

Pemodelan Karakteristik Penggunaan Air Conditioner (Ac) Menggunakan Metode Reinforcement Learning Untuk Meningkatkan Efisiensi Penggunaan Listrik

1st Muhammad Idris Syafrizal

Fakultas Teknik Elektro

Telkom University

Bandung, Indonesia

idrissyafrialm@student.telkomuniversity.ac.id

2nd Muhammad Ary Murti

Fakultas Teknik Elektro

Telkom University

Bandung, Indonesia

arymurti@telkomuniversity.ac.id

3rd Junartha Halomoan

Fakultas Teknik Elektro

Telkom University

Bandung, Indonesia

junartha@telkomuniversity.ac.id

Penggunaan Air Conditioner (AC) yang tidak terkelola secara optimal menjadi salah satu penyebab utama tingginya konsumsi energi listrik pada gedung perkuliahan. Penelitian ini bertujuan untuk memodelkan karakteristik penggunaan AC serta mengevaluasi potensi peningkatan efisiensi energi melalui pendekatan simulasi berbasis Reinforcement Learning, khususnya algoritma Soft Actor-Critic (SAC). Penelitian difokuskan pada simulasi pengendalian suhu AC menggunakan data operasional historis, tanpa penerapan kontrol secara real-time pada perangkat fisik. Data suhu ruangan, suhu luar, setpoint AC, dan konsumsi daya diperoleh dari sistem berbasis Internet of Things (IoT) yang terdiri dari sensor DHT22, kWh meter DDS238-2 ZN/S, dan mikrokontroler ESP32. Data tersebut digunakan untuk membangun model dinamika AC sebagai lingkungan simulasi. Selanjutnya, algoritma Soft Actor-Critic dilatih secara offline untuk mempelajari kebijakan pengaturan setpoint suhu yang mampu menyeimbangkan antara kenyamanan termal dan efisiensi energi. Hasil simulasi menunjukkan bahwa algoritma Soft Actor-Critic mampu menghasilkan strategi pengendalian yang lebih stabil dibandingkan pengaturan manual. Rata-rata error kenyamanan suhu pada simulasi berbasis SAC tercatat sebesar 0,25°C, lebih rendah dibandingkan metode manual yang mencapai 1,46°C. Dari sisi konsumsi energi, simulasi menunjukkan potensi penghematan energi sebesar 10–17%, dengan penurunan konsumsi dari 1,606 kWh menjadi 1,326 kWh pada skenario pengujian selama 10 jam. Berdasarkan hasil tersebut, dapat disimpulkan bahwa pendekatan simulasi berbasis Reinforcement Learning dengan algoritma Soft Actor-Critic efektif untuk memodelkan karakteristik penggunaan AC dan mengevaluasi potensi efisiensi energi. Penelitian ini dapat menjadi dasar pengembangan sistem pengendalian AC yang lebih cerdas dan adaptif pada tahap implementasi selanjutnya.

Kata kunci: Air Conditioner, Simulasi, Reinforcement Learning, Soft Actor-Critic, Efisiensi Energi.

I. PENDAHULUAN

Dalam era modern, kehidupan manusia tidak dapat dipisahkan dari penggunaan peralatan elektronik, khususnya Air Conditioner (AC). Di Indonesia, penggunaan AC tidak hanya terbatas pada rumah tangga, tetapi juga digunakan secara luas pada gedung perkantoran dan institusi pendidikan [1]. Penggunaan AC dengan pengaturan suhu yang tepat terbukti mampu meningkatkan kenyamanan termal serta produktivitas kerja, karena kondisi lingkungan yang sejuk dan stabil berpengaruh langsung terhadap kinerja dan konsentrasi pengguna [2].

Berdasarkan data Kementerian Energi dan Sumber Daya Mineral (ESDM) tahun 2023, sektor bangunan dan rumah tangga menyumbang sekitar 37% konsumsi listrik nasional,

di mana lebih dari 45% di antaranya digunakan oleh sistem pendingin udara (AC) [3]. Hal ini menunjukkan bahwa sistem pengendalian AC memiliki peranan yang sangat signifikan terhadap efisiensi energi nasional. Di lingkungan Universitas Telkom, hampir seluruh ruang kelas dan kantor telah dilengkapi dengan AC, bahkan beberapa ruangan menggunakan lebih dari satu unit. Meskipun hal ini meningkatkan kenyamanan pengguna, hingga saat ini belum tersedia sistem yang mampu memantau dan mengendalikan konsumsi energi AC secara terintegrasi, sehingga penggunaan energi menjadi tidak terkontrol dan berpotensi meningkatkan biaya operasional listrik.

Berbagai penelitian sebelumnya telah mencoba mengoptimalkan penggunaan AC menggunakan metode machine learning. Yu et al. menerapkan metode Decision Tree untuk pemodelan konsumsi energi bangunan karena memiliki waktu komputasi yang relatif cepat, namun metode ini memiliki keterbatasan dalam menangani data numerik kontinu secara akurat [4]. Pendekatan Artificial Neural Network (ANN) juga banyak digunakan karena kemampuannya dalam memodelkan hubungan non-linear antara variabel lingkungan dan konsumsi energi, tetapi membutuhkan data pelatihan yang besar serta

waktu komputasi yang panjang [5]. Selain itu, sebagian besar penelitian tersebut masih berfokus pada prediksi konsumsi energi, bukan pada pengendalian AC secara adaptif dan otomatis [6].

