un agente se representa en la figura 3.

Figura 3. Estructura básica de un agente



Fuente: RUSSELL, S., NORVIG, P. *Inteligencia Artificial: Un Enfoque Moderno*. p. 38.

2.2. Aprendizaje profundo

El Aprendizaje Profundo o «*Deep Learning*» es una rama del aprendizaje automático que permite, mediante modelos computacionales compuestos de múltiples capas de procesamiento, aprender la representación de datos con múltiples niveles de abstracción, es decir, dichos modelos llegan a analizar la información a tal punto que pueden realizar predicciones²⁷.

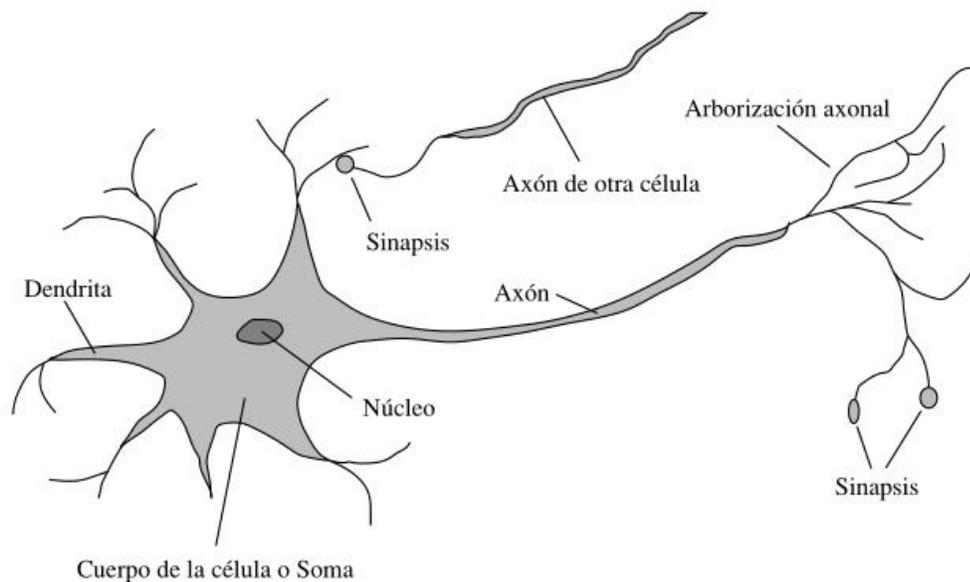
²⁶ RUSSELL, S., NORVIG, P. *Inteligencia Artificial: Un Enfoque Moderno*. p. 37-40.

²⁷ LECUN, Y. *Deep learning*. p. 436-437.

2.2.1. Neurona

Una neurona, en el ámbito de la IA, es una unidad computacional que toma valores de entrada, los procesa y devuelve valores diferentes de salida. Este proceso se basó en el funcionamiento real de las neuronas biológicas, donde ellas toman, por medio de sus dendritas, señales eléctricas provenientes de otras neuronas, procesan estas señales en sus núcleos y luego proceden a enviar señales resultantes por sus axones y demás dendritas hacia otras neuronas, efectuando así una sinapsis.

Figura 4. Estructura de una célula nerviosa o neurona



Fuente: RUSSELL, S. AND NORVIG, P. Inteligencia Artificial: Un Enfoque Moderno. España: Pearson Educación, 2004. p. 13. ISBN 842054003X.

Tal como se observa en la figura 4, la sinapsis es la conexión entre el axón de una neurona y la dendrita de otra cercana. Aunque el funcionamiento más detallado de las neuronas sigue siendo un misterio, mediante la

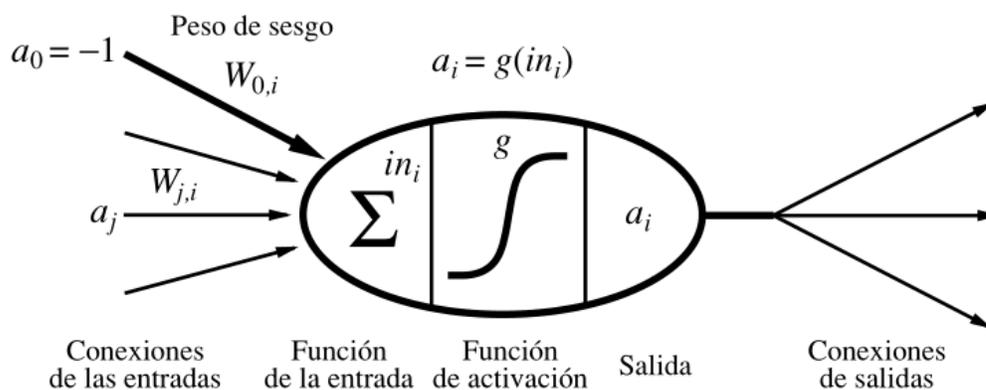
matemática se puede representar este proceso (ver figura 5). Una conexión desde una neurona j a otra neurona i , propaga la activación a_j de j a i como:

$$a_i = g \left(\sum_{j=0}^n W_{j,i} a_j \right)$$

Donde:

- a_i es la activación de la salida de la neurona actual
- g es una función de activación
- $W_{j,i}$ es el peso de la conexión de la unidad j a esta unidad
- a_j es la activación de la salida de la unidad j

Figura 5. **Modelo matemático de una neurona simple**



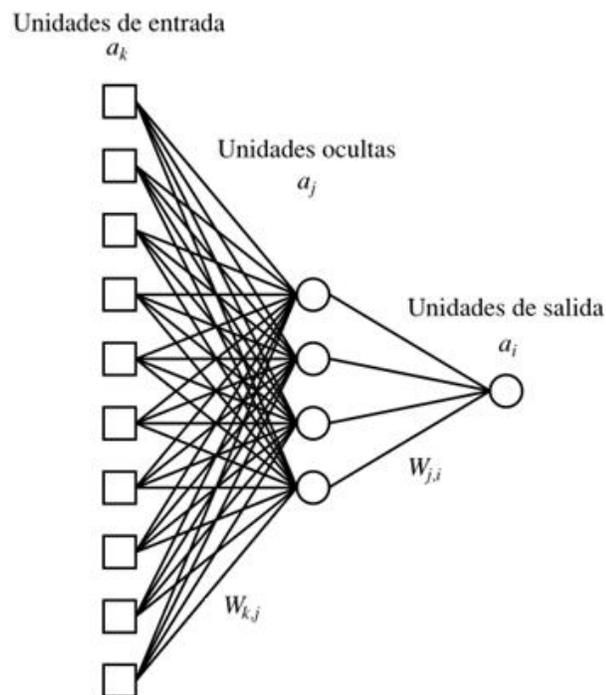
Fuente: RUSSELL, S. AND NORVIG, P. Inteligencia Artificial: Un Enfoque Moderno.
España: Pearson Educación, 2004. p. 839. ISBN 842054003X.

2.2.2. Red neuronal

Una red neuronal está compuesta por nodos o unidades (neuronas) conectadas a través de conexiones dirigidas, las cuales se basan en las redes

entramadas de millones de neuronas en el sistema nervioso de los seres vivos. En su forma más básica, una red neuronal se conforma de una capa con unidades de entrada, una serie de capas con unidades escondidas y una capa con unidades de salida. Si la cantidad de capas y de unidades se incrementa considerablemente, la red neuronal pasa a denominarse red neuronal profunda.

Figura 6. **Estructura básica de una red neuronal**



Fuente: RUSSELL, S., NORVIG, P. *Inteligencia Artificial: Un Enfoque Moderno*. p. 847.

En términos generales, se le puede considerar a una red neuronal como una función que puede aprender, es decir, una vez entrenada la red neuronal, calculará el resultado esperado al darle nuevos valores de entrada. Este entrenamiento se realiza al proveerle una considerable cantidad de datos que ejemplifican el comportamiento deseado para que propague la información en toda su estructura de capas.

2.3. Aprendizaje por refuerzo

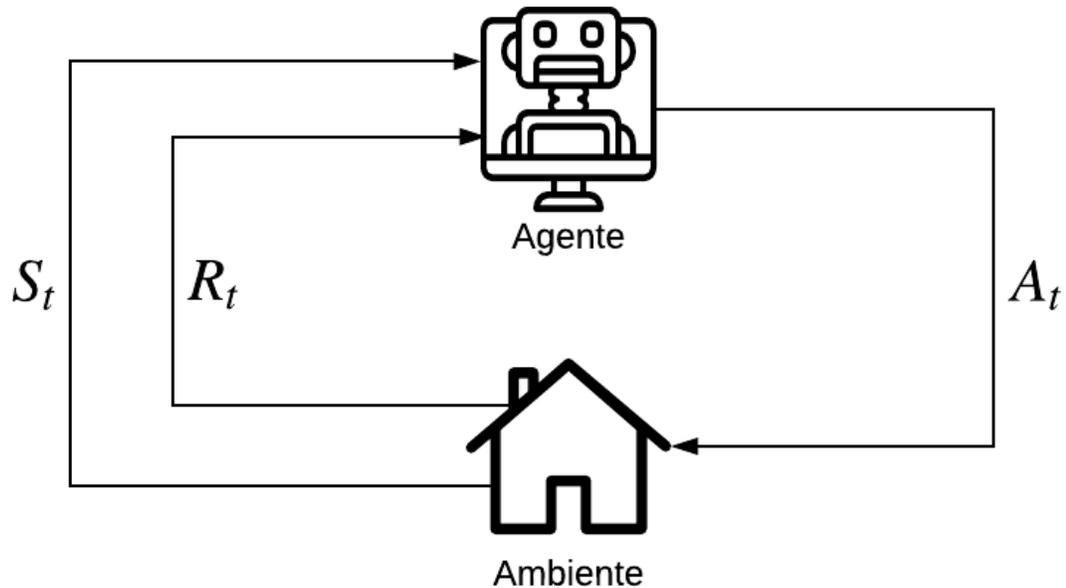
El Aprendizaje por Refuerzo («*Reinforcement Learning*» o RL) es una rama del aprendizaje automático que tiene como objetivo imitar la manera en que adquieren nuevos conocimientos los humanos, es decir, realizar acciones aleatorias ya sean acertadas o fallidas y acumular experiencia sin importar lo positivo o negativo de los resultados, proceso que se repite indefinidamente con el ambiente hasta alcanzar un alto nivel de entendimiento y dominio del sistema.

Es posible formalizar matemáticamente un problema de RL como una tupla del Proceso de Decisión Markoviano: $\langle S, A, T, R \rangle$ donde S es un conjunto finito de estados, A es un conjunto finito de acciones, T es la función de transición y R es la función de recompensa²⁸. Sin embargo, cuando S tiende a ser infinito, este algoritmo tiene la desventaja de generalizar las posibles recompensas basadas en sus experiencias pasadas, lo cual conlleva a predicciones no tan precisas y exactas.

Como se puede observar en la figura 7, el proceso de aprendizaje de RL consiste, por medio de un agente, en tomar información directamente de un ambiente en forma de estados S , efectuar acciones aleatorias A dentro de éste y analizar las consecuencias, positivas o negativas, en forma de una recompensa R . Dicho proceso se repite indefinidamente a través del tiempo hasta formar un mecanismo que utilice la experiencia acumulada y que sirva para predecir comportamientos y efectos futuros.

²⁸ OTTERLO, M. *Reinforcement Learning and Markov Decision Processes*. p. 3-42.

Figura 7. Diagrama del proceso de aprendizaje por refuerzo



Fuente: elaboración propia, empleando Lucidchart.

2.3.1. Aprendizaje-Q

Una técnica RL más avanzada es el Aprendizaje-Q o «*Q-learning*», que se basa en encontrar una metodología óptima para seleccionar acciones de acuerdo con la calidad de la recompensa que devuelven²⁹. Aquí, los valores-Q, donde Q deriva del inglés «*quality*», es decir, calidad, se definen para representar la recompensa total acumulada al realizar una acción a en un estado s . Luego, utilizando tablas para guiar la mejor acción a hacia cada estado s , considera todas las recompensas R pasadas alcanzables posibles hasta el estado s presente. Es importante destacar que esta técnica está limitada a sistemas con un conjunto reducido y discreto de acciones y estados.

²⁹ WATKINS, C., DAYAN, P. *Q-learning in Machine Learning*. p. 279-292.

2.4. Aprendizaje por refuerzo profundo

El Aprendizaje por Refuerzo Profundo («*Deep Reinforcement Learning*» o DRL) es el resultado de combinar el RL convencional con el aprendizaje profundo, donde se utilizan modelos neuronales para representar funciones y políticas de valor. En 2013, la compañía DeepMind propuso este tipo de algoritmo y demostró que, al entrenar un agente capaz de ganar en diferentes videojuegos de Atari, se pueden maximizar recompensas sin las limitaciones del RL tradicional, principalmente cuando las dimensiones del sistema de estados S crecen exponencialmente³⁰.

2.4.1. Aprendizaje-Q profundo

Una técnica de DRL que cabe destacar es el Aprendizaje-Q Profundo («*Deep-Q Learning*» o DQL), también introducida por DeepMind³¹. Esta técnica hace uso de una Red-Q Profunda («*Deep Q-Network*» o DQN), que es una red neuronal con muchas capas que estima los valores-Q. La utilidad de incluir redes neuronales surge cuando los conjuntos de acciones y estados son muy extensos.

2.4.2. Aprendizaje-Q doble

Recientemente, surgieron varias mejoras en los algoritmos originales de Aprendizaje-Q, una de ellas es la técnica llamada Aprendizaje-Q Doble o «Double Q-Learning», que consiste en la variante del algoritmo DRL que se implementó en este trabajo de graduación. La principal característica del algoritmo es la presencia de dos DQN para contrarrestar la sobreestimación de

³⁰ MNIH, V. *Playing atari with deep reinforcement learning*. p. 18.

³¹ Ídem.

los valores-Q que pueden conducir a un resultado no óptimo. Esta DQN adicional es una copia exacta de la otra, pero solo se sincroniza en cada τ pasos (un número arbitrario) y se utiliza para calcular los valores-Q objetivos³². La implementación de Wei et al. de este algoritmo usa un agente para controlar la velocidad del aire de una unidad de AC con una temperatura fija. Su experimento se llevó a cabo utilizando un sistema de AC centralizado para controlar tres ambientes diferentes³³. Emplearon el software de simulación EnergyPlus³⁴ para capacitar al agente con cien meses de datos. Los detalles de la implementación del método anterior están representados por Simonini³⁵, quien demostró la aplicabilidad de este algoritmo en el videojuego Doom utilizando la biblioteca de aprendizaje automático Tensorflow³⁶.

El presente estudio propone una implementación DRL para la optimización de energía considerando el confort térmico y el control del aire interior, que combina la implementación del aprendizaje-Q doble de Wei et al. y las mejoras detalladas por Simonini. Se usaron entornos simulados y reales para el entrenamiento, y los resultados experimentales se compararon con un método de control clásico.

³² HASSELT, H. *Deep reinforcement learning with double Q-Learning*. p. 2094-2100.

³³ WEI, T. *Deep Reinforcement Learning for Building HVAC Control*. p. 1-6.

³⁴ CRAWLEY, D. *EnergyPlus Program: creating a new-generation building energy simulation program*. p. 319 - 331.

³⁵ SIMONINI, T. *A free course in Deep Reinforcement Learning from beginner to expert*. https://simoninithomas.github.io/Deep_reinforcement_learning_Course/. Consulta: 13 de agosto de 2018.

³⁶ ABADI, M. *TensorFlow: a system for large-scale machine learning*. p. 265-283.

3. ELEMENTOS DE LA SOLUCIÓN

El aprendizaje-Q doble es el algoritmo DRL seleccionado para solucionar la problemática descrita en la primera sección. Este algoritmo, al ser del tipo RL, permite tomar información directamente de un ambiente, efectuar acciones aleatorias y analizar los efectos resultantes. El objetivo de realizar este proceso repetidamente es el de entrenar redes neuronales profundas que servirán de mecanismo para elegir la mejor acción a realizar según las condiciones dadas del ambiente. En la práctica, sin embargo, por la naturaleza de los sistemas de ventilación y AC, obtener información de los ambientes reales involucra periodos de espera muy largos para analizar los efectos de las acciones previas, atrasando así el proceso de entrenamiento. Por tanto, se recurrió al uso de herramientas de simulación, lo cual acelera considerablemente el entrenamiento e implementación del algoritmo. El detalle de las herramientas utilizadas para implementar el algoritmo DRL se presentan en la tabla IV.

Tabla IV. **Herramientas empleadas para implementación del algoritmo**

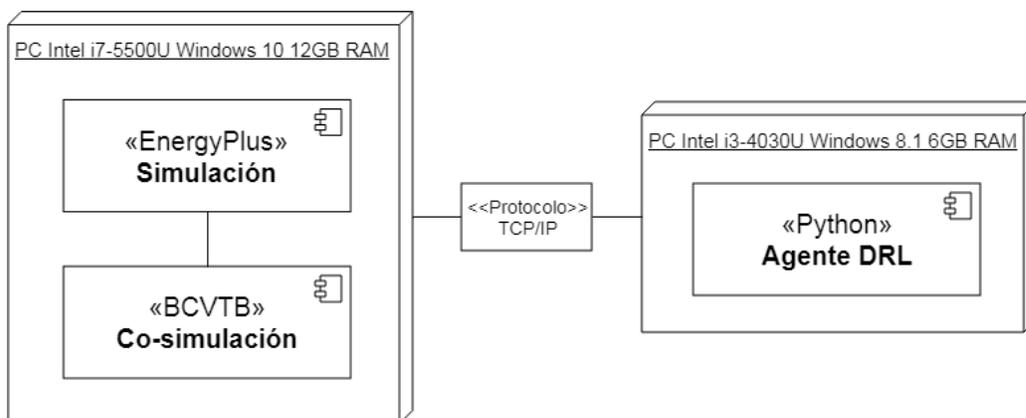
Componente	Herramienta	Versión
Simulación	Sketchup	Edición Make 2017
	OpenStudio	1.12.4
	EnergyPlus	8.5
Co-simulación	BCVTB	1.6
Agente DRL	Python	3
	TensorFlow	1.9
Internet de las Cosas	Java	8

Fuente: elaboración propia.

Así mismo, para sincronizar la simulación con el algoritmo, es necesario efectuar una co-simulación en tiempo real, que permita la comunicación entre las acciones y recompensas que calcula el agente DRL con las simulaciones de los ambientes. Además, se realizaron dos tareas principales para optimizar más el tiempo total de espera al ejecutar las simulaciones y el proceso de entrenamiento. En primer lugar, la carga computacional se distribuyó en dos computadoras conectadas a través de sockets de red; y, en segundo lugar, esta interfaz de sockets de red se agregó a la funcionalidad interna de la herramienta de co-simulación al modificar su código fuente, evitando un mayor retraso al llamar a una tercera aplicación para establecer la misma conexión.

Estas tareas de optimización permitieron el uso de equipos computacionales de gama media o baja. Se tarda entre 20 y 25 minutos en finalizar un solo año de entrenamiento, y entre 2 y 4 horas para 10 años. Estos tiempos de entrenamiento se vieron directamente afectados por los valores de los hiperparámetros, por ejemplo, cuanto más bajo es el valor del intervalo de tiempo Δt_c , más tiempo se necesita. El despliegue del proceso de entrenamiento correspondiente se muestra en la figura 8.

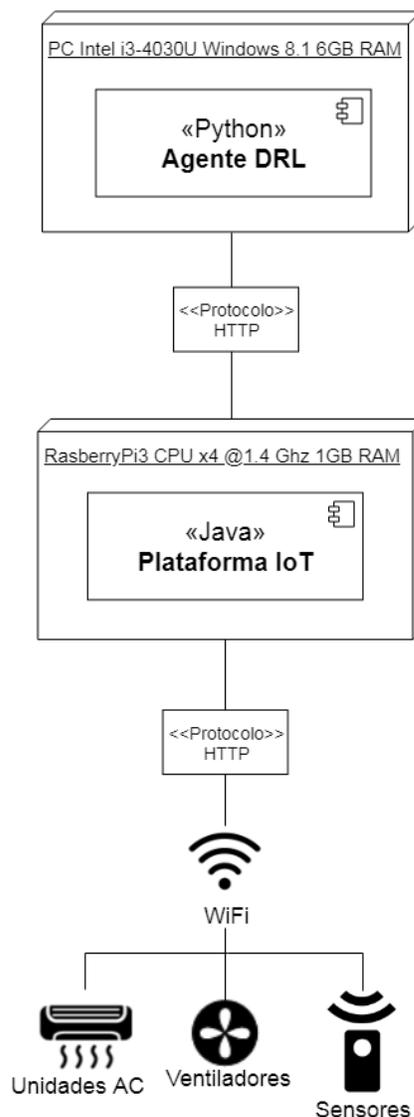
Figura 8. **Diagrama de despliegue para el proceso de entrenamiento**



Fuente: elaboración propia, empleando draw.io.

Una vez entrenado y probado el agente, se implementó en los ambientes reales mediante una plataforma de internet de las cosas, la cual se instaló en un Raspberry Pi que por comunicación wifi recibe información de los sensores, y envía señales a los ventiladores y unidades de AC. El despliegue correspondiente de la solución en los ambientes reales se detalla en la figura 9.

Figura 9. **Diagrama de despliegue en los ambientes reales**



Fuente: elaboración propia, empleando draw.io.

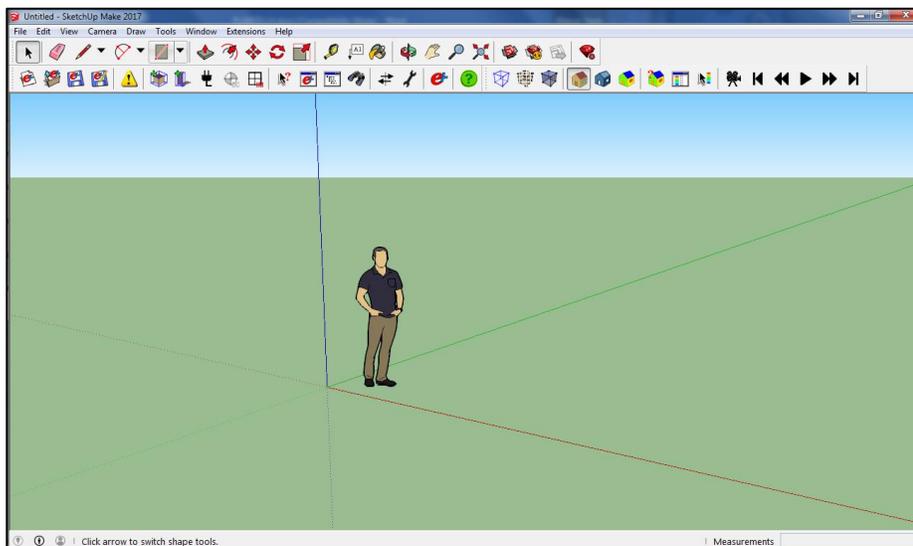
3.1. Simulación

En el ámbito de la informática, una simulación consiste en un programa que realiza una representación basada en sistemas reales, y ejecuta todos los procesos involucrados para imitar sus comportamientos. Gracias al poder computacional disponible en la actualidad, es posible realizar un gran número de simulaciones en periodos cortos de tiempo. Para la implementación del algoritmo DRL, se usaron las herramientas SketchUp Make 2017, OpenStudio v1.12.4 y EnergyPlus v8.5.

3.1.1. SketchUp

Es un programa para diseño gráfico y modelado en tres dimensiones (3-D) de estructuras arquitectónicas. Usando esta herramienta, se modelaron las dimensiones físicas de los ambientes reales y, gracias a complementos adicionales.

Figura 10. Entorno de trabajo de SketchUp



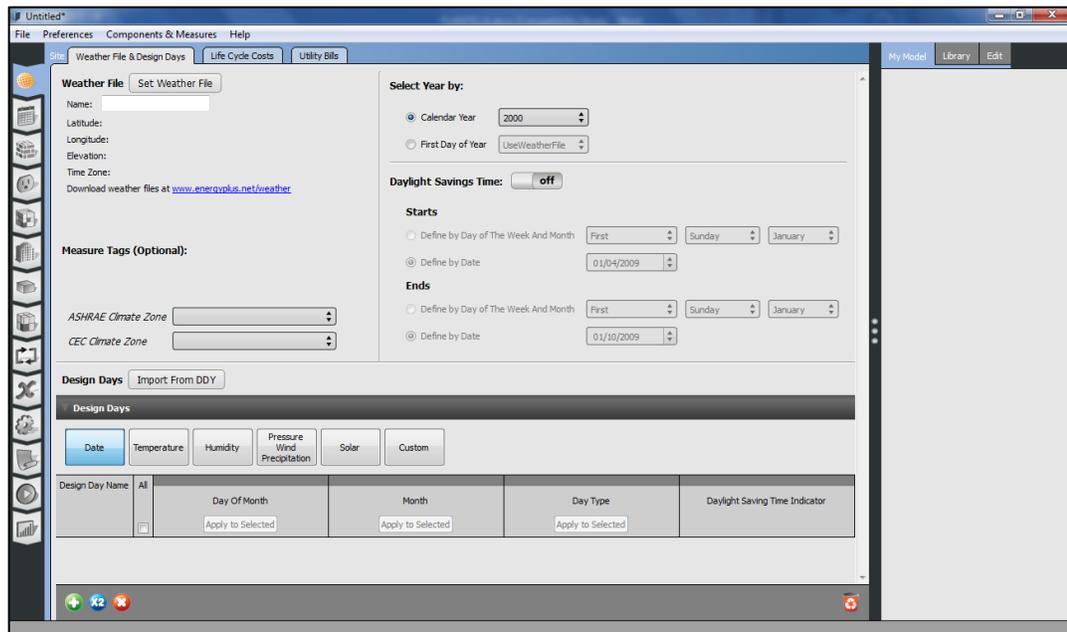
Fuente: captura de pantalla del programa SketchUp Make 2017.

Posee la gran ventaja que sus modelos se pueden exportar a otras herramientas como OpenStudio y EnergyPlus³⁷.

3.1.2. OpenStudio

Es una colección de programas multiplataforma de código abierto para apoyar la modelización energética de edificios usando EnergyPlus. Dispone de una interfaz de usuario muy completa e intuitiva, lo cual facilita la configuración de simulaciones³⁸. OpenStudio al igual que EnergyPlus, pueden recibir modelos físicos realizados en SketchUp.

Figura 11. Entorno de trabajo de OpenStudio



Fuente: captura de pantalla del programa OpenStudio v1.12.4.

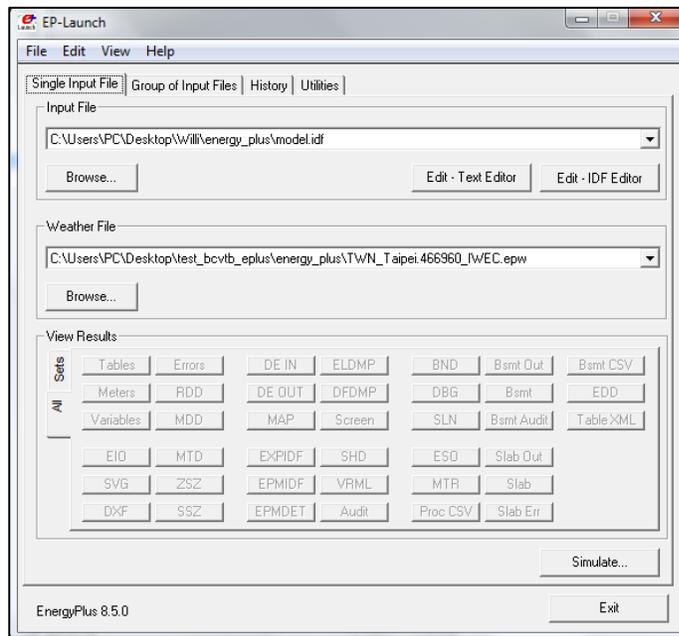
³⁷ SketchUp. <https://www.sketchup.com>. Consulta: 22 de enero 2019.

³⁸ GUGLIELMETTI, R. *OpenStudio: an open source integrated analysis platform*. p. 1-3.

3.1.3. EnergyPlus

Es un programa de código abierto para simulación de energía de edificios que ingenieros, arquitectos e investigadores utilizan para modelar el consumo de energía (para calefacción, refrigeración, ventilación, iluminación, circuitería y procesos) y el uso de agua en edificios³⁹.

Figura 12. Entorno de ejecución de EnergyPlus



Fuente: captura de pantalla del programa EnergyPlus v8.5.

Un archivo de clima específico de una región con extensión EPW y un archivo de texto de configuración con extensión IDF, son necesarios para ejecutar una simulación en EnergyPlus. Los archivos de clima pueden ser adquiridos gratuitamente desde el sitio oficial de EnergyPlus⁴⁰ y los archivos de

³⁹ CRAWLEY, D. *EnergyPlus: creating a new-generation building energy simulation program*. p. 319-331.

⁴⁰ ENERGYPLUS. *Weather Data by Region*. https://energyplus.net/weather-region/asia_wmo_region_2/TWN%20%20. Consulta: 26 de agosto de 2018.

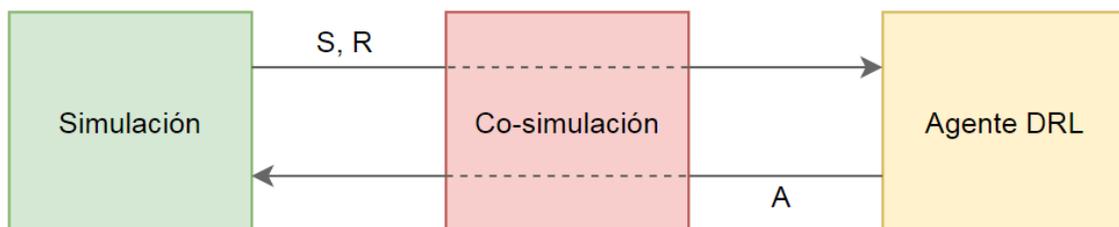
configuración se pueden crear con el programa IDFEditor que provee el paquete de instalación. Luego, en cuestión de segundos, EnergyPlus puede ejecutar una simulación y generar varios archivos de salida, incluyendo archivos web HTML y archivos separados por comas CSV.

Para configurar correctamente un archivo IDF, es necesario ser conocedor de sistemas de ventilación y AC, materiales y sus propiedades, procesos físicos y termodinámicos, entre otros. Estudiantes de ingeniería mecánica de la Universidad Nacional Chiao Tung brindaron soporte y apoyo en el uso correcto de esta herramienta.

3.2. Co-simulación

Se le denomina co-simulación a la sincronización de dos o más programas de simulación independientes, intercambiando información entre sí en tiempo de simulación. Usualmente se hacen uso de protocolos de red TCP/IP para establecer comunicaciones entre estos programas. La comunicación entre EnergyPlus y el algoritmo DRL se hizo mediante la herramienta BCVTB v1.6. La arquitectura de dicha comunicación se detalla en la figura 13.

Figura 13. **Arquitectura de comunicación por co-simulación**

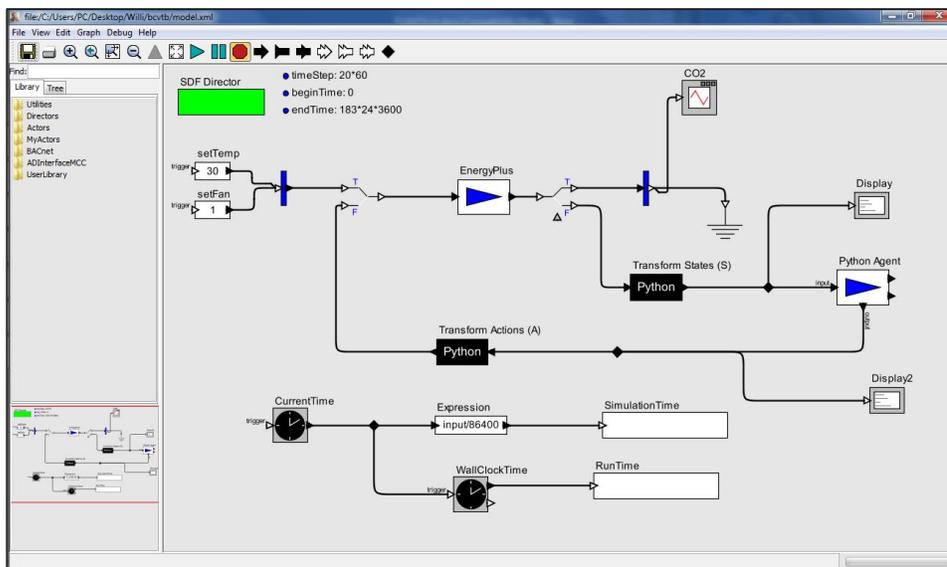


Fuente: elaboración propia, empleando draw.io.

3.2.1. BCVTB

Banco de Prueba para Controles de Construcción Virtual («*Building Controls Virtual Test Bed*» o BCVTB) es un programa de código abierto de co-simulación que permite dirigir, comunicar y sincronizar distintos programas de simulación, por ejemplo: Matlab, EnergyPlus, Modelica, Python, Java, entre otros⁴¹. Por ser un programa de código abierto, permite modificar y agregar funcionalidades personalizadas, esto empleando el lenguaje de programación Java en que fue escrito. Dicha característica se aprovechó y se creó un módulo específico para comunicar con el agente DRL, logrando optimizar más el tiempo de entrenamiento.

Figura 14. Entorno de trabajo de BCVTB



Fuente: captura de pantalla del programa BCVTB v1.6.

⁴¹ WETTER, M. *Co-simulation of building energy and control systems with the Building Controls Virtual Test Bed*. p. 185-203.

3.3. Agente DRL

Python es un lenguaje de programación interpretado, interactivo, orientado a objetos y modular. Por sus capacidades en los ámbitos de la matemática y estadística, se ha convertido en el lenguaje de programación preferido para la implementación de algoritmos de IA, dando soporte a muchas de sus ramas mediante una amplia colección de paquetes y librerías, donde TensorFlow, con su paquete visual TensorBoard, es una de las más completas y flexibles librerías para implementación de algoritmos de aprendizaje automático. Se empleó TensorFlow v1.9.0 para la implementación del algoritmo DRL.

3.3.1. TensorFlow

Es una librería de código abierto especializada en el aprendizaje automático para el lenguaje de programación Python. Sirve como un marco de trabajo que facilita la implementación de redes neuronales. Se caracteriza por ser rápido, flexible y listo para producción. TensorFlow fue desarrollado originalmente por investigadores e ingenieros de Google con el propósito de llevar a cabo investigaciones en los campos del aprendizaje automático y de las redes neuronales⁴².

3.3.2. TensorBoard

Es un paquete visual de TensorFlow que permite graficar en tiempo real el flujo de información producto de procesos de entrenamiento de algoritmos. Se compone de un conjunto de aplicaciones web para inspeccionar y comprender

⁴² ABADI, M. *TensorFlow: a system for large-scale machine learning*. p. 265-283.

ejecuciones y gráficos de TensorFlow. Gracias a esta herramienta es posible analizar el rendimiento y ejecución de algoritmos de aprendizaje automático.

3.4. Internet de las cosas

Ante el número creciente de dispositivos que se conectan a internet, surgen oportunidades para utilizarlos en maneras más avanzadas y complejas, donde varios dispositivos participan en la recolección de información o en la ejecución de comandos. De esta tendencia surgió el Internet de las Cosas («*Internet of Things*» o IoT), el cual se define como una red abierta y completa de objetos inteligentes que tienen la capacidad de autoorganizarse, compartir información, datos y recursos, reaccionando y actuando frente a situaciones y cambios en el ambiente⁴³.