Dalam beberapa tahun terakhir, pendekatan Reinforcement Learning (RL) mulai banyak diterapkan pada sistem Heating, Ventilation, and Air Conditioning (HVAC). Penelitian oleh Al Sayed et al. menunjukkan bahwa metode RL mampu mengurangi konsumsi energi HVAC secara signifikan dibandingkan metode kontrol konvensional, karena agen RL dapat belajar langsung dari interaksi dengan lingkungan tanpa memerlukan model matematis sistem yang kompleks [7]. Penelitian lain oleh Sun et al. membuktikan bahwa algoritma Soft Actor-Critic (SAC) mampu menjaga stabilitas suhu ruangan sekaligus meningkatkan efisiensi energi, berkat kemampuannya dalam menangani ruang aksi kontinu serta mekanisme eksplorasi kebijakan yang lebih stabil [8].

Meskipun memiliki keunggulan tersebut, penerapan Reinforcement Learning juga memiliki beberapa keterbatasan. Proses pembelajaran RL membutuhkan data dalam jumlah besar serta waktu pelatihan yang relatif panjang

[9]. Selain itu, pada fase awal pembelajaran, agen RL dapat menghasilkan aksi yang kurang optimal akibat proses eksplorasi, sehingga berpotensi menurunkan kenyamanan termal apabila tidak dilengkapi dengan mekanisme pembatasan atau pengamanan yang sesuai [10].

Berdasarkan kelebihan dan keterbatasan tersebut, penelitian ini mengadopsi algoritma Soft Actor-Critic (SAC) karena memiliki stabilitas pembelajaran yang lebih baik dibandingkan algoritma RL lainnya serta mampu menangani pengaturan suhu AC yang bersifat kontinu [8], [11]. Algoritma SAC kemudian diintegrasikan dengan sistem Internet of Things (IoT) untuk melakukan pengambilan data suhu dan konsumsi daya secara real-time. Data tersebut digunakan untuk memodelkan karakteristik penggunaan AC secara dinamis dan menghasilkan strategi pengendalian yang adaptif. Dengan pendekatan ini, diharapkan sistem yang dikembangkan mampu meningkatkan efisiensi energi, menurunkan konsumsi listrik, serta tetap menjaga kenyamanan termal pengguna, sehingga mendukung implementasi konsep smart campus yang berkelanjutan.

II. KAJIAN TEORI

A. Cara Kerja AC Split

Air Conditioner (AC) split bekerja berdasarkan siklus kompresi uap (Vapor Compression Cycle) yang terdiri dari empat komponen utama: evaporator, kompresor, kondensor, dan katup ekspansi. Sistem ini terbagi menjadi dua bagian yaitu unit indoor dan unit outdoor.

1. Evaporator

Udara panas dalam ruangan melewati pipa berisi refrigeran bertekanan rendah. Panas diserap oleh refrigeran, sehingga udara keluar menjadi lebih dingin dan kering.

2. Kompresor

Gas refrigeran bertekanan rendah dikompresi menjadi gas bersuhu dan bertekanan tinggi agar siap membuang panas ke lingkungan.

3. Kondensor

Gas panas dilewatkan ke kondensor dan melepaskan panas ke udara luar. Setelah itu, refrigeran berubah menjadi cair bertekanan tinggi.

4. Katup Ekspansi

Refrigeran cair bertekanan tinggi melewati katup ekspansi, tekanannya turun drastis sehingga suhunya menjadi rendah sebelum kembali ke evaporator untuk menyerap panas lagi.

B. Kenyamanan Thermal

Kenyamanan termal menurut *Szokolay* (1973) pada 'Manual of Tropical Housing and Building' merupakan proses yang melibatkan kondisi fisik fisiologis dan psikologis. Kenyamanan termal adalah hasil pemikiran seseorang yang mengekspresikan mengenai kepuasan dirinya terhadap lingkungan termalnya. *ASHRAE (American Society of Heating Refrigerating Air Conditioning Engineer)* mendefinisikan kenyamanan termal sebagai suatu kondisi dimana ada kepuasan terhadap keadaan termal di sekitarnya. Sedangkan kenyamanan termal menurut *Snyder* (1989) merupakan keadaan lingkungan/alam yang dapat mempengaruhi manusia. Dari pernyataan tersebut dapat dinyatakan bahwa kenyamanan termal merupakan rumusan empirik yang merupakan sebuah pengalaman terhadap rasa

dimana kondisi yang dirasakan dapat berbeda antara satu orang dengan yang lainnya. Dalam menentukan kenyamanan sebuah zona / area, dapat dinyatakan dengan melihat persepsi dominan yang dirasakan oleh sekelompok sampel pada area tersebut. Menurut rekomendasi *ASHRAE* kenyamanan dapat diperoleh dalam rentang suhu $23^{\circ}\text{C} - 26^{\circ}\text{C}$ dengan kelembaban $50\% - 70\%$

TABEL 1
Jenis Kenyamanan.

Kondisi	Temperatur Efektif
Sejuk nyaman	$20,5^{\circ}\text{C} - 22,8^{\circ}\text{C}$
Nyaman optimal	$22,8^{\circ}\text{C} - 25,8^{\circ}\text{C}$
Hangat nyaman	$25,8^{\circ}\text{C} - 27,1^{\circ}\text{C}$

Berikut, tingkat kelembaban yang disarankan berada pada rentang $40\% - 50\%$ dalam kondisi ruangan yang normal. Apabila ruangnya memiliki jumlah orang yang dianggap padat, maka kelembaban yang dianjurkan berkisar antara $55\% - 60\%$. Perhitungan kebutuhan AC dalam ruangan dengan metode BTU membutuhkan data berupa panjang, lebar, tinggi ruangan dalam satuan feet ($1\text{ m} = 3,28\text{ feet}$), besaran I dengan melihat apakah ruangan tersebut berada di lantai atas atau di lantai bawah dan barat, selatan, ataupun timur. Berikut merupakan contoh perhitungan kebutuhan AC pada ruangan Spheroidene dengan menggunakan metode BTU: Terdapat beberapa jenis daya AC yang dipasarkan secara umum antara lain seperti pada Tabel

C. Reinforcement Learning

Reinforcement Learning merupakan salah satu cabang Machine Learning yang berfokus pada proses pembelajaran melalui interaksi antara agen dan lingkungan untuk mencapai suatu tujuan tertentu. Tidak seperti supervised learning yang menggunakan data berlabel, RL belajar berdasarkan pengalaman dengan menerima umpan balik berupa reward (hadiah) atau punishment (hukuman) dari tindakan yang dilakukan. Dalam konteks penelitian ini, agent adalah model pembelajaran (policy) yang menentukan setpoint suhu AC, sedangkan environment adalah ruangan yang dipengaruhi oleh perubahan suhu, kelembaban, dan konsumsi daya listrik AC. Tujuan dari agent adalah meminimalkan penggunaan energi listrik sekaligus menjaga kenyamanan termal pengguna.[1]

Sistem RL dapat dimodelkan dengan Markov Decision Process

(MDP) yang terdiri atas:

- *State (s)*: kondisi sistem pada waktu tertentu (suhu ruangan, kelembaban, dan daya listrik).
- *Action (a)*: tindakan yang dapat dilakukan agent (misal: menaikkan/menurunkan setpoint suhu).
- *Reward (r)*: umpan balik dari lingkungan terhadap aksi yang diambil (misal: efisiensi energi meningkat \rightarrow reward positif).
- *Policy (π)*: strategi agent untuk menentukan aksi berdasarkan keadaan saat ini.
- *Value Function (v)*: nilai harapan dari total reward masa depan.

Persamaan Bellman untuk fungsi nilai didefinisikan sebagai berikut:

$$V\pi = (st) = E_{\pi}[rt + \gamma V\pi(st + 1)] \quad (2)$$

di mana $\gamma \in [0,1]$ adalah faktor diskonto yang menentukan seberapa besar pengaruh reward di masa depan terhadap keputusan saat ini.

D. Mekanisme Reinforcement Learning

Secara umum, proses pembelajaran Reinforcement Learning terbagi menjadi dua fase

1. Eksplorasi (Exploration): agent mencoba berbagai aksi untuk mengenal lingkungan.
 2. Eksploitasi (Exploitation): agent menggunakan pengalaman terbaik untuk memaksimalkan reward.
- Dalam implementasinya, Reinforcement Learning dapat dikategorikan menjadi beberapa pendekatan:

- Value-based: seperti Q-Learning dan Deep Q-Network (DQN), yang berfokus mempelajari nilai aksi terbaik.
- Policy-based: seperti REINFORCE, yang langsung mempelajari kebijakan tanpa menghitung nilai Q.

- Actor-Critic: kombinasi keduanya; Actor menentukan aksi, sedangkan Critic mengevaluasi aksi tersebut. Pendekatan Actor-Critic digunakan dalam penelitian ini karena mampu menangani action space kontinu, yang sesuai dengan karakteristik suhu Air Conditioner yang tidak diskret.

E. Soft Actor-Critic

Soft Actor-Critic merupakan algoritma Reinforcement Learning berbasis off-policy actor-critic yang dirancang untuk aksi kontinu. Soft Actor-Critic dikembangkan untuk memaksimalkan dua hal secara bersamaan:

- Nilai reward jangka panjang return
- Entropi kebijakan, yaitu tingkat acaknya keputusan agen agar tetap mampu mengeksplorasi lingkungan.