Tabla V. Elementos IoT utilizados

Nombre		Elemento
Identificación	Nombramiento	Estático
	Direccionamiento	IPv4
Percepción		Sensores de temperatura, CO ₂ , humedad, de globo negro
Comunicación		Wifi
Computación	Hardware	Raspberry Pi 3
	Software	Rasbian OS, Java
Servicio		Unidades de AC Ventiladores
Semántica		Agente DRL

Fuente: elaboración propia.

⁴³ MADAKAM, S. *Internet of Things (IoT): A literature review*. p. 164.

El modelo básico de IoT consiste en tres capas: aplicación, red y percepción, y los elementos básicos que lo componen son: identificación, percepción, comunicación, computación, servicio y semántica⁴⁴. Los elementos de la solución IoT implementada se detallan en la tabla V.

3.4.1. Capa de aplicación

Esta capa es la que define la utilidad final de toda la plataforma IoT. En esta solución, mediante el agente DRL, se controlan los ventiladores y unidades AC de los ambientes, cuyas especificaciones técnicas se detallan en la tabla VI y tabla VII.

Tabla VI. **Configuración de unidades AC**

Capacidad de refrigeración	7,3 kW	(2,92 – 8,03 kW)
Capacidad de calefacción	7,3 kW	(2,92 – 8,03 kW)
Coefficiente de rendimiento (COP)	Refrigeración: 3,81	
Refrigerante	R410A	1,88 kg
Tamaño de unidad interior	1 070 mm * 180 mm * 330 mm	

Fuente: elaboración propia.

Tabla VII. **Configuración de ventiladores**

Potencia	21,7 W
Fuente de potencia	100 V - 240 V / 50 - 60 Hz

⁴⁴ AL-FUQAHA, A. *Internet of things: A survey on enabling technologies, protocols, and applications*. p. 2347-2376.

Continuación de la tabla VII.

Tamaño de instalación	400 mm * 400 mm
Flujo de escape	1 105 m ³ /h

Fuente: elaboración propia.

3.4.2. Capa de red

Esta capa se ocupa de establecer conectividad entre las otras capas del modelo IoT. Para esta solución, el módulo principal es un Raspberry Pi 3, el cual es un ordenador autocontenido, portable y de bajo costo, que junto a su sistema operativo (Rasbian OS), se configuraron para que mediante comunicación wifi recabaran información de la capa de percepción y enviaran comandos de vuelta a la capa de aplicación. Se utilizó el lenguaje de programación Java, que se caracteriza por ser un lenguaje orientado a objetos y multiplataforma, para definir la conectividad de la plataforma IoT.

3.4.3. Capa de percepción

Esta capa se encarga de recoger información del ambiente. En esta solución, cada 20 minutos se recababan datos del ambiente y se enviaban acciones de vuelta a través de la capa de red. Las configuraciones de los sensores de temperatura, temperatura de globo negro, CO₂ y humedad utilizados, están detalladas en la tabla VIII y tabla IX.

Tabla VIII. **Configuración de sensores (temperatura, CO₂, humedad)**

25 °C	Temperatura	Humedad	CO₂
Precisión	±0,3 °C	±2 %	±30 ppm
Rango	-10~60 °C	0~95 %	0~2 000 ppm
Tamaño	113,57 mm * 80 mm * 28,79 mm		
Detecta cambios una vez por segundo			
Balanceo dinámico cada 16 segundos			

Fuente: elaboración propia.

Tabla IX. **Configuración térmica de globo negro**

25 °C	Temperatura	Humedad	Temperatura de globo
Precisión	±0,6 °C	±3 %	±1 °C
Rango	0~50	0~100 %	0~80 °C
Diámetro	75 mm		
Tasa de radiación	0,95		

Fuente: elaboración propia.

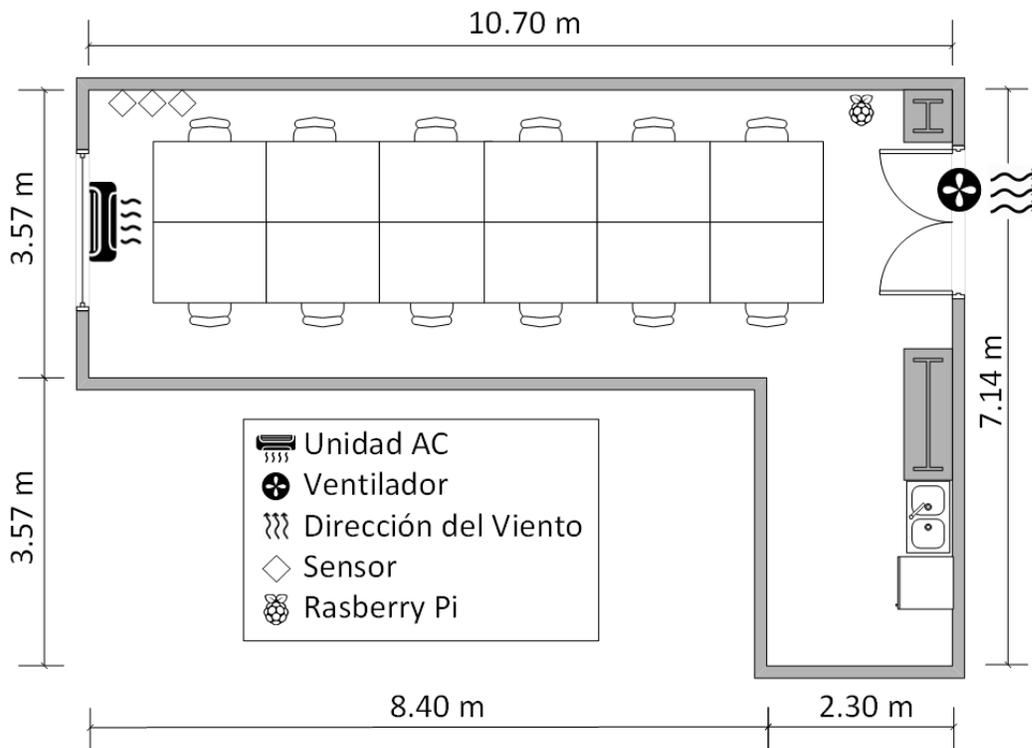
Así mismo, se usó un enchufe inteligente para monitorear el funcionamiento del ventilador, lo que permite obtener información de forma inalámbrica sobre el consumo energético, la corriente eléctrica, el voltaje y el control de encendido/apagado; las especificaciones del enchufe inteligente se detallan en la tabla X. Por otro lado, para monitorear las unidades de AC, se usaron sensores internos provistos por el fabricante. Todo el equipo presentado anteriormente, se instaló en el laboratorio 509 (figura 15) y en el aula 132 (figura 16) con fines experimentales.

Tabla X. **Configuración del enchufe inteligente**

Estándar de red	IEEE 802.11 b/g/n
Protocolo de red	TCP/IP
Banda de frecuencia inalámbrica	2,4 GHz
Señal de operación	AC 110 V 60 Hz
Consumo energético	1,5 W
Carga energética máxima	15 A / 110 V AC
Modelo	Danton Wireless TSP-151BB

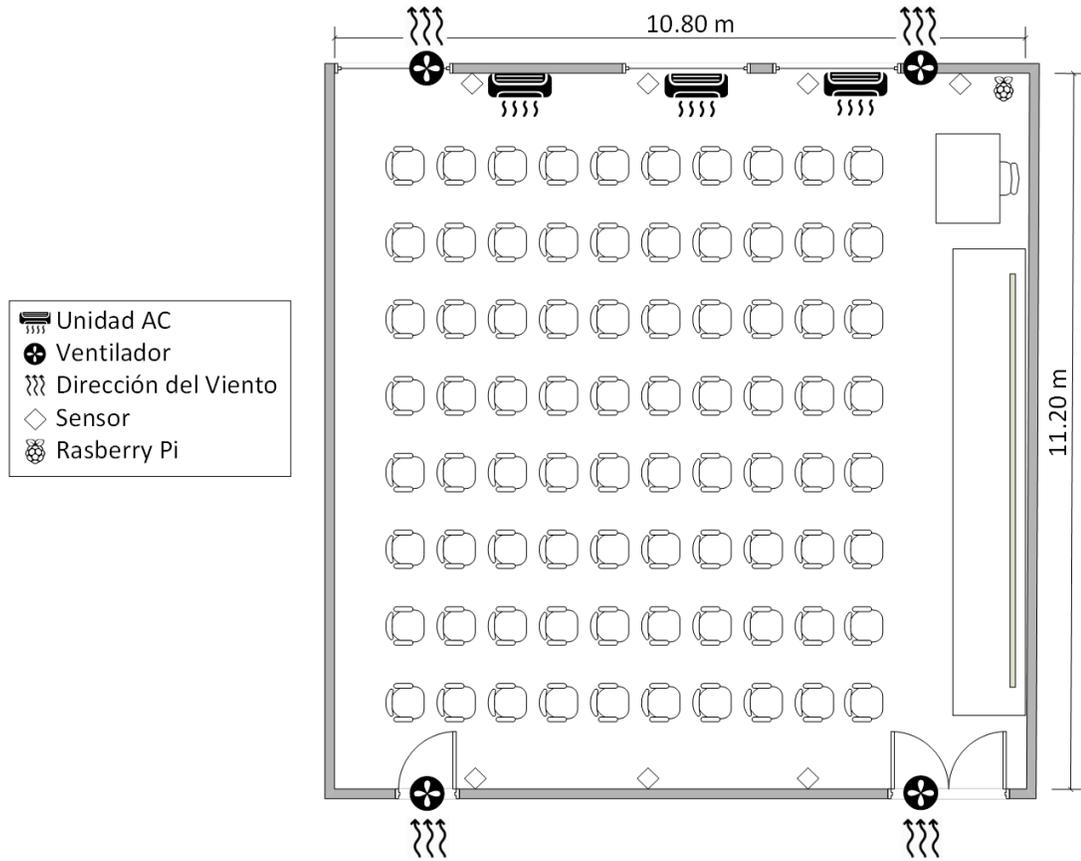
Fuente: elaboración propia.

Figura 15. **Plano del laboratorio 509**



Fuente: elaboración propia, empleando Microsoft Visio 2016.

Figura 16. Plano del salón 132



Fuente: elaboración propia, empleando Microsoft Visio 2016.

4. SISTEMA DRL PARA CONTROL DE VENTILACIÓN Y AIRE ACONDICIONADO

La razón principal de seleccionar el algoritmo DRL para sistemas de ventilación y aire acondicionado es su capacidad para manejar interacciones complejas en diferentes entornos. Para este estudio, aparecen interacciones complejas dentro de los ambientes: PMV, niveles de CO₂ y consumo de energía sujetos a unidades de AC y ventiladores. Sin embargo, se vuelve aún más complejo en asociación con el cambio de las condiciones ambientales interiores y exteriores. Por lo tanto, la utilización del aprendizaje Q-doble con el proceso Markoviano de decisión, es una solución viable. Más detalles del algoritmo serán tratados en las secciones subsiguientes.

4.1. Estado del sistema

En toda implementación de un algoritmo DRL es necesario definir la forma que tendrá el estado del sistema S , el cual no es más que un conjunto de n características C , que representan el estado actual del ambiente en un instante de tiempo t determinado.

$$S_t = \{C_1, C_2, \dots, C_n\}, n \neq \infty$$

En este caso el ambiente es un salón de clases o un laboratorio y está conformado por las siguientes cantidades físicas: la temperatura interior del ambiente, la temperatura exterior, el nivel de CO₂, el índice de PMV, la humedad relativa, la temperatura radiante promedio y la cantidad de personas dentro del salón de clases o laboratorio.

Aunque, los valores anteriores describen muy bien el estado actual del ambiente, no proporcionan información sobre cómo cambia con relación al tiempo. Por eso, se incluye también la diferencia de algunos valores del estado anterior con los del estado actual, con lo cual se espera que el algoritmo descubra la dirección y la velocidad de los cambios en el ambiente. Los valores a los que únicamente se les aplicó la diferencia corresponden a la temperatura interior, temperatura exterior, nivel de CO₂ e índice PMV. Los elementos que conforman al estado se pueden observar en la tabla XI.

Tabla XI. **Valores que conforman el estado del sistema**

Símbolo	Característica	Dimensional
t_i	Temperatura interior	°C
t_e	Temperatura exterior	°C
co_2	Nivel de concentración de CO ₂	ppm
pmv	Índice PMV	
h	Humedad relativa	%
r	Temperatura radiante promedio	°C
n	Cantidad de personas dentro del ambiente	
Δt_i	Cambio de temperatura interior	°C
Δt_e	Cambio de temperatura exterior	°C
Δco_2	Cambio en el nivel de concentración de CO ₂	ppm
Δpmv	Cambio del índice PMV	

Fuente: elaboración propia.

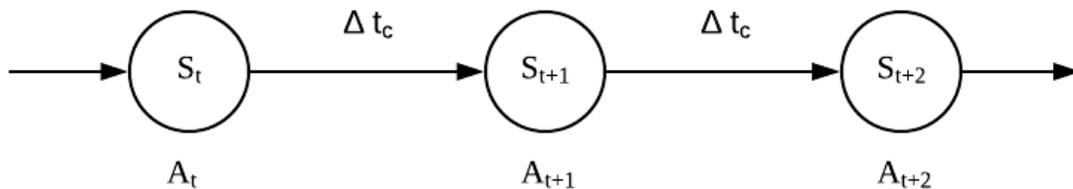
Es importante resaltar que el conjunto de estados posibles es infinito, esto debido a la naturaleza continua de las magnitudes físicas que conforman cada estado.

$$S = \{S_1, S_2, S_3, \dots S_n\}, |S| = \infty$$

4.2. Acciones de control

Las acciones A representan el medio por el cual el algoritmo interactúa con el ambiente, ya que estas influyen directamente en los valores que adquirirá el estado del sistema en el futuro. Además, las acciones desencadenan una transición de estados en el ambiente y tienen una duración de un tiempo arbitrario Δt_c .

Figura 17. **Secuencia de control de estados S y acciones A**



Fuente: elaboración propia, empleando Lucidchart.

Se poseen dos tipos de acciones: las del sistema de AC y las del sistema de ventilación. El sistema de AC se puede ajustar a una temperatura con un conjunto discreto de valores, indicado como $T = \{g \in (22, 23, 24, 25, 26, 27)\}$ donde cada elemento está dado en $^{\circ}\text{C}$. Mientras que el sistema de ventilación se puede configurar a un estado v , que está dado por el conjunto $V = \{v \in (0, 1)\}$ donde cada elemento representa si el sistema de ventilación está apagado o encendido. El conjunto de todas las acciones posibles está determinado por la expresión:

$$A = T \times V = \{(g, v): v \in V \wedge g \in T\}$$

Dicho en palabras, todas las combinaciones posibles de los elementos que conforman ambos conjuntos. Por ejemplo $A_t = (25, 1)$, correspondería a una acción realizada en el instante t en la que el sistema de AC tendrá una

temperatura de 25 °C, y los ventiladores del sistema de ventilación estarán encendidos.

4.3. Función de recompensa

Indica al algoritmo qué tan favorable es haber llegado al estado donde se encuentra actualmente el ambiente a consecuencia de la acción anterior. La función de recompensa definida para el algoritmo es la siguiente:

$$R(A_{t-1}, S_t) = w_P \cdot R_P(S_t) + w_C \cdot R_C(S_t) + w_A \cdot R_A(A_{t-1}) + w_V \cdot R_V(A_{t-1})$$

Donde:

- A_{t-1} representa la acción anterior que llevó al ambiente al estado actual.
- S_t representa el estado actual del ambiente.
- $R_P(S_t)$ representa la penalización asociada al índice PMV.
- $R_C(S_t)$ representa la penalización asociada al nivel de CO₂.
- $R_A(A_{t-1})$ representa la penalización del consumo de energía del sistema de AC.
- $R_V(A_{t-1})$ representa la penalización del consumo de energía del sistema de ventilación.
- w_P, w_C, w_A y w_V son constantes arbitrarias que sirven para regularizar los valores de penalización, referirse a la tabla XII para más detalles.

Debido a que las variables de las cuales depende la función de recompensa (índice PMV, nivel de CO₂ y consumos de energía) pueden abarcar valores con magnitudes demasiado grandes y con diferentes unidades, cada función de penalización dentro de la función de recompensa se ha normalizado, por lo cual solamente proporcionan valores de 0 a 1 en la mayoría de las circunstancias. Al aplicar la normalización se evita que una de las

variables tome valores mucho más grandes que otras y domine el resultado de la función de recompensa.

Tabla XII. **Valores de los pesos de la función de recompensa**

Nombre	Valor
w_P	1
w_C	1
w_A	$ R_P(0,4) \approx 0,64$
w_V	$ R_C(900) \approx 0,60$

Fuente: elaboración propia.

La función de recompensa actúa como una penalización, cada vez que el ambiente se encuentre fuera de los parámetros adecuados (valor de PMV aceptable, nivel bajo de CO₂) o se utiliza una gran cantidad de energía. La función tomará valores demasiado bajos, por lo que el algoritmo siempre tratará de maximizar el valor de esta función, siendo la única manera de lograrlo, el acercar el ambiente al estado ideal que le proporcione una penalización mínima.

Las constantes asociadas al consumo eléctrico del sistema de AC y del sistema de ventilación están asociadas a una condición en la cual el consumo de energía se vuelve menos favorable. Por ejemplo, el algoritmo podrá utilizar el sistema de AC a su máxima potencia hasta un PMV de $\pm 0,4$, después de este umbral, la penalización por el consumo de energía será mayor que la penalización generada por el valor del PMV. Un caso análogo sucede con el sistema de ventilación y el nivel de CO₂, después del umbral de 900 ppm, ya no será factible la utilización del sistema de ventilación.