Dengan demikian, SAC lebih stabil dibandingkan algoritma Reinforcement Learning lain seperti DDPG dan PPO karena tidak mudah terjebak pada solusi lokal (local optimum).

a) Return (Reward Kumulatif)

Return merupakan total reward yang diperoleh agen sejak waktu tertentu hingga akhir episode, dengan mempertimbangkan faktor diskonto.

$$G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k} \quad (3)$$

G_t : return (reward kumulatif)

r_t : reward pada waktu ke- t

γ : discount factor ($0 < \gamma < 1$)

Penjelasan:

Rumus ini digunakan untuk mendefinisikan tujuan utama Reinforcement Learning, yaitu memaksimalkan total reward jangka panjang. Faktor diskonto γ digunakan untuk menyeimbangkan antara reward saat ini dan reward di masa depan.

b.) Fungsi Nilai Aksi (Action-Value Function)

Fungsi nilai aksi menyatakan ekspektasi return ketika agen berada pada state tertentu dan mengambil aksi tertentu.

$$Q^{\pi}(s_t, a_t) = \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k} \mid s_t, a_t \right] \quad (4)$$

- s_t : state lingkungan
- a_t : aksi yang diambil agen
- π : kebijakan (policy)

Penjelasan:

Fungsi ini digunakan oleh Critic Network untuk mengevaluasi seberapa baik suatu aksi pengendalian AC pada kondisi suhu tertentu.

c.) Fungsi Reward

Reward dirancang untuk mencerminkan tujuan pengendalian, yaitu menjaga kenyamanan suhu dan meminimalkan konsumsi energi.

$$r_t = -\alpha |T_t - T_{set}| - \beta P_t \quad (5)$$

- T_t : suhu ruangan saat t
- T_{set} : suhu target
- P_t : daya listrik AC
- α, β : bobot penalti

Penjelasan:

Reward bernilai negatif jika terjadi deviasi suhu dari setpoint atau konsumsi daya meningkat. Dengan desain ini, agen didorong untuk menjaga suhu tetap nyaman sekaligus hemat energi.

d.) Fungsi Objektif Soft Actor-Critic

Berbeda dengan RL konvensional, SAC memaksimalkan reward dan entropi kebijakan.

$$J(\pi) = \sum_t \mathbb{E} [r(s_t, a_t) + \alpha \mathcal{H}(\pi(\cdot | s_t))] \quad (6)$$

- \mathcal{H} : entropi kebijakan
- α : koefisien entropi

Penjelasan:

Penambahan entropi bertujuan untuk menjaga eksplorasi agar agen tidak cepat terjebak pada solusi lokal, sehingga pembelajaran menjadi lebih stabil.

e.) Target Q-Value pada SAC

Target nilai Q digunakan dalam proses pembaruan jaringan Critic.

$$y_t = r_t + \gamma \left(\min_{i=1,2} Q_{\theta_i}'(s_{t+1}, a_{t+1}) - \alpha \log \pi(a_{t+1} | s_{t+1}) \right) \quad (7)$$

Penjelasan:

Penggunaan dua Critic Network dan operator minimum bertujuan untuk mengurangi bias estimasi nilai Q (*overestimation bias*).

f.) Fungsi Loss Critic Network

$$L_Q(\theta) = \mathbb{E}[(Q_{\theta}(s_t, a_t) - y_t)^2] \quad (8)$$

Penjelasan:

Loss ini digunakan untuk melatih Critic Network agar estimasi nilai Q mendekati target Q-value.

g.) Fungsi Loss Actor Network

$$L_{\pi} = \mathbb{E}[\alpha \log \pi(a_t | s_t) - Q(s_t, a_t)] \quad (9)$$

Penjelasan:

Actor dilatih untuk menghasilkan aksi yang memaksimalkan nilai Q sekaligus mempertahankan eksplorasi melalui entropi.

h.) Soft Update Target Network

$$\theta' \leftarrow \tau \theta + (1 - \tau) \theta' \quad (10)$$

- τ : parameter soft update

Penjelasan:

Mekanisme ini digunakan untuk menjaga stabilitas pembelajaran dengan memperbarui target network secara perlahan.

III. METODE

A. Desain Sistem

Penelitian ini bertujuan untuk membangun sistem yang dapat memodelkan karakteristik penggunaan Air Conditioner AC menggunakan metode Reinforcement Learning RL, khususnya algoritma Soft Actor-Critic, guna meningkatkan efisiensi konsumsi listrik berdasarkan data real-time. Sistem ini dirancang untuk memantau parameter lingkungan seperti suhu dalam ruangan, suhu luar, kelembapan, konsumsi energi, dan setpoint suhu, kemudian memproses data tersebut untuk menghasilkan rekomendasi pengaturan AC yang optimal.

Gambar 3. 1 Diagram Blok

Arsitektur sistem secara umum terbagi menjadi empat lapisan utama:

a) Lapisan Perangkat IoT (Data Collection Layer):

Terdiri dari mikrokontroler ESP32 dan sensor DHT22 yang bertugas mengumpulkan data suhu (T_{in}), kelembapan (H), daya penggunaan Ac (P), dan setpoint.

b) Lapisan Cloud dan Penyimpanan Data (Data Storage Layer):

Data sensor dikirim ke platform Firebase Realtime Database yang berfungsi sebagai media penyimpanan dan komunikasi antara perangkat IoT dengan sistem analitik.

c) Lapisan Pemodelan dan Pembelajaran (Learning Layer):

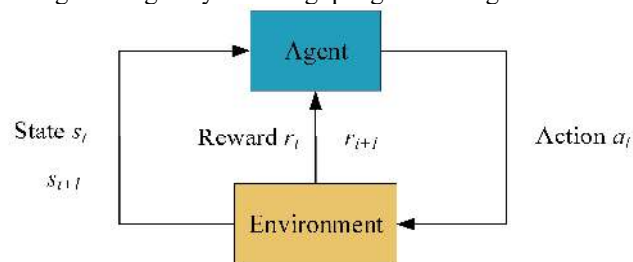
Data yang tersimpan di Firebase kemudian diolah menggunakan Python. Model dinamika AC dilatih untuk memprediksi hubungan antara parameter input dan konsumsi energi. Setelah itu, algoritma Soft Actor-Critic (SAC) dilatih untuk menghasilkan kebijakan pengaturan suhu yang optimal.

d) Lapisan Rekomendasi dan Aplikasi (Application Layer):

Hasil rekomendasi dari model RL dapat ditampilkan melalui antarmuka dashboard atau dikirim kembali ke perangkat untuk mengendalikan AC secara otomatis.

B. Perancangan Soft Actor-Critic

Algoritma Soft Actor-Critic (SAC) diterapkan sebagai pengendali adaptif pada sistem ini untuk menentukan setpoint AC berdasarkan kondisi ruangan secara real time. SAC memodelkan kendali AC sebagai proses Markov (MDP) dengan state berupa kondisi termal dan konsumsi daya, action berupa setpoint AC, serta reward yang dirancang berdasarkan kenyamanan termal dan efisiensi energi. Arsitektur SAC terdiri dari satu jaringan aktor dan dua jaringan kritis sebagai penilai kualitas aksi. Model SAC dilatih menggunakan data historis melalui mekanisme replay buffer, kemudian diterapkan sebagai pengendali otomatis melalui platform IoT untuk menghasilkan setpoint AC yang hemat energi sekaligus nyaman bagi penghuni ruangan.



GAMBAR 1
Lapisan Soft Actor-Critic

1) State (S)

State merupakan kondisi ruangan saat ini yang diterima oleh agent. State terdiri dari:

- Suhu dalam ruangan (T_{in})
- Suhu luar ruangan (T_{out})
- Kelembapan (RH)
- Konsumsi daya AC (P)
- Setpoint AC sebelumnya ($T_{set} - 1$)

2) Action (A)

Action adalah keputusan nilai setpoint AC yang akan diterapkan oleh sistem, Aksi bersifat kontinu, kemudian dibulatkan agar sesuai remote AC Panasonic

3) Reward (R)

Reward menggambarkan kualitas keputusan agent pada waktu tersebut Reward dirancang untuk:

- meminimalkan konsumsi energi
- menjaga jarak suhu ruangan terhadap suhu kenyamanan

Reward dirumuskan sebagai:

$$R_t = -\alpha \cdot |T_{in} - T_{comfort}| - \beta \cdot P_t \quad (12)$$

Di mana:

- $T_{in} - T_{comfort}$ adalah penalti ketidaknyamanan
- P_t adalah konsumsi daya listrik AC
- α dan β adalah bobot.

Semakin tinggi reward \rightarrow semakin baik kendali SAC

C. Alur Kerja Sistem

Alur kerja sistem pemodelan karakteristik penggunaan Air Conditioner berbasis IoT dan Reinforcement Learning dengan algoritma Soft Actor-Critic.

- 1) Sensor DHT22 mengukur suhu dalam ruangan (T_{in}), suhu luar (T_{out}), dan kelembapan (Humidity).
- 2) Sensor kWh DDS238-2-ZN-S mengukur konsumsi energi listrik AC secara real-time.
- 3) Kedua sensor dihubungkan ke mikrokontroler ESP32, yang berfungsi mengumpulkan data dan mengirimkannya ke Firebase Realtime Database melalui koneksi Wi-Fi.
- 4) Firebase Database menyimpan seluruh data sensor dan menjadi penghubung antara perangkat keras (ESP32) dan sistem analisis berbasis machine learning.
- 5) Model RL SAC mengambil data dari Firebase untuk melakukan analisis dan memberikan rekomendasi pengaturan suhu (setpoint) yang optimal agar konsumsi energi lebih efisien tanpa mengurangi kenyamanan.
- 6) Hasil rekomendasi tersebut ditampilkan pada dashboard atau aplikasi yang berfungsi sebagai antarmuka pengguna untuk memantau data dan menyesuaikan suhu AC.

Secara keseluruhan, sistem ini bekerja secara otomatis dan adaptif dalam memantau serta mengoptimalkan penggunaan energi listrik pada AC berbasis data real-time.

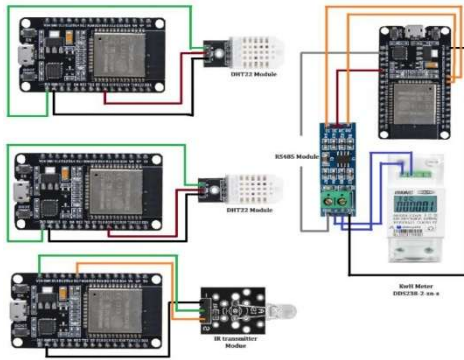
D. Fungsi Dan Fitur

Fungsi utama dari sistem yang dirancang adalah sebagai berikut:

- 1) **Monitoring Data Lingkungan:** Mengumpulkan parameter T_{in} , T_{out} , H, Energy, dan Setpoint secara real-time.
- 2) **Pemodelan Dinamika AC:** Memprediksi perilaku AC berdasarkan kondisi lingkungan.
- 3) **Reinforcement Learning:** Melatih model *Soft Actor-Critic* untuk menentukan kebijakan suhu optimal.