4.3.1. Penalización de índice PMV

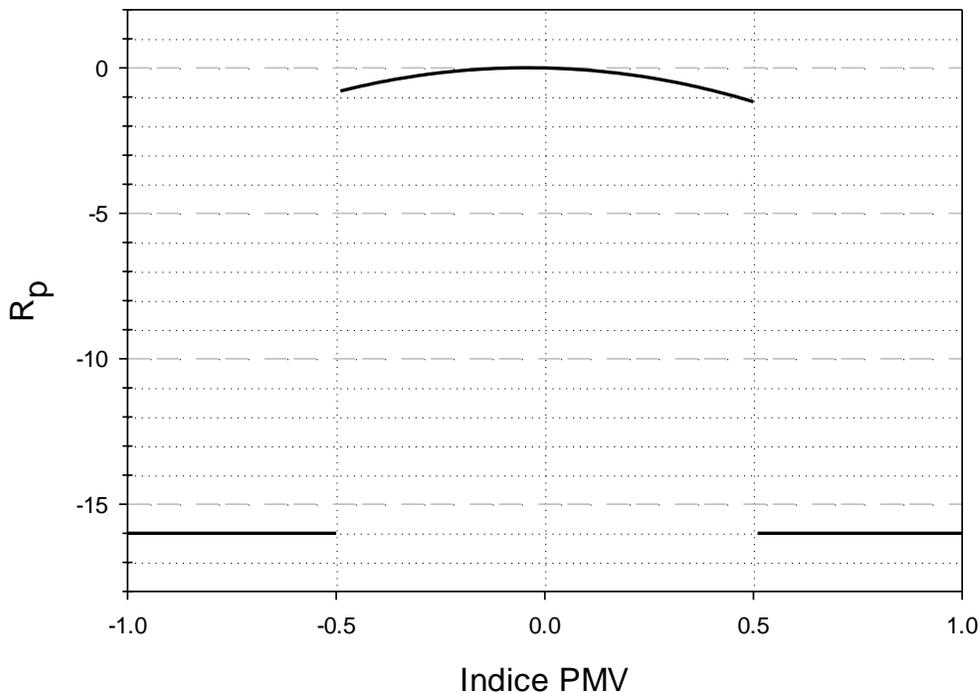
La penalización correspondiente al índice PMV del ambiente está representada en la función de recompensa por la expresión:

$$R_p(S_t) = \begin{cases} 18,09(e^{-(0,03pmv^4+0,22pmv^2)} - 1) - \frac{pmv}{3}, & |pmv| \leq 0,5 \\ -L, & |pmv| > 0,5 \end{cases}, \quad pmv \in S_t$$

Donde:

- pmv es el índice PMV que pertenece al estado actual del ambiente S_t .
- L es el corte arbitrario en la función de recompensa.

Figura 18. Penalización del índice PMV



Fuente: elaboración propia, empleando SigmaPlot 14.

El valor máximo de recompensa se da cuando el ambiente posee un índice PMV igual a 0, por lo que el algoritmo tratará de mantener el PMV en un rango cercano a dicho valor, coincidiendo con la recomendación de los estándares ANSI/ASHRAE 55 e ISO 7730 de $(-0,5, +0,5)$.

Para los valores inaceptables $[-3, -0,5] \cup [+0,5, +3]$, la función posee un corte L que sirve para que el algoritmo se centre en aprender únicamente los mejores valores de PMV y no en qué tan mal es un valor con respecto a otro. Por ejemplo, que no diferencie que un valor de -1 es mejor que $-1,5$, ya que ambos son valores no deseados.

4.3.2. Penalización del nivel de CO₂

La penalización correspondiente al nivel de CO₂ del ambiente está representado en la función de recompensa por la expresión:

$$R_C(S_t) = \begin{cases} 0, & co_2 < 750 \\ -\frac{co_2}{250} + 3, & 750 \leq co_2 \leq 1\,000, \\ -L, & co_2 > 1\,000 \end{cases} \quad co_2 \in S_t$$

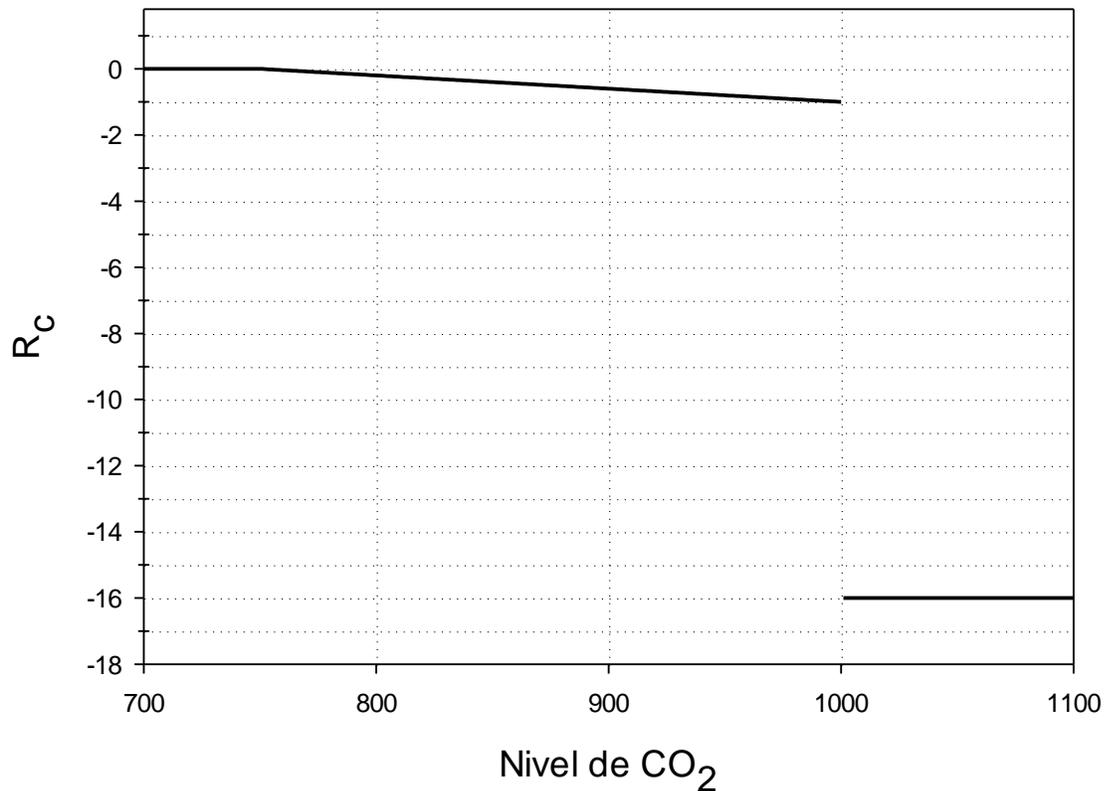
Donde:

- co_2 es el nivel de CO₂ que pertenece al estado S_t actual del ambiente en ppm.
- L es el corte arbitrario en la función de recompensa.

El valor máximo de la función se da cuando el nivel de CO₂ es menor a 750, por lo que el algoritmo tratará de mantener la concentración de CO₂ cerca de ese umbral y no bajar a menos de 750, ya que eso desperdiciaría energía eléctrica del sistema de ventilación.

Para los valores inaceptables mayores a 1 000, la función posee un corte L que sirve para que el algoritmo se centre en aprender únicamente los mejores valores de CO_2 y no en qué tan mal es un valor con respecto a otro. Por ejemplo, que no diferencie que un valor de 1 200 es mejor que 1 500, ya que ambos son también valores no deseados.

Figura 19. **Penalización del nivel de CO_2**



Fuente: elaboración propia, empleando SigmaPlot 14.

4.3.3. Penalización del sistema de AC

La penalización del sistema de AC se da cuando el ambiente cambia del estado S_{t-1} al estado S_t , la transición la causa la realización de una acción A_{t-1} en el instante $t - 1$. La realización de tal acción conlleva a la utilización de una

cantidad de energía por parte del sistema de AC durante el tiempo que dura la transición, es decir entre los instantes $t - 1$ y t . La penalización de dicha energía está representada en la función de recompensa por la siguiente expresión:

$$R_A(A_{t-1}) = -\frac{\bar{E}_{AC}(t-1, t) - Min_{AC}}{Max_{AC} - Min_{AC}}$$

Donde:

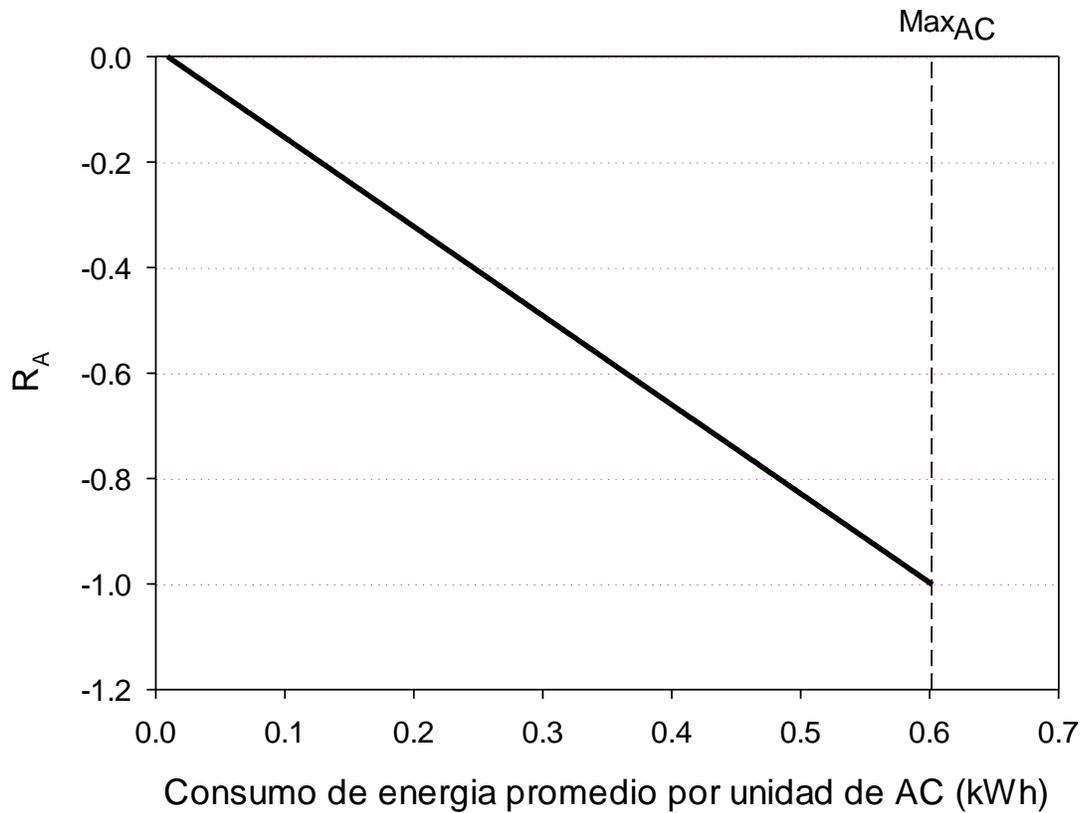
- \bar{E}_{AC} representa el consumo de energía promedio por unidad de AC en el tiempo que duró la acción, es decir entre los instantes $t - 1$ y t , y el valor de energía se da en kilowatt-hora (kWh).
- Min_{AC} es el consumo de energía mínimo que se puede utilizar en una unidad de AC durante una transición de estado del ambiente, siendo este valor de $0,009 kWh$ en este estudio.
- Max_{AC} es el consumo de energía de una unidad de AC trabajando a su capacidad nominal durante una transición de estado en el ambiente, siendo este valor de $0,6015 kWh$ en este estudio.

El valor de Min_{AC} es diferente a 0 porque las unidades de AC utilizadas nunca dejan de consumir energía, debido a que utilizan un mínimo de energía para alimentar los circuitos de control y comunicación en su interior. Los valores de Max_{AC} y Min_{AC} son propios de las unidades de aire acondicionado utilizadas y fueron obtenidos de datos históricos del funcionamiento de las unidades de AC.

Cuando todas las unidades de AC están en su capacidad nominal, la penalización tenderá al valor -1 ; en otras palabras, la penalización será máxima cuando las unidades estén trabajando a su máxima capacidad. Por otro lado,

cuando el sistema de AC no se esté utilizando para enfriar el ambiente, el resultado de la penalización tenderá al valor de 0.

Figura 20. **Penalización de la energía del sistema de AC**



Fuente: elaboración propia, empleando SigmaPlot 14.

La energía promedio utilizada por cada unidad de AC entre dos instantes de tiempo (t_0, t_f) está dada por la siguiente expresión:

$$\bar{E}_{AC}(t_0, t_f) = \frac{1}{N} \sum_{n=1}^N E_n(t_0, t_f)$$

Donde:

- N representa la cantidad de unidades de AC que hay en el sistema.
- E_n representa la energía consumida por la n -ésima unidad de AC del sistema.
- t_o representa el instante en el que inicia la medición de energía en segundos (s).
- t_f representa el momento de finalización de la medición de energía en segundos (s).

Por último, se define la energía consumida por la n -ésima unidad de AC, la cual únicamente depende de ella misma, sin verse afectada por el funcionamiento de las otras unidades.

$$E_n(t_o, t_f) = \frac{1}{C_f} \int_{t_o}^{t_f} k_n(t) dt$$

Donde:

- $k_n(t)$ representa la potencia instantánea en watts (W) que está utilizando la n -ésima unidad de AC en el instante t .
- C_f es el factor de conversión para transformar la energía de joules a kilowatt-hora ($1 kWh = 3,6 * 10^6 J$)⁴⁵.

En la industria de equipos electrodomésticos, los fabricantes utilizan la dimensional de kilowatt-hora ya que los valores de energía en Joules tienden a

⁴⁵ FREEDMAN, Y., ZEMANSKY, S. *Física Universitaria*. 12a ed. p. 146.

poseer una gran magnitud, y al momento de realizar informes o gráficas, se dificulta su debida representación e interpretación.

4.3.4. Penalización del sistema de ventilación

Toda acción que realice el algoritmo también conlleva a la utilización de una determinada cantidad de energía por parte del sistema de ventilación. Tal consumo de energía está representado en la función de recompensa por la expresión:

$$R_V(A_t) = -v, \quad v \subset A_t$$

Donde:

- v representa la acción del sistema de ventilación que fue realizada por la acción A_t durante la transición del ambiente entre los instantes $t - 1$ y t .

Cuando el algoritmo encienda el sistema de ventilación en una transición, la penalización será de -1 ; cuando el sistema de ventilación se mantenga apagado durante una transición, la penalización será de 0 ya que no existió consumo de energía alguno.

A diferencia de las unidades de AC, los ventiladores que conforman el sistema de ventilación poseen un consumo de energía que podría considerarse constante. Por lo que, independientemente de las condiciones del ambiente, siempre consumirán la misma cantidad de energía. El algoritmo, en vez de optimizar la energía del sistema de ventilación, lo que hace es optimizar su uso, encendiendo o manteniendo apagados los ventiladores si las condiciones del ambiente lo ameritan.

4.4. Arquitectura de la implementación

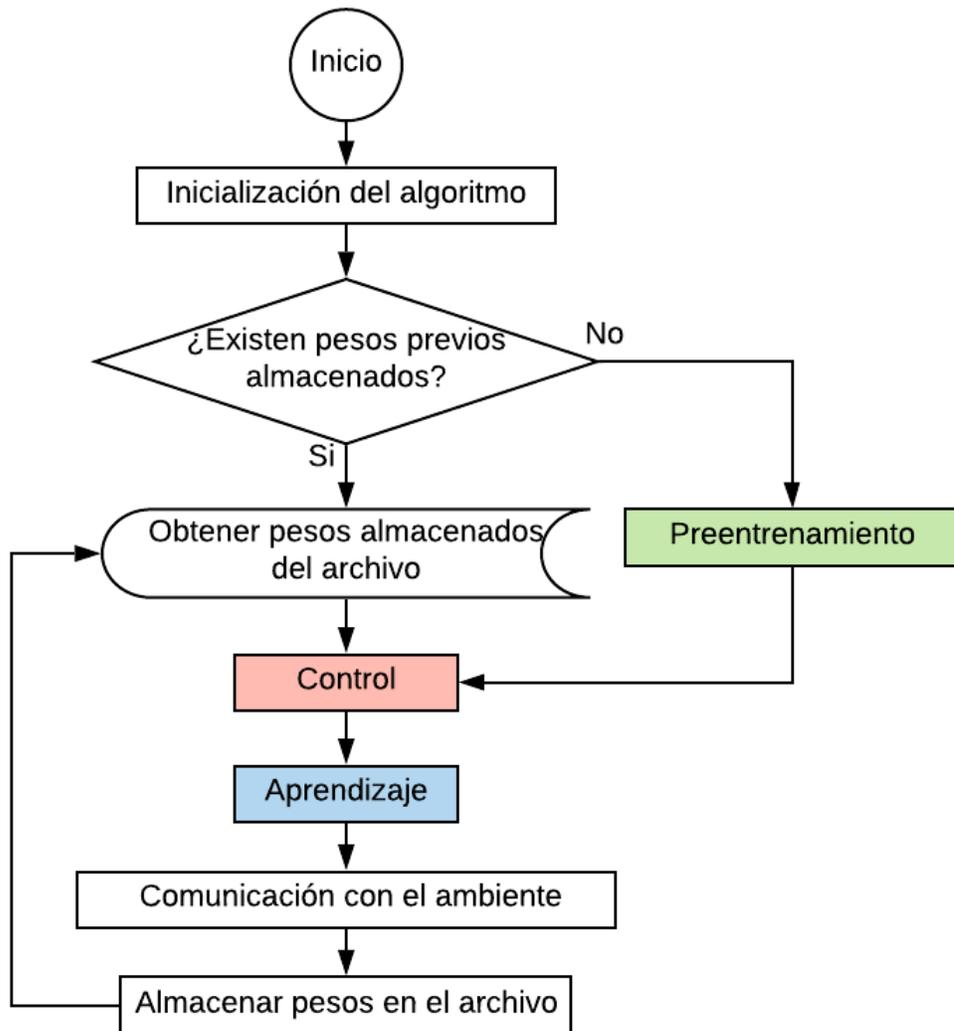
La implementación del algoritmo de aprendizaje-Q profundo está conformada por los siguientes componentes:

- Memoria de experiencias previas M
- Red neuronal principal Q
- Red neuronal auxiliar \hat{Q}
- Ambiente
- Archivo de pesos

El algoritmo consiste en una serie de procesos que dictan cómo se relacionan entre sí los componentes anteriores, entre los cuales hay que destacar tres procesos importantes: el de preentrenamiento, el de control y el de entrenamiento o aprendizaje. Estos procesos serán explicados con mayor detalle más adelante.