- 4) Rekomendasi Pengaturan: Memberikan saran kenaikan atau penurunan suhu agar efisien.
- 5) Analisis Konsumsi Energi: Menyediakan analisis hubungan antara pengaturan suhu dan konsumsi listrik.

E. Desain Perangkat Keras

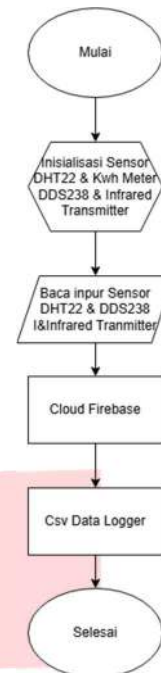


GAMBAR 2

Desain Perangkat Keras

Perangkat keras yang digunakan dalam penelitian ini dirancang dengan pendekatan arsitektur terdistribusi, di mana setiap komponen memiliki fungsi spesifik dalam proses akuisisi data dan pengendalian AC. Sistem terdiri dari empat node utama berbasis ESP32 yang saling terhubung melalui jaringan Wi-Fi dan berkomunikasi dengan layanan cloud Firebase sebagai pusat penyimpanan dan pengolahan data. Node pertama berfungsi sebagai sensor suhu dalam ruangan, menggunakan modul DHT22 untuk mengukur parameter temperatur (Tin) dan kelembapan relatif (RH). Node kedua memiliki konfigurasi serupa dan digunakan untuk mengukur suhu lingkungan luar ruangan (Tout). Kedua node ini memberikan informasi kondisi termal aktual yang menjadi masukan utama bagi algoritma pengendalian. Node ketiga bertugas melakukan pengukuran konsumsi daya listrik AC menggunakan KWh Meter DDS238 yang dihubungkan melalui modul RS485 dan protokol Modbus RTU. Data daya sesaat (P) dan energi kumulatif (Wh) dikirimkan secara berkala ke Firebase sebagai basis analisis efisiensi energi. Node keempat berfungsi sebagai aktuator yang mengendalikan AC menggunakan modul infrared transmitter. Node ini menerima perintah setpoint suhu, mode operasi, dan pengaturan kipas yang dihasilkan oleh algoritma Soft Actor-Critic (SAC) dan disimpan pada Firebase. Perintah tersebut dikonversi menjadi sinyal inframerah sesuai protokol AC Panasonic, sehingga kontrol dapat dilakukan tanpa modifikasi perangkat AC. Keempat node bekerja secara terpadu untuk menyediakan data operasional yang lengkap bagi proses pembelajaran dan evaluasi model SAC. Dengan desain ini, sistem mampu melakukan monitoring, analisis, dan pengendalian AC secara otomatis berdasarkan kondisi lingkungan aktual, sehingga dapat meningkatkan efisiensi energi sekaligus menjaga kenyamanan termal pengguna.

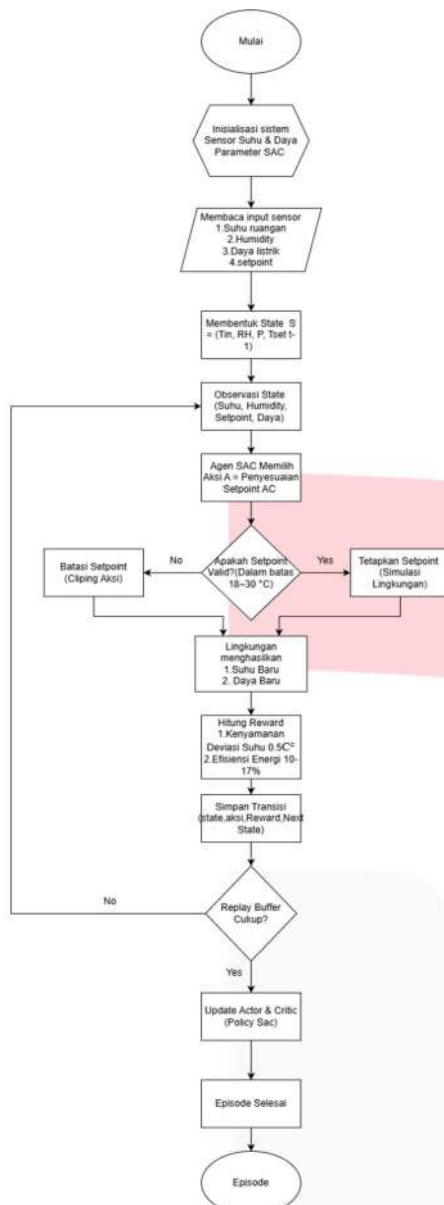
C. Flowchart



GAMBAR 3

Flowchart Sistem Pemodelan Reinforcement Learning

Flowchart sistem diawali dengan proses inisialisasi, di mana sensor suhu dan kelembapan (DHT22 dan DS3238) serta infrared transmitter diaktifkan untuk memastikan sistem siap beroperasi. Setelah itu, sistem melakukan pembacaan data suhu ruangan dan parameter pendukung kondisi operasional Air Conditioner (AC). Data hasil pembacaan sensor dikirimkan ke cloud Firebase sebagai media penyimpanan data secara real-time dan historis. Selanjutnya, data tersebut diekspor dalam bentuk dataset berformat CSV melalui proses data logging menggunakan Visual Studio Code (VS Code). Dataset CSV digunakan sebagai lingkungan *Reinforcement Learning* (environment), dengan state berupa suhu ruangan, setpoint AC, dan konsumsi daya listrik. Aksi didefinisikan sebagai perubahan setpoint AC, sedangkan reward dirancang untuk menyeimbangkan kenyamanan termal dan efisiensi energi. Environment berinteraksi dengan agen *Soft Actor-Critic* (SAC) untuk mempelajari kebijakan pengendalian AC yang optimal. Hasil pembelajaran agen SAC kemudian dievaluasi melalui simulasi dan analisis efisiensi energi, sehingga diperoleh karakteristik penggunaan AC yang lebih efisien. Proses diakhiri dengan tahap selesai.



GAMBAR 5

Fowchart Sistem Pemodelan Reinforcement Learning

Proses pemodelan diawali dengan tahap mulai sebagai penanda awal sistem. Selanjutnya, dataset historis penggunaan Air Conditioner (AC) yang meliputi suhu ruangan, kelembapan, setpoint AC, dan konsumsi daya listrik diproses sebagai dasar pembentukan model simulasi. Dataset tersebut digunakan untuk membangun lingkungan Reinforcement Learning (environment) yang merepresentasikan karakteristik operasional AC. Setelah lingkungan terbentuk, agen Soft Actor-Critic (SAC) diinisialisasi, yang terdiri dari jaringan actor dan critic untuk mempelajari kebijakan pengambilan keputusan optimal. Pada setiap episode, lingkungan di-reset untuk menentukan state awal. Agen kemudian mengamati state yang mencakup suhu, kelembapan, setpoint, dan daya listrik, lalu memilih aksi berupa penyesuaian setpoint AC. Aksi ini diterapkan pada lingkungan sehingga menghasilkan state baru. Selanjutnya, sistem menghitung reward yang merepresentasikan kenyamanan termal dan efisiensi energi. Transisi state-aksi-reward-state berikutnya digunakan untuk memperbarui kebijakan agen SAC. Proses ini berlangsung secara iteratif

hingga episode selesai, dan setelah seluruh episode simulasi berakhir, proses pemodelan dinyatakan selesai.

VI HASIL DAN PEMBAHASAN

A. Grafik Pengujian Reward Soft Actor-Critic

Pada grafik reward SAC, terlihat bahwa nilai reward pada awal episode (0–100 menit) masih relatif rendah dengan variasi cukup besar (≈ -0.16 hingga -0.10). Ini mencerminkan bahwa kontrol yang diambil oleh agen SAC pada tahap awal masih belum optimal dalam menyeimbangkan error kenyamanan dan konsumsi energi. Memasuki menit ke 150–300, reward meningkat dan stabil mendekati nilai -0.06 hingga -0.08 . Pola ini menunjukkan bahwa agen berhasil menemukan strategi kontrol yang lebih konsisten, di mana kombinasi setpoint dan konsumsi daya mulai lebih efisien. Selain itu, reward kumulatif yang secara gradual turun namun dengan slope yang menurun memperlihatkan bahwa keputusan SAC semakin dekat ke target optimal lingkungan. Hal penting adalah reward lebih stabil dibandingkan periode awal, menandakan bahwa SAC telah mempelajari perilaku termal ruangan dan beban pendinginan secara adaptif.

B. Grafik Pengujian Distribusi Setpoint SAC Terhadap Penggunaan Manual

Distribusi setpoint manual sangat terbatas pada nilai tertentu (24°C dan 28°C), mencerminkan pola kontrol statis dan tidak adaptif terhadap perubahan kondisi ruangan. Sebaliknya SAC menghasilkan distribusi setpoint yang lebih luas (25°C – 27°C) dan adaptif terhadap dinamika suhu luar dan beban pendinginan. Hal ini menunjukkan dua keunggulan utama:

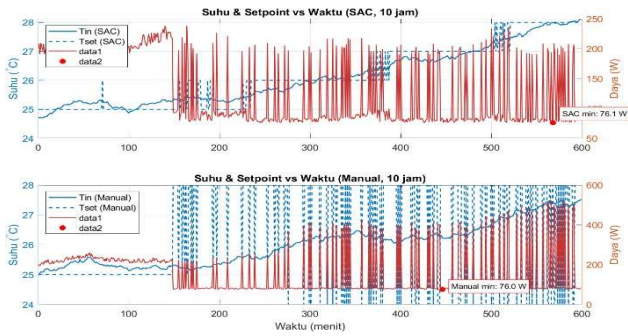
- 1) SAC dapat menyesuaikan setpoint secara real-time berdasarkan kondisi termal yang berubah.
- 2) Setpoint SAC lebih stabil, karena tidak mengalami switching ekstrem seperti metode manual.

Distribusi SAC yang lebih halus ini menghasilkan sistem yang tidak hanya lebih hemat energi tetapi juga lebih nyaman.