La inicialización del algoritmo mencionado en la figura 21, corresponde a la creación de la memoria M , donde se almacenará la experiencia que vaya adquiriendo el algoritmo durante su ejecución. Se inicia la comunicación con el ambiente y se crean las redes neuronales que utilizará. La condicional sobre la experiencia previa corresponde a la verificación de la existencia del archivo de pesos que, en caso de no existir, se realizará el proceso de preentrenamiento para generarlo. El archivo de pesos es el encargado de almacenar los pesos de las redes neuronales con el fin de no perder lo aprendido por el algoritmo en las futuras ejecuciones.

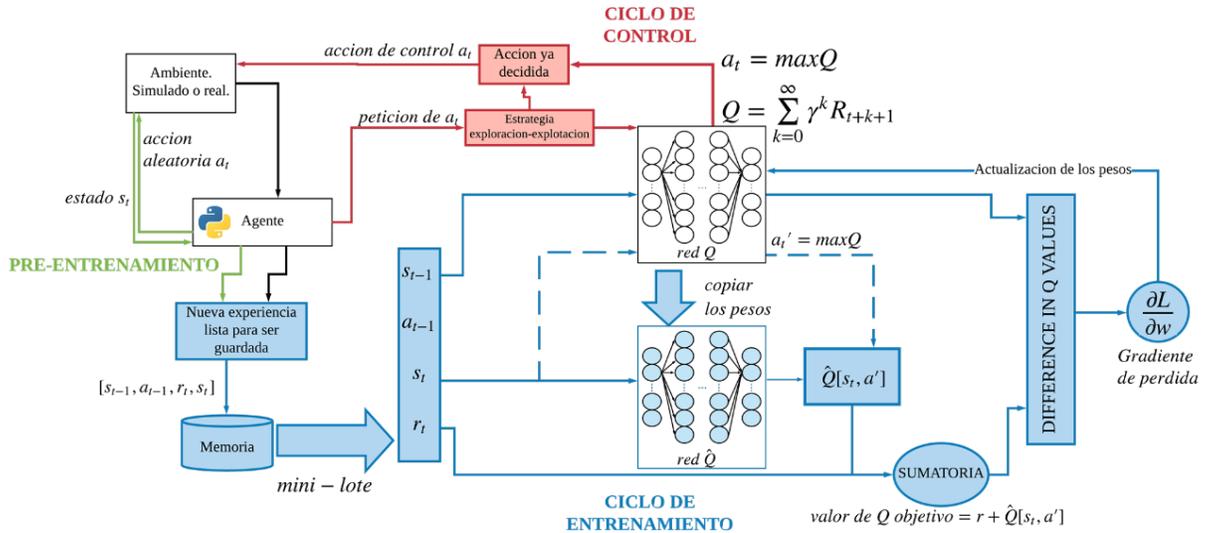
Figura 21. Diagrama de flujo principal del algoritmo



Fuente: elaboración propia, empleando Lucidchart.

La ejecución del algoritmo termina con la verificación de la comunicación con el ambiente, los procesos de aprendizaje y control se repiten indefinidamente si aún hay comunicación; mientras que, si no lo hay, se sobrescribe el archivo de pesos y se finaliza la ejecución.

Figura 22. Arquitectura del algoritmo



Fuente: elaboración propia, empleando Lucidchart.

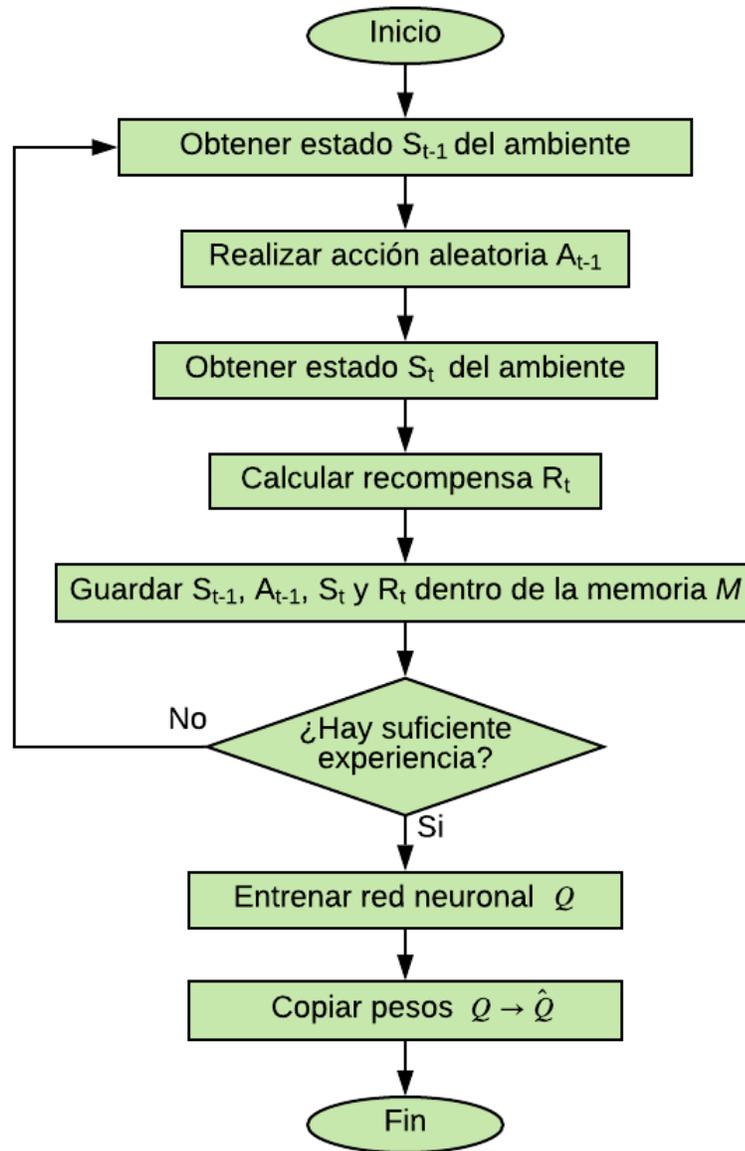
La implementación del algoritmo de aprendizaje-Q profundo utilizada (vista en la figura 22) se basa en la publicación de la investigación *Deep Reinforcement Learning for Building HVAC Control*⁴⁶. En ella, implementaron el algoritmo de aprendizaje-Q profundo para controlar el flujo de aire de los sistemas HVAC de distintas zonas para mantener la temperatura dentro de un rango adecuado y reducir el consumo de energía eléctrica.

4.4.1. Proceso de pre-entrenamiento

Este proceso únicamente se desarrolla cuando el algoritmo nunca ha sido ejecutado. No existe ningún archivo asociado a los pesos de las redes neuronales y la memoria M donde guarda las experiencias previas está vacía, por lo que su experiencia con el ambiente es nula. Los pasos del proceso de preentrenamiento se pueden observar en la figura 23.

⁴⁶ WEI, T. *Deep Reinforcement Learning for Building HVAC Control*. p. 1-6.

Figura 23. Diagrama de flujo del proceso de preentrenamiento



Fuente: elaboración propia, empleando SigmaPlot 14.

Esta fase es de exploración, ya que el algoritmo explora el conjunto de acciones. Específicamente, realiza acciones aleatorias y guarda en la memoria M cómo el ambiente responde a ellas. Cada experiencia de la memoria está compuesta por cuatro elementos: el estado del ambiente antes de realizar la

acción, la acción realizada, el estado del ambiente después de la acción, y qué tan favorable fue la transición (valor de la función de recompensa).

$$Experiencia = [S_{t-1}, A_{t-1}, S_t, R_t]$$

El algoritmo continúa ejecutando acciones aleatoriamente hasta que la memoria alcanza la cantidad mínima de experiencias necesarias para poder entrenar la red neuronal principal Q . Una vez alcanzada esa cantidad, se procede a inicializar los pesos de la red neuronal, para posteriormente entrenarla y copiar los pesos hacia la red auxiliar \hat{Q} . Por último, se almacenan los pesos en un archivo para las ejecuciones posteriores del algoritmo.

4.4.2. Proceso de control

El proceso de control corresponde al conjunto de pasos en los que el algoritmo predice la acción que llevará el ambiente al estado más favorable posible, esto quiere decir que dicho estado tendrá un buen nivel de CO₂, un índice adecuado de PMV, y utilizará la menor cantidad de energía posible.

Esta fase se denomina fase de explotación, ya que el algoritmo se encuentra aprendiendo de sus experiencias pasadas y dependiendo de las circunstancias, buscará sacarles el máximo provecho. En este proceso se maneja un valor de probabilidad ε llamado factor codicioso o «*greedy factor*». Este valor indica la probabilidad de explorar nuevamente el conjunto de acciones, en otras palabras, la probabilidad de realizar una acción de forma aleatoria. Por otro lado, el complemento, la probabilidad de $1 - \varepsilon$, representa la probabilidad de explotar las experiencias previas, es decir, predecir la mejor acción con base en las experiencias acumuladas. El valor del factor ε varía de

forma descendente en función de la cantidad de acciones que ha realizado el algoritmo. El factor está dado por la siguiente expresión:

$$\varepsilon(n) = \varepsilon_{min} + (1 - \varepsilon_{min}) \cdot e^{-\varepsilon_{decay} \cdot n}$$

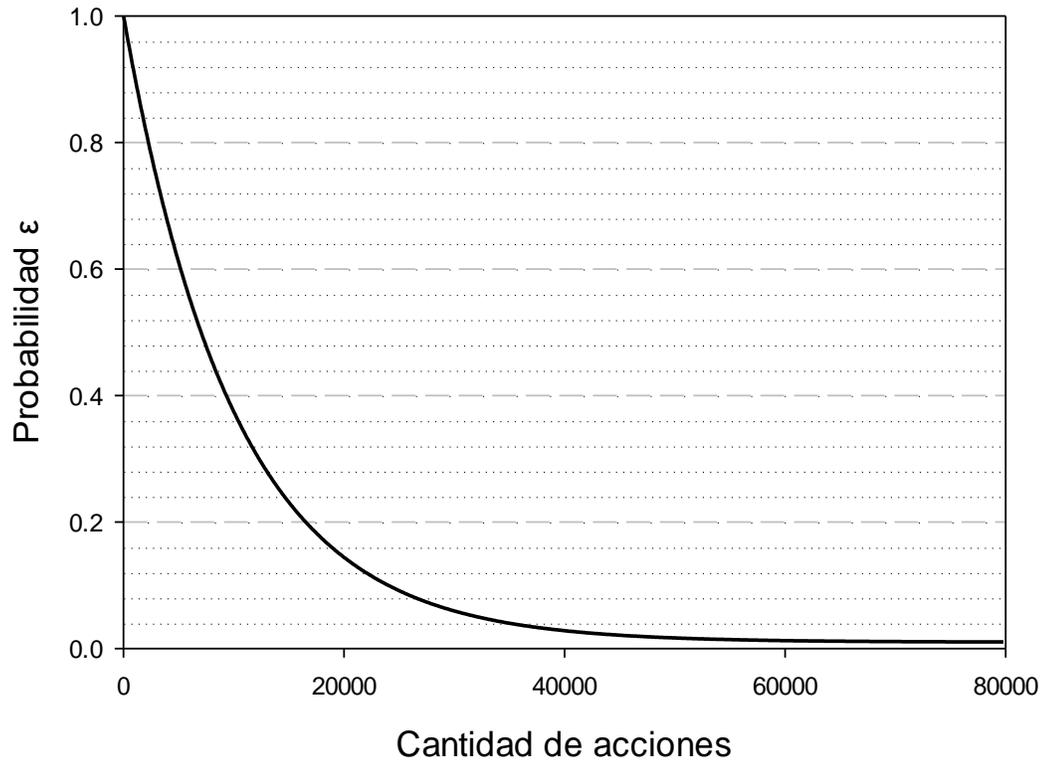
Donde:

- ε_{min} representa el valor mínimo al que puede decaer la probabilidad del factor codicioso.
- ε_{decay} representa la rapidez a la que decae exponencialmente el factor codicioso conforme se realizan las acciones.
- n representa la cantidad de acciones realizadas por el algoritmo.

Al principio, debido a la poca experiencia que tiene el algoritmo, la probabilidad de exploración será muy alta, por lo que la mayoría de las acciones que haga el algoritmo serán aleatorias; pero conforme el algoritmo vaya aprendiendo, cada vez irá haciendo menos acciones aleatorias y tratará más de predecir la mejor acción con base en su experiencia acumulada. El valor de ε nunca decae a 0, por lo que, a largo plazo, el algoritmo tendrá siempre una probabilidad mínima ε_{min} de realizar una acción aleatoria.

El proceso de control inicia recibiendo el estado S_{t-1} del ambiente, se calcula el valor del factor codicioso y, dependiendo de la decisión tomada por el factor ε , se elegirá una acción aleatoriamente, o la mejor acción con base en la predicción de la red neuronal principal Q . La acción elegida A_{t-1} se envía al ambiente para su realización.

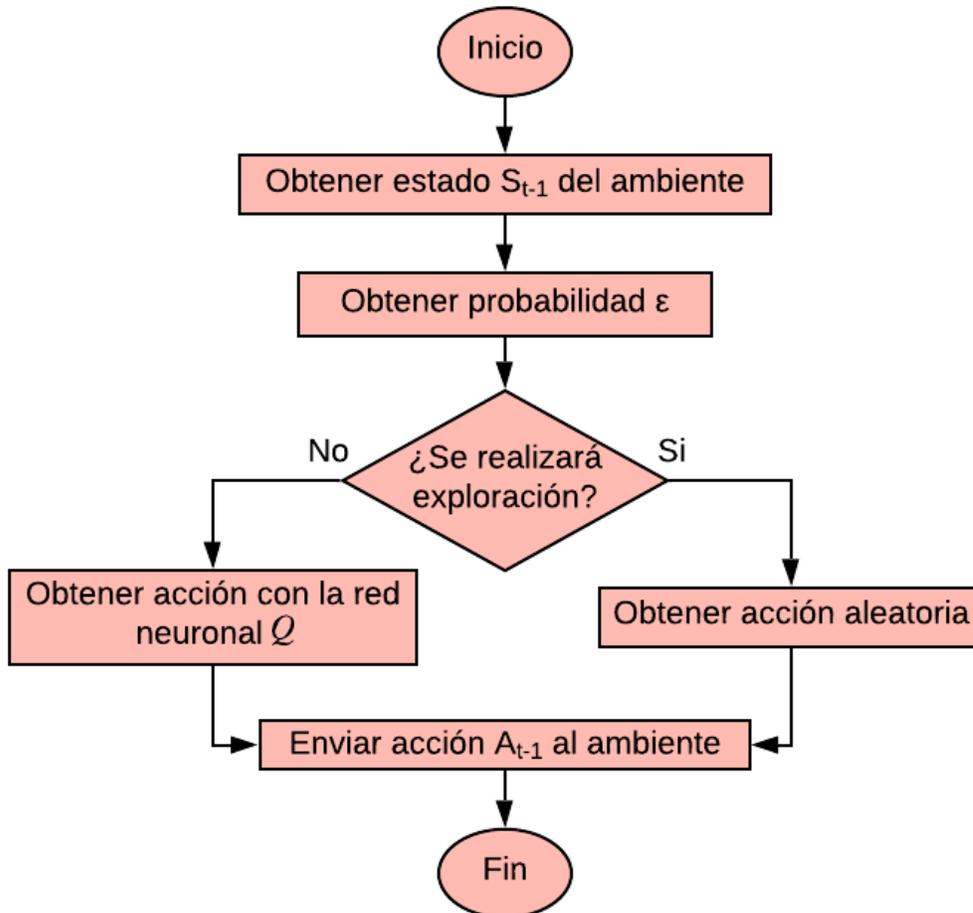
Figura 24. Valor del factor codicioso



Fuente: elaboración propia, empleando SigmaPlot 14.

La predicción de la mejor acción se realiza al ingresar los datos del estado S_{t-1} a la red neuronal Q . Esta red devuelve el valor-Q para cada una de las acciones del espacio de estados. Para determinar cuál es la mejor, únicamente se elige la acción que tenga asociado el valor-Q más alto. Los pasos del proceso de control se pueden observar en la figura 25.

Figura 25. Diagrama de flujo del proceso de control

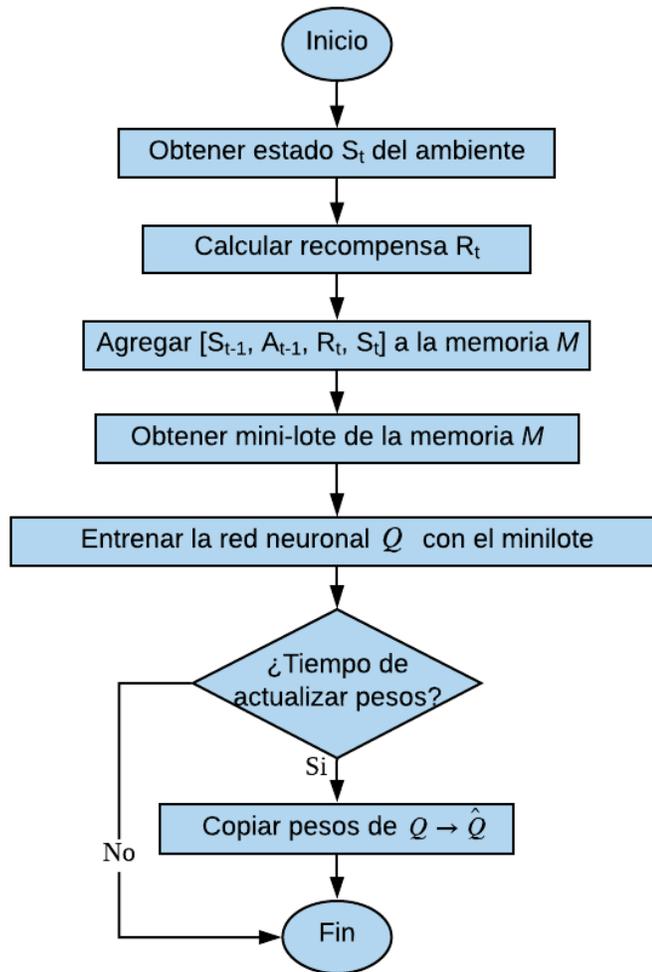


Fuente: elaboración propia, empleando Lucidchart.

4.4.3. Proceso de aprendizaje

El proceso de aprendizaje corresponde al conjunto de pasos que le permiten al algoritmo aprender de las experiencias previas. Cuando se dice que el algoritmo aprende, lo que sucede realmente es que los pesos de las redes neuronales Q y \hat{Q} se modifican con base en las experiencias almacenadas en M . Este proceso ocurre inmediatamente después del proceso de control, por lo que es posible decir que el algoritmo siempre está aprendiendo. Los pasos del proceso de aprendizaje se pueden observar en la figura 26.

Figura 26. Diagrama de flujo del proceso de aprendizaje



Fuente: elaboración propia, empleando Lucidchart.

Como se observa en la figura 26, primero se obtiene del ambiente el estado S_t , el cual corresponde al estado al que llegó el ambiente después de haber realizado la acción A_{t-1} en el proceso de control, Δt_c segundos antes. Se calcula la recompensa R_t que representa qué tan favorable fue la transición realizada con la acción A_{t-1} . Con lo anterior, se tienen todos los elementos para guardar este nuevo conjunto de experiencia $[S_{t-1}, A_{t-1}, R_t, S_t]$ en M . Debido a

que la cantidad de experiencias almacenadas en la memoria pueden ser demasiadas, ralentizando así la ejecución del algoritmo, se entrena con una cantidad menor de experiencias. A este grupo de datos se le conoce como mini-lote, y son elegidos de forma aleatoria de la memoria.

Se entrena la red Q con el mini-lote elegido, de modo que las predicciones de la red principal Q se parezcan a las predicciones realizadas por la red auxiliar \hat{Q} . Es importante recalcar que durante el proceso de control no se utilizó la red \hat{Q} , debido a que el único fin de esta es apoyar en el aprendizaje de la red principal Q . Para finalizar el proceso de aprendizaje, cada cierto número arbitrario de acciones τ del algoritmo, los pesos de la red neuronal Q se copian a la red neuronal \hat{Q} para que lo aprendido por una red neuronal le sea transmitido a la otra.

4.5. Hiperparámetros

A lo largo de toda la implementación del algoritmo se han dejado muchas constantes solamente expresadas junto con otras que son por defecto del tipo de algoritmo utilizado. Al conjunto de estas constantes se les conoce en inteligencia artificial como hiperparámetros, y tanto la cantidad como los valores de estos varían de una implementación a otra.

Los valores de los hiperparámetros influyen directamente en el comportamiento que adquiere el algoritmo por lo que, diferentes conjuntos de valores, producen resultados muy diferentes. En la práctica, buscar el conjunto adecuado de valores conlleva un proceso lento de prueba y error. La configuración de valores usada en los hiperparámetros de este estudio está detallada en la tabla XIII.

Tabla XIII. Valores de hiperparámetros

Nombre	Valor
Δt_c	20 minutos
γ	0,95
Mini-lote	128
ε_{min}	0,01
τ	640
α	0,0003
L	16
ε_{decay}	0,0001
Unidades en las capas de Q y \hat{Q}	100, 150, 200, 250, 400, 400
Activación en Q y \hat{Q}	eLU
Optimización en Q y \hat{Q}	RMSprop
Inicialización en Q y \hat{Q}	Xavier

Fuente: elaboración propia.

5. ENTRENAMIENTO DEL SISTEMA

Las acciones realizadas por el algoritmo tardan varios minutos en afectar el ambiente real, ya que la temperatura y el nivel de CO₂ cambian muy lentamente. Por ello, no es viable ejecutar el algoritmo desde cero en el ambiente real, ya que se tardaría demasiado tiempo en entrenar, configurar y realizar las pruebas correspondientes. Para solventar este problema, se ha realizado una simulación de software de los ambientes para reducir drásticamente el tiempo de entrenamiento del algoritmo.

5.1. Desarrollo de la simulación

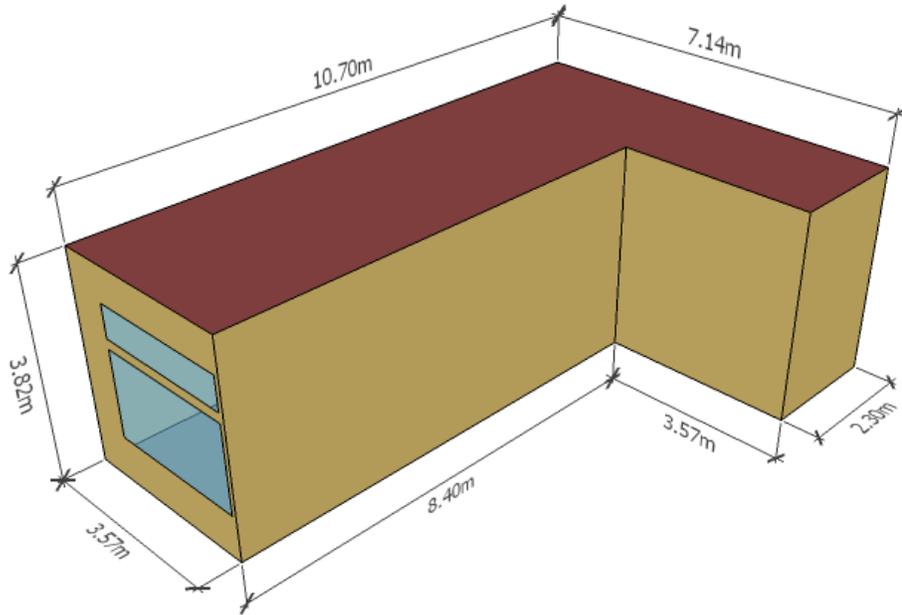
Para el desarrollo de las simulaciones se llevaron a cabo los pasos siguientes: la creación del modelo, la adaptación del modelo, la creación de los sistemas y, por último, la comunicación entre el algoritmo con las simulaciones.

5.1.1. Creación del modelo

Para la creación del modelo se utilizó la herramienta Sketchup en conjunto con el complemento de Open Studio. En esta fase se crea un archivo con extensión IDF, que contiene las propiedades espaciales de los ambientes reales, entre las cuales se pueden mencionar:

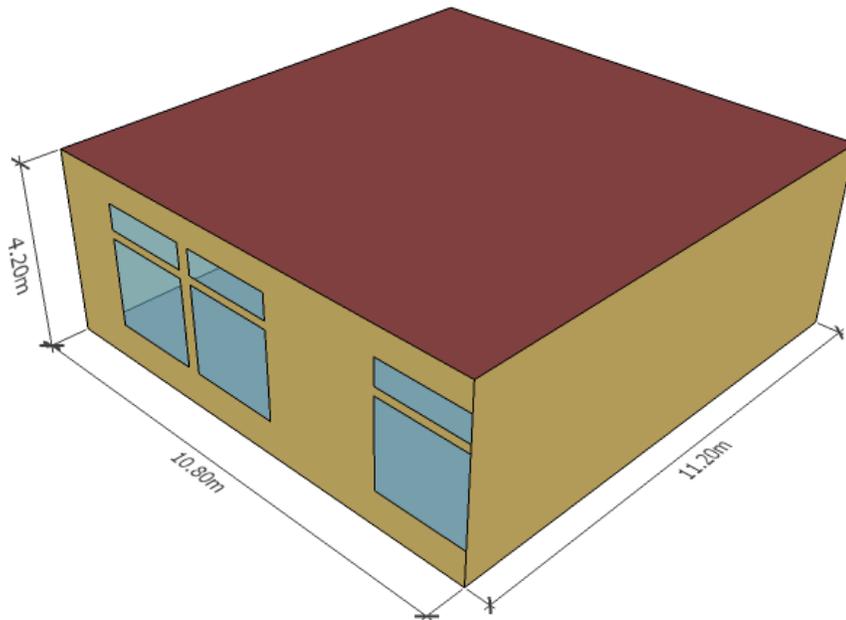
- Espacio tridimensional
- Orientación con respecto a los puntos cardinales
- Coordenadas geográficas del ambiente
- Localización de puertas y ventanas

Figura 27. **Modelo 3D del laboratorio 509**



Fuente: elaboración propia, empleando SketchUp Make 2017.

Figura 28. **Modelo 3D del salón 132**



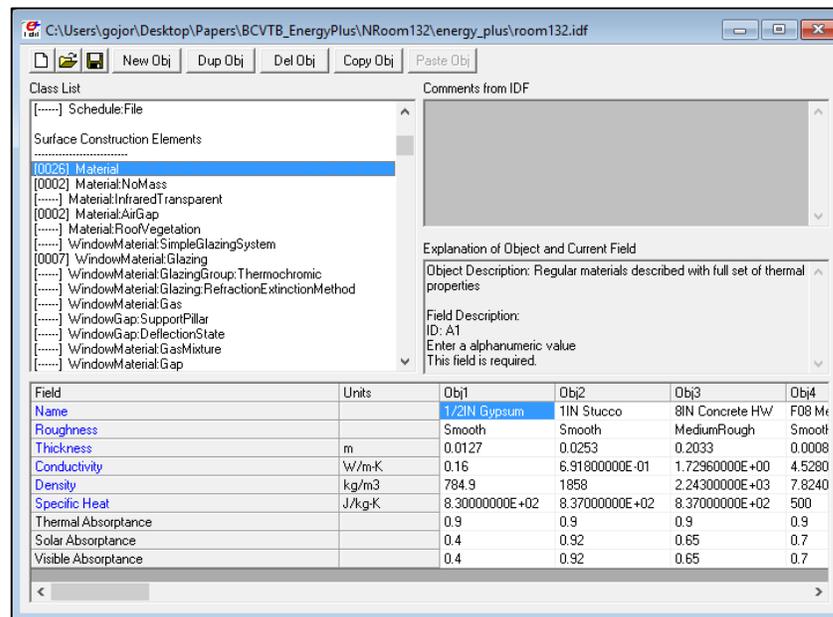
Fuente: elaboración propia, empleando SketchUp Make 2017.

5.1.2. Adaptación del modelo

En esta parte se agregan características físicas y se definen los algoritmos que determinan sus comportamientos, por ejemplo, el nivel de CO₂ promedio del exterior y el algoritmo que define su comportamiento. Se utilizó el programa IDFEditor que viene incluido en el paquete de instalación de EnergyPlus, el cual permite agregar características a los archivos IDF de una manera más amigable. Entre algunas características se puede mencionar:

- Algoritmo de cálculo de la sombra
- Algoritmo de balance de calor
- Algoritmo de balance de contaminación del aire
- Horarios de ocupación y uso del equipo eléctrico
- Materiales de construcción en paredes, puertas, ventanas, techo, entre otros.

Figura 29. Interfaz de IDFEditor



Fuente: captura de pantalla del programa IDFEditor v1.48b.

5.1.3. Modelación de los sistemas

Por último, solamente queda representar los sistemas de ventilación y AC en el modelo de simulación. Nuevamente con el programa IDFEditor se agregan estas características en su sección correspondiente dentro de los archivos IDF. Entre las características que se agregan están:

- Caudal de flujo de los ventiladores
- Presión de aire de los ventiladores
- Eficiencia energética de los ventiladores
- Capacidad total de enfriamiento de las unidades de AC
- Factor COP de las unidades de AC
- Eficiencia energética de las unidades de AC

Las características fueron obtenidas de los fabricantes de los sistemas que se utilizan en el ambiente real.

5.1.4. Comunicación

Debido a que el tiempo dentro de la simulación no es el mismo tiempo en el que corre la ejecución del algoritmo, es necesaria la utilización de un intermediario para sincronizar ambos procesos. En este caso, se utilizó el programa BCVTB, el cual permite la sincronización y transferencia de datos de distintos ambientes de simulación entre sí.

5.2. Entrenamiento

Para el entrenamiento en la simulación se utilizó un archivo con los datos climáticos de la ciudad de Taipéi, este archivo se obtuvo del sitio web oficial de

EnergyPlus⁴⁷. Se utilizaron los datos de esta ciudad debido a que es el archivo de clima disponible más cercano a la ciudad donde se realizaron las pruebas (Hsinchu, Taiwán).

El tiempo de ejecución de la simulación se divide en periodos de entrenamiento. Cada periodo de entrenamiento consiste en la simulación de los datos climáticos del mes de mayo al mes de octubre. La razón de usar únicamente estos meses del archivo de clima es porque son los meses más calurosos del año y, en los otros meses, los sistemas de AC no son muy utilizados por la población de Taiwán. El resumen de la configuración de la simulación se puede encontrar en la tabla XIV.

Tabla XIV. **Detalles de configuración de las simulaciones**

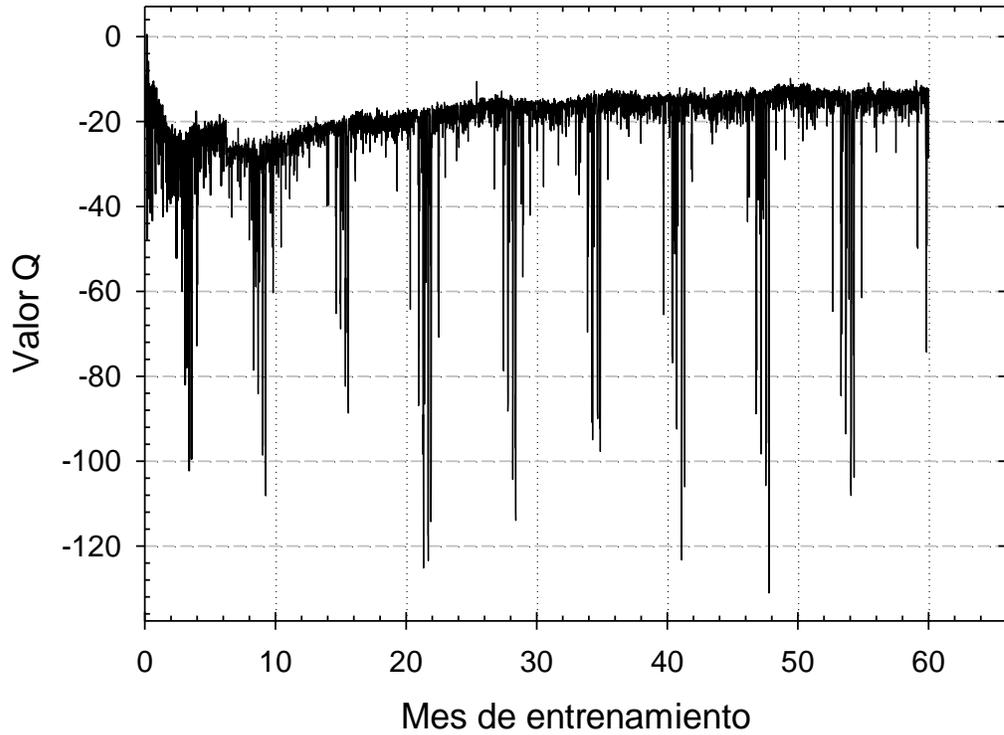
Nombre	Valor
Paso de simulación	20 minutos
Periodo de entrenamiento	Mayo a Octubre
Número de iteraciones	10
Archivo de clima	Taipéi

Fuente: elaboración propia.

El número de iteraciones corresponde a la cantidad de periodos de entrenamiento que fueron realizados en la simulación. Los diez periodos de entrenamiento equivalen a 60 meses de aprendizaje por parte del algoritmo. El desempeño del algoritmo a lo largo del periodo de entrenamiento en la simulación se visualiza en la figura 30.

⁴⁷ ENERGYPLUS. *Weather Data by Region*. https://energyplus.net/weather-region/asia_wmo_region_2/TWN%20. Consulta: 29 de agosto de 2018]

Figura 30. Entrenamiento del algoritmo



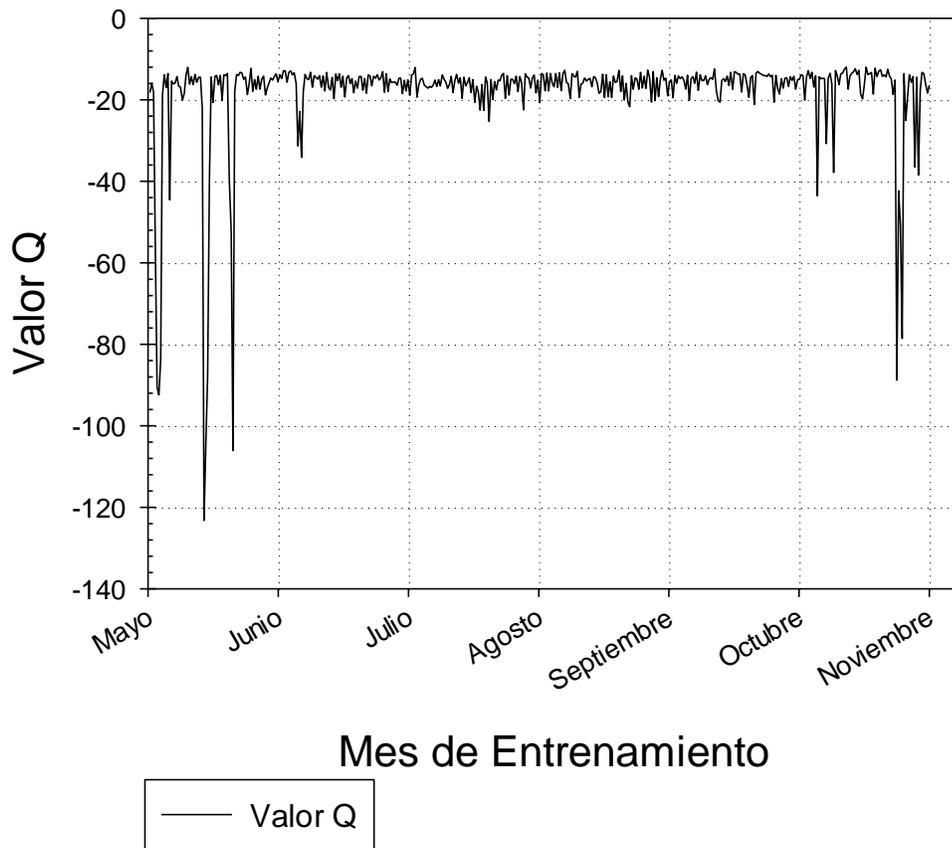
Fuente: elaboración propia, empleando SigmaPlot 14.

Inicialmente, los valores-Q de la red principal Q son muy bajos debido al alto nivel de exploración y poca explotación del algoritmo. No obstante, conforme pasa el tiempo, el agente aprende a predecir las mejores acciones y la explotación de las experiencias previas aumenta, siendo prueba de ello el aumento del valor-Q en los últimos meses.

Los picos que posee la gráfica equivalen específicamente a los primeros días de mayo y a los últimos días de octubre de cada periodo de entrenamiento, ya que en estos días la temperatura exterior es demasiado baja, por lo que cualquier acción que realice el algoritmo tendrá una baja recompensa. Lo

anterior se puede observar en la figura 31, donde se muestran los datos del octavo periodo de entrenamiento.

Figura 31. **Octavo periodo de entrenamiento**



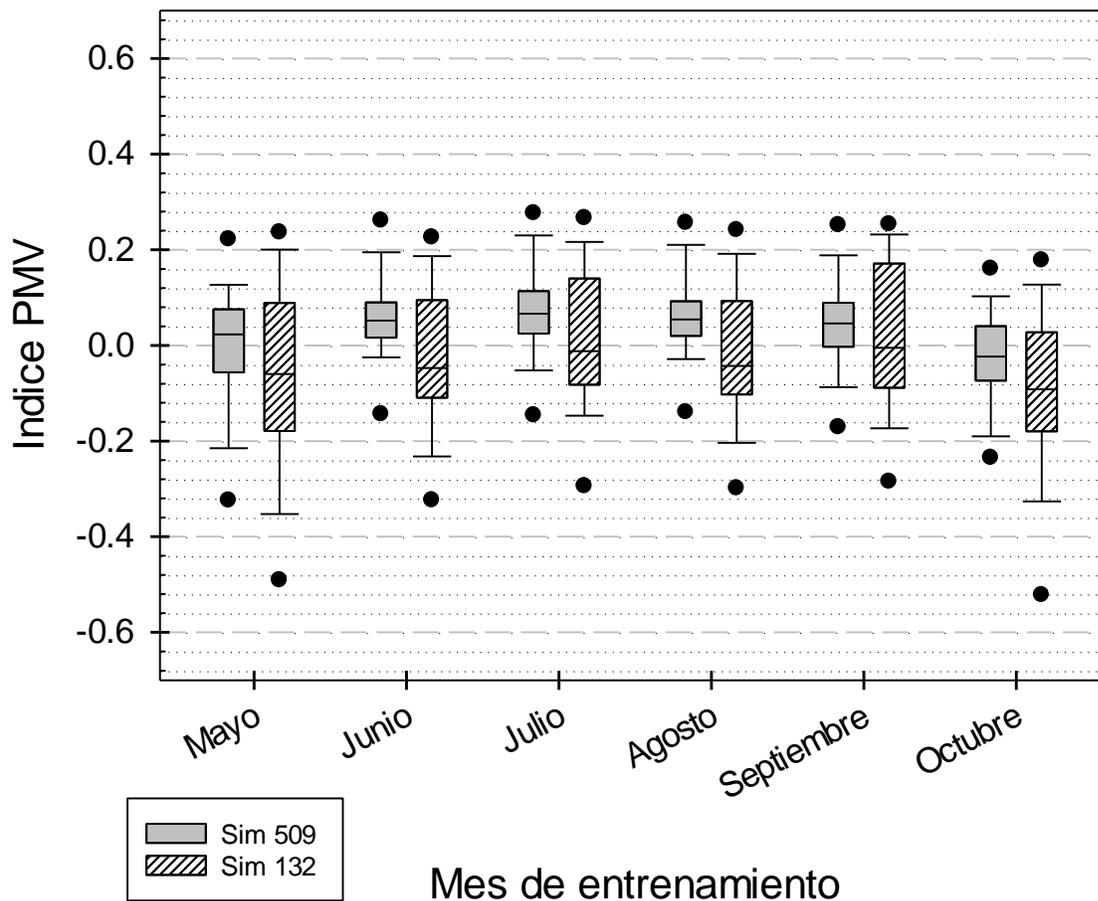
Fuente: elaboración propia, empleando SigmaPlot 14.

5.3. Pruebas y validación

Después de haber entrenado al algoritmo por diez periodos de entrenamiento, se analizó su comportamiento en un periodo de entrenamiento adicional para determinar el desempeño general del algoritmo después de

haber sido entrenado. Los resultados obtenidos se muestran en la figura 32 y figura 33.

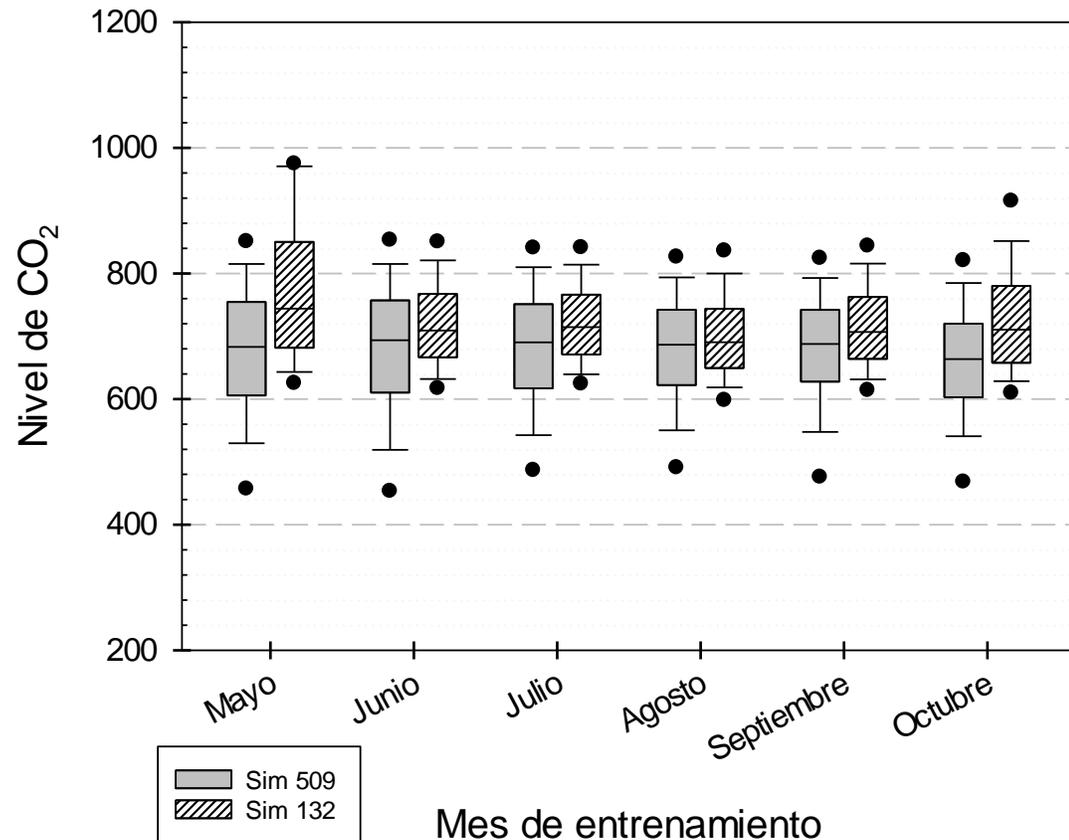
Figura 32. **PMV promedio en ambientes simulados**



Fuente: elaboración propia, empleando SigmaPlot 14

Como se puede observar en la figura 32, para ambos ambientes de simulación (509 y 132) el algoritmo puede controlar de forma correcta el índice PMV, ya que la mayor parte de los percentiles estuvieron alrededor de 0.

Figura 33. Nivel de CO₂ promedio en ambientes simulados



Fuente: elaboración propia, empleando SigmaPlot 14.

En el caso de la concentración de CO₂ promedio, los resultados también son satisfactorios, como puede observarse en la figura 33, donde en todo momento los niveles de CO₂ caen por debajo de los 800 ppm.

En resumen, los promedios de PMV y CO₂ son los siguientes: para la simulación del salón 132, el PMV promedio es de $-0,02259$ y un nivel de CO₂ promedio de 758,27 ppm, por otro lado, la simulación del laboratorio 509 tiene un PMV promedio de $+0,03409$ y un nivel de CO₂ de 676,45 ppm.

6. RESULTADOS EXPERIMENTALES Y DISCUSIÓN

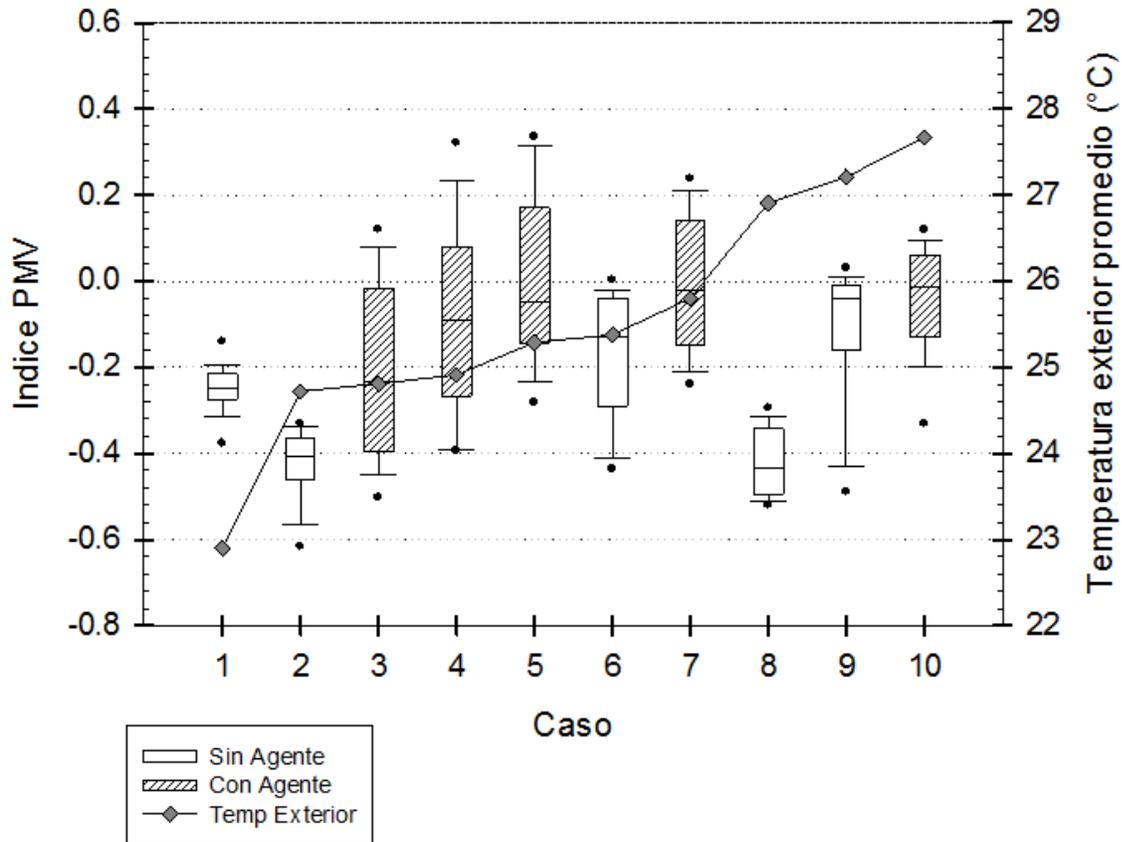
Una vez entrenado al agente, se realizaron pruebas reales en el laboratorio 509 y en el salón 132, cuyas configuraciones experimentales se detallan en la figura 15 y figura 16.

Aunque se realizaron pruebas en el salón 132, se tenía conciencia de que la configuración experimental estaba influenciada por demasiados factores fuera de nuestro control, como que el controlador de la unidad de AC estaba al alcance de cualquiera, lo que les permitía cambiar la temperatura manualmente y, en consecuencia, algunos de los resultados fueron corrompidos. Por lo tanto, solo los datos que estaban libres de interferencia se presentan en este estudio.

A pesar de esta limitación, los resultados de las pruebas con y sin el agente con respecto al índice de PMV y los niveles de CO₂ en el laboratorio 509 están disponibles y se muestran en la figura 34 y en la tabla XV. El total de diez casos experimentales se usan para comparación, y la tabla XII enumera la información resultante de los experimentos. Los casos 1, 2, 6, 8 y 9 están sin la presencia del agente, y el resto son operados por el agente. Todos los datos se recopilaron de octubre a noviembre de 2018, de 1:00 pm a 6:00 pm.

En resumen de la figura 34 y la tabla XV, el PMV promedio con el agente encendido fue de $-0,125$, y $-0,274$ con el agente apagado; el promedio de CO₂ fue de 689,5 ppm con el agente encendido y 1 039,1 ppm con el agente apagado. Estos resultados demuestran la capacidad del agente de controlar un ambiente, manteniendo un CO₂ más bajo y un PMV mejor, siempre dentro de los valores recomendados.

Figura 34. Valor PMV y temperatura exterior en el laboratorio 509



Fuente: elaboración propia, empleando SigmaPlot 14.

Tabla XV. Casos experimentales en el laboratorio 509 con y sin agente

Caso	Uso del agente	Personas promedio	$\overline{CO_2}$ (ppm)	\overline{PMV}	$\overline{T_{ext}}$ (°C)	Total E_{AC} (kWh)	Total $E_{ventilación}$ (kWh)
1	No	4	603,25	-0,24	22,90	4,27	0
2	No	5	900,47	-0,42	24,72	4,07	0
3	Sí	6	671,03	-0,21	24,81	3,39	0,07
4	Sí	5	686,02	-0,09	24,91	4,29	0,04
5	Sí	8	630,52	+0,01	25,29	5,46	0,05
6	No	8	1 407,75	-0,17	25,37	3,50	0
7	Sí	6	693,16	-0,28	25,79	4,09	0,08

Continuación de la tabla XV.

8	No	7	1 202,65	-0,42	26,91	5,32	0
9	No	6	1 081,33	-0,10	27,20	3,97	0
10	Sí	5	766,72	-0,04	27,67	4,92	0,09

Fuente: elaboración propia.

Tabla XVI. **Resultados resumidos obtenidos en el laboratorio 509**

Presencia del Agente DRL	Promedio Total de PMV	Promedio Total de CO ₂ (ppm)	Promedio Total de E_{AC} (kWh)	Promedio Total de $E_{ventilación}$ (kWh)
Sí	-0,125	689,5	4,43	0,07
No	-0,274	1 039,1	4,26	0

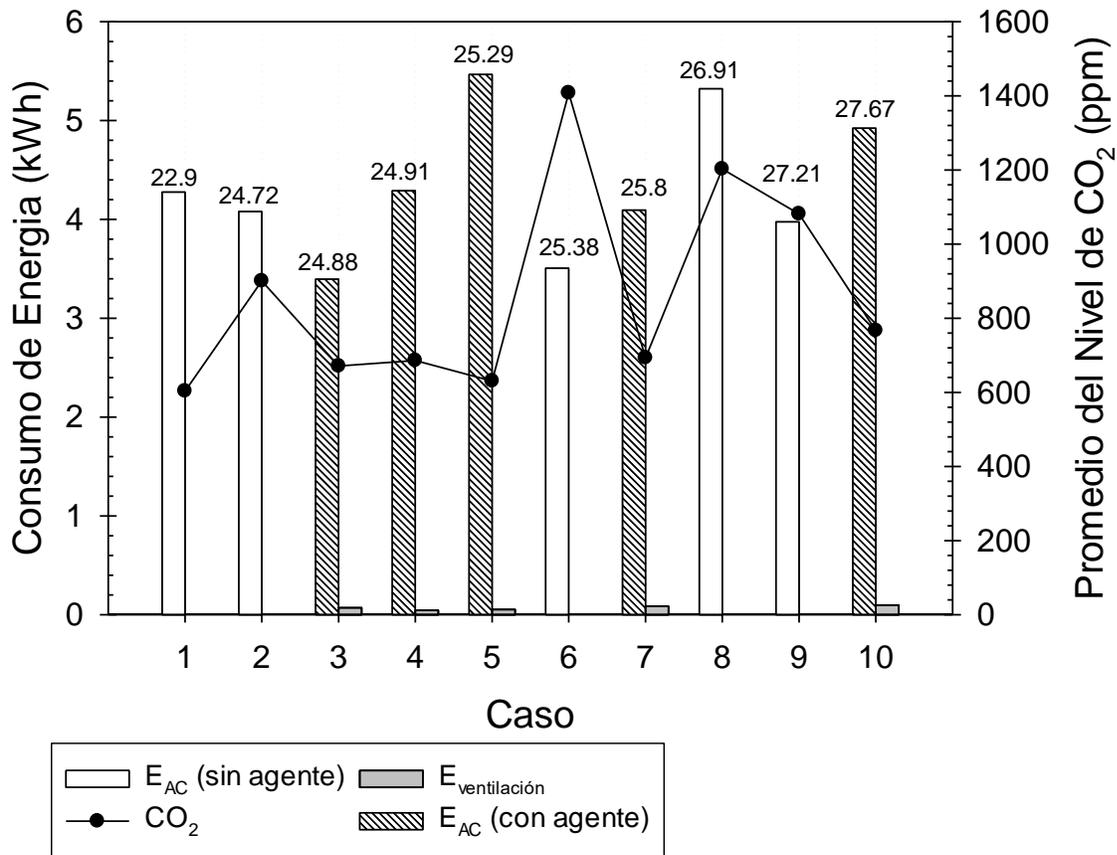
Fuente: elaboración propia.

Comparando el caso 2 y el caso 4 de la tabla XV, el consumo de energía es ligeramente mayor (alrededor del 5 %), mientras que el nivel de CO₂ es 220 ppm más bajo, en ambas situaciones con el agente encendido. Para el caso 3 y el caso 6, con aproximadamente la misma temperatura exterior, el consumo de energía para el caso 3 también es ligeramente más bajo que con el agente apagado, mientras que el nivel de CO₂ disminuyó considerablemente. Esto se debe, principalmente, a la diferencia en la cantidad de los ocupantes presentes cuando se efectuaron los experimentos.

El consumo energético de AC correspondiente durante estos períodos se muestra en la figura 35, donde se puede apreciar que el consumo de energía se incrementa considerablemente con el aumento de la temperatura exterior y, que consumió un poco más de energía para controlar el ambiente con el agente encendido, comparando días con similares temperaturas externas promedio.

Observar que el número sobre las barras es la temperatura exterior correspondiente.

Figura 35. Consumo energético y niveles de CO₂ en el laboratorio 509



Fuente: elaboración propia, empleando SigmaPlot 14.

Los resultados previos demuestran que el control de CO₂ aumenta la carga del sistema de AC al introducir aire del exterior. Para obtener una estimación más detallada sobre su influencia en el consumo de energía, se realizó una simulación de 6 meses del laboratorio 509. El consumo de energía resultante sujeto a varios controles de CO₂ se muestra en la figura 36, notar que SC significa sistema de control.

Se simularon siete casos con una temperatura fija de 25 °C. El caso de referencia es sin la presencia del agente y sin control de CO₂ (denominado AC), y se hacen cuatro casos adicionales similares; pero sujetos al control de CO₂, encendiendo los ventiladores cuando el límite de nivel de CO₂ en el ambiente supera los 900 ppm (SC_900), 808 ppm (SC_808), 737 ppm (SC_737) y 700 ppm (SC_700), respectivamente.

Los otros dos casos corresponden a simulaciones donde se utilizaron dos agentes DRL (Agente-1 y Agente-2) para las comparaciones de los cinco casos anteriores con el control inteligente. Los dos agentes utilizan dos penalizaciones de peso diferentes para controlar la ventilación, estas penalizaciones se detallan en la tabla XVII.

Tabla XVII. **Comparación de pesos entre Agente-1 y Agente-2**

Agente-1		Agente-2	
w_P	1	w_P	1
w_C	1	w_C	1
w_A	$ R_P(0,4) \approx 0,64$	w_A	$ R_P(0,4) \approx 0,64$
w_V	$ R_C(900) \approx 0,60$	w_V	$ R_C(950) \approx 0,80$

Fuente: elaboración propia.

Tabla XVIII. **Configuración de los casos de prueba simulados**

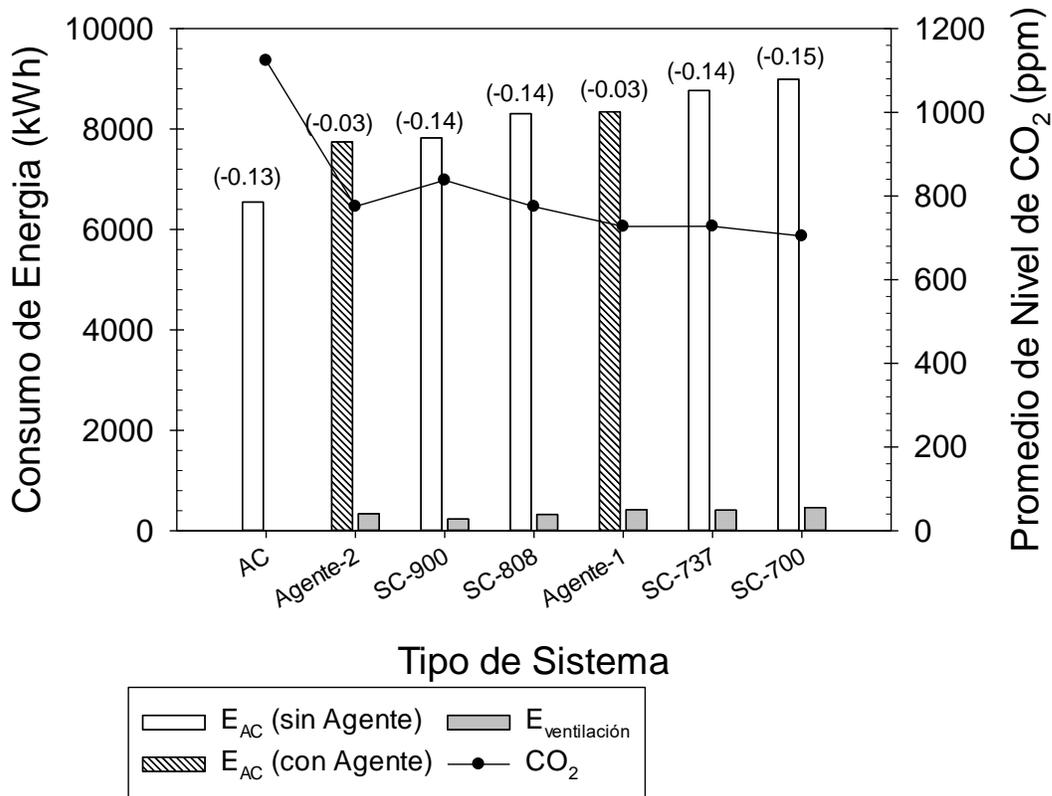
Nombre del Caso	Presencia del Agente DRL	Límite de CO ₂ (ppm)	Configuración de AC (°C)
AC	No	Sin control	25
CS-700	No	700	25
CS-737	No	737	25
CS-808	No	808	25
CS-900	No	900	25

Continuación de la tabla XVIII.

Agente-1	Sí	Automático	Variable
Agente-2	Sí	Automático	Variable

Fuente: elaboración propia.

Figura 36. Consumo energético de AC para diferentes casos



Fuente: elaboración propia, empleando SigmaPlot 14.

Como se puede observar en la figura 36, el aumento del consumo energético del sistema de AC, con agente y sin él, no es tan pronunciado como se esperaba, y este consumo está directamente relacionado con los niveles de CO₂ cuando el agente no está activado. Cabe destacar que el caso de

referencia muestra el menor consumo de energía, pero se acompaña con un nivel inaceptable de CO₂ de 1 150 ppm y un PMV de -0,13 comparativamente bajo.

En cambio, con el control inteligente, tanto los niveles de PMV como de CO₂ fueron mejores y se logra sin una penalización adicional del consumo de energía. Esto puede quedar claro para la comparación entre el Agente-1 y el SC_737, donde el nivel de CO₂ controlado es el mismo (~ 737 ppm), pero el PMV correspondiente para el caso Agente-1 es -0,03, y -0,14 para el caso SC_737.

No obstante, el consumo de energía con el control del Agente-1 es aproximadamente un 4 % a 5 % más bajo que el de SC_737. Similarmente, para el caso Agente-2 y SC_808 donde el nivel controlado de CO₂ es el mismo; pero el consumo de energía para el caso Agente-2 es un 4 % a 6 % menos. Por otro lado, para el mismo consumo de energía, tomando el caso Agente-1 y SC_808 como ejemplo, el PMV correspondiente con Agente-1 es -0,03 mientras que es -0,14 sin control inteligente. Sin embargo, el nivel correspondiente de CO₂ con control inteligente es aproximadamente un 10 % más bajo.

En resumen, los agentes propuestos muestran un mejor rendimiento que el sistema de AC convencional con y sin control del nivel de CO₂, lo cual indica que el agente DRL puede equilibrar las necesidades de confort térmico, calidad del aire interior CO₂ y consumo de energía.

CONCLUSIONES

1. El consumo de energía eléctrica logrado es de un 4 % a 5 % más bajo que los demás casos cuando no se utilizó el algoritmo DRL, a pesar de que es un pequeño ahorro de energía, los valores de PMV y nivel de CO₂ tienen una mejora considerable.
2. La implementación del algoritmo DRL propuesto puede controlar y equilibrar las necesidades de confort térmico, según los estándares de ANSI/ASHRAE 55 e ISO 7730, ya que el valor de PMV estuvo dentro del rango de $[-0,28, +0,01]$.
3. Los resultados del nivel de concentración de CO₂ en el ambiente real también son satisfactorios, ya que los niveles promedios de CO₂ decaen a menos de 800 ppm cuando se usa la implementación del algoritmo DRL.
4. En lo que respecta al control del índice PMV y el nivel de CO₂ en el ambiente de simulación, la implementación del algoritmo logra un consumo menor de energía que el realizado por un control separado de ambas magnitudes, en el que la configuración del aire acondicionado es constante y el sistema de ventilación es controlado por un sistema de control.
5. La simulación de ambientes en EnergyPlus fue una parte importante de la investigación, reduciendo drásticamente el tiempo de entrenamiento y logrando que el experimento fuera viable.

RECOMENDACIONES

1. Utilizando una configuración distinta de hiperparámetros a la presentada en este trabajo de graduación, es posible crear un algoritmo más eficiente.
2. Si se desea utilizar un solo agente para múltiples ambientes, es necesario agregar las características que hacen diferente a los ambientes dentro del estado del algoritmo, por ejemplo, el área, la cantidad de unidades AC, cantidad de ventiladores, entre otros.
3. Entre más variedad de ambientes simulados son utilizados para el entrenamiento del agente, más viable se vuelve la posibilidad de hacer un paquete de instalación estándar para distribuirlo comercialmente como un software que sea compatible con sistemas típicos de ventilación y AC.

BIBLIOGRAFÍA

1. ABADI, M. *TensorFlow: a system for large-scale machine learning*. EE.UU.: Advanced Computing Systems Association (USENIX), 2016. 21 p.
2. AL-FUQAHA, A. *Internet of things: A survey on enabling technologies, protocols, and applications*. IEEE Communications Surveys Tutorials, 2015. 29 p. ISBN 1553-877X.
3. ARNABAT, I. *¿Cómo funciona el aire acondicionado? Infografía*. [en línea] <<https://www.caloryfrio.com/aire-acondicionado/como-funciona-el-aire-acondicionado-infografia.html>> [Consulta: 12 de enero de 2019].
4. ASHRAE. *Standard 55-2010 Thermal environmental conditions for human occupancy*. American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2010. 36 p.
5. ASHRAE. *Standard 62.1-2010 Ventilation for acceptable indoor air quality*. American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2010. 50 p.
6. AUSTRALIA, W. *Adopted National Exposure Standards for Atmospheric Contaminants in the Occupational Environment*. [en línea] <<https://www.safeworkaustralia.gov.au/doc/adopted-national-expos>

ure-standards-atmospheric-contaminants-occupational-environment-nohsc-1003> [Consulta: 22 de noviembre de 2018].

7. BODUCH, M. and FINCHER, W. *Standards of human comfort: relative and absolute*. EE. UU.: Universidad de Texas, 2009. 12 p.
8. BONINO, S. *Carbon Dioxide Detection and Indoor Air Quality Control*. [en línea] <<https://ohsonline.com/articles/2016/04/01/carbon-dioxide-detection-and-indoor-air-quality-control.aspx>> [Consulta: 16 de enero de 2019]
9. CEN. *15251: 2007 Indoor environmental input parameters for design assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics*. European Committee for Standardization, 2007. 52 p.
10. CRAWLEY, D. *EnergyPlus: creating a new-generation building energy simulation program*. Energy and Buildings, 2001. 12 p. ISBN 0378-7788.
11. DE LA LENGUA ESPAÑOLA, R. A. *Diccionario de la lengua española*. 23a ed. España, 2014. 4432 p.
12. EDUCATION, D. o. *Building Bulletin 101 Guidelines on ventilation, thermal comfort and indoor air quality in schools*. [en línea] <<https://www.gov.uk/government/publications/building-bulletin-101-ventilation-for-school-buildings>> [Consulta: 23 de noviembre de 2018].

13. ENERGYPLUS. *Weather Data by Region*. [en línea] <https://energyplus.net/weather-region/asia_wmo_region_2/TWN%20%20> [Consulta: 17 de agosto de 2018].
14. FANGER, P. O. *Thermal comfort. Analysis and applications in environmental engineering*. 1970. 244 p.
15. FERRY, T. *The Number-one Concern: Electric Power*. Taiwán: Taiwan Business TOPICS, 2018. 46 p.
16. FREEDMAN, Y. and ZEMANSKY, S. *Física Universitaria*. 12a ed. México: Addison-Wesley, 2009. 763 p.
17. GOOGLE. *SketchUp*. [En línea] <<https://www.sketchup.com>> [Consulta: 12 de febrero 2019].
18. GUGLIELMETTI, R. *OpenStudio: an open source integrated analysis platform*. Australia: National Renewable Energy Laboratory (NREL), 2011. 18 p.
19. HASSELT, H. *Deep reinforcement learning with double Q-Learning*. EE. UU.: Association for the Advancement of Artificial Intelligence (AAAI) Press, 2016. 7 p.
20. HWANG, R. *Comfort Temperature and Preferred Temperature in Taiwan*. Singapur: Springer, 2018. 8 p. ISBN 978-981-10-8465-2.
21. HWANG, R. *Occupants' Behavior in Taiwan*. Singapur: Springer, 2018. 35 p. ISBN 978-981-10-8465-2.

22. ISO. *Standard 7730:2005 Ergonomics of the thermal environment*. International Organization for Standardization, 2005. 11 p.
23. KELLY. *Taiwan Weather, Taiwan Climate in Spring, Summer, Autumn and Winter*. [en línea] <<https://www.chinahighlights.com/taiwan/weather.htm>> [Consulta: 23 de octubre de 2018].
24. KUO-LIANG, L. *Energy Consumption Analysis for Concrete Residences—A Baseline Study in Taiwan*. Multidisciplinary Digital Publishing Institute, 2017. 13 p.
25. LECUN, Y. *Deep learning*. Nature, 2015. 15 p. ISBN 1476-4687.
26. LIAO, S. *The Non-Linear Relationship between Electricity Consumption and Temperature in Taiwan: An Application for STR (Smooth Transition Regression) Model*. Scientific Research Publishing, 2018. 18 p.
27. MADAKAM, S. *Internet of Things (IoT): A literature review*. Journal of Computer Communications, 2015. 164 p.
28. MNIH, V. *Playing atari with deep reinforcement learning*. Arxiv Preprint, 2013. 9 p.
29. MYHRVOLD, A. N. *Indoor environment in schools—pupils health and performance in regard to CO2 concentrations*. Indoor Air, 1996. 374 p.

30. NG, A. *Neural Networks and Deep Learning, Deep Learning Specialization*. [en línea] <<https://www.coursera.org/learn/neural-networks-deep-learning>> [Consulta: 24 de noviembre de 2018]
31. OSHA. *NIOSH 1988 OSHA PEL Project Documentation*. [en línea] <<https://www.cdc.gov/niosh/pel88/124-38.html>> [Consulta: 13 de septiembre de 2018]
32. OTTERLO, M., WIERING, M. *Reinforcement Learning and Markov Decision Processes*. Alemania: Springer, 2012. 50 p.
33. RUSSELL, S., NORVIG, P. *Inteligencia Artificial: Un Enfoque Moderno*. España: Pearson Educación, 2004. 65 p. ISBN 842054003X.
34. SATISH, U. *Is CO2 an indoor pollutant? Direct effects of low-to-moderate CO2 concentrations on human decision-making performance*. *Environmental Health Perspectives*, 2012. 1890 p.
35. SIMONINI, T. *A free course in Deep Reinforcement Learning from beginner to expert*. [en línea] <https://simoninithomas.github.io/Deep_reinforcement_learning_Course/> [Consulta: 14 de agosto de 2018]
36. WATKINS, C., DAYAN, P. *Q-learning in Machine Learning*. 1992. 320 p.
37. WEI, T. *Deep Reinforcement Learning for Building HVAC Control*. Association for Computing Machinery (ACM), 2017. 11 p.

38. WETTER, M. *Co-simulation of building energy and control systems with the Building Controls Virtual Test Bed*. Taylor & Francis, 2011. 250 p. ISBN 1940-1493.

39. YUAN, E. *Environmental Protection Agency of the Executive Yuan*. [en línea] <<https://oaout.epa.gov.tw/law/LawContent.aspx?id=FL068252>> [Consulta: 24 de noviembre de 2018].

APÉNDICES

Apéndice 1. **Salón de clases 132, edificio 5, NCTU, Taiwán**



Fuente: elaboración propia.

Apéndice 2. Laboratorio 509, edificio 5, NCTU, Taiwán



Fuente: elaboración propia.