C. Grafik Pengujian Terhadap Kontrol Suhu Antara Penggunaan Manual dengan Penggunaan Soft Actor-Critic

Pada grafik kontrol SAC, kurva suhu ruangan (T_{in}) mengikuti setpoint dengan jarak penyimpangan yang kecil. SAC mempertahankan suhu $\pm 0.5^{\circ}\text{C}$ dari target yang diinginkan, dan ketika terjadi perubahan kondisi lingkungan, SAC menyesuaikan setpoint secara bertahap. Sebaliknya, grafik kontrol manual menunjukkan pola setpoint yang sangat agresif (frekuensi switching tinggi), mengindikasikan kontrol yang tidak stabil. T_{in} pada mode manual juga sering mengalami overshoot hingga 2 – 3°C dari setpoint, menyebabkan ketidaknyamanan termal dan peningkatan konsumsi energi. SAC jauh lebih stabil dan halus, sedangkan manual cenderung liar dan tidak teratur.

D. Grafik Pengujian Terhadap T_{in} (Temperature Indoor) – T_{set} (Setpoint AC) – Power Penggunaan selama 10 Jam



GAMBAR 6

Grafik Pengujian Terhadap Tin (Temperature Indoor-Setpoint-Power Penggunaan 10 Jam

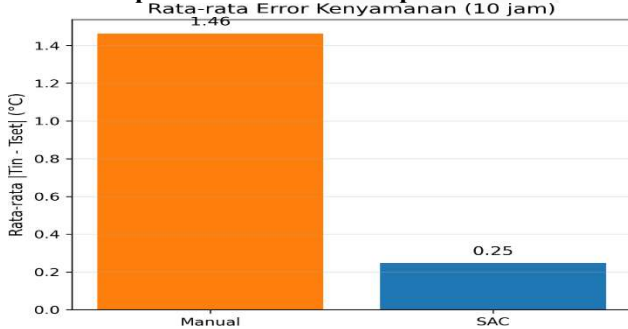
Grafik ini menunjukkan dinamika simultan suhu dalam, setpoint, dan konsumsi daya. Mode SAC

- Daya SAC berada pada range 75–230 W, konsisten dan rendah.
- SAC menggunakan daya minimal lebih sering dan menghindari lonjakan konsumsi.
- Nilai daya minimum tercatat sekitar 76.1 W dengan label pada grafik.
- Ketika suhu ruangan mulai naik, SAC menaikkan output secara bertahap, bukan drastis.

Mode Manual

- Daya mode manual sering melonjak hingga 400–550 W.
 - Banyak spike daya menunjukkan kontrol on-off yang tidak efisien.
 - Meskipun daya minimum sama (≈ 76 W) frekuensi penggunaan daya rendah jauh lebih sedikit.
- Perbandingan keduanya menunjukkan bahwa SAC lebih hemat dan lebih stabil, sementara mode manual cenderung membuang energi karena kontrol yang tidak teratur.

E. Grafik Pengujian Rata Error Terhadap Kenyamanan Antara Temperatur Indoor dan Setpoint Ac



GAMBAR 7

Grafik Pengujian Rata-Rata Error Terhadap Kenyamanan antara Temperature Indoor Dan Setpoint Ac

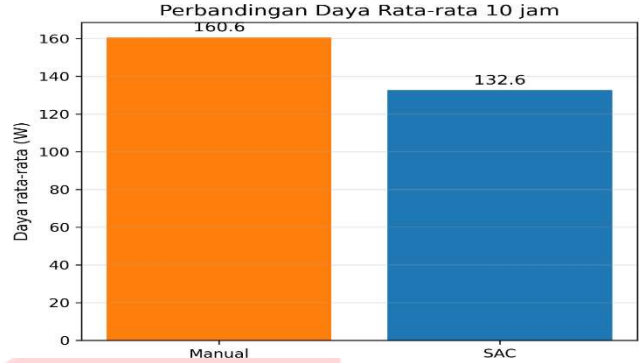
- Rata-rata error manual: 1.46°C
- Rata-rata error SAC: 0.25°C

Dengan demikian SAC memberikan kenyamanan 82.9% lebih baik. Pada grafik time series:

- Manual sering overshoot hingga 3°C.
- SAC mempertahankan error sangat kecil (0–0.6°C).

Ini merupakan bukti kuat bahwa SAC tidak hanya menghemat energi, tetapi juga meningkatkan kenyamanan ruangan.

F. Grafik Pengujian dan Perbandingan Energi Antara Mode Manual Dan Mode Soft Actor-Critic



GAMBAR 8

Grafik Pengujian dan Perbandingan Energi Antara

Mode Manual Dan Mode Soft Actor-Critic Dari grafik konsumsi energi total 10 jam:

- Manual: 1.606 kWh
- SAC: 1.326 kWh

Selisih penghematan = 0.280 kWh. Angka ini sangat signifikan untuk sistem HVAC (Heating Ventilation and Air Conditioning), mengingat AC umumnya adalah beban terbesar dalam konsumsi listrik harian. Hal yang menyebabkan penghematan tersebut:

I KESIMPULAN

Berdasarkan hasil penelitian, penerapan algoritma Soft Actor-Critic (SAC) pada sistem pengendalian AC berhasil merepresentasikan dinamika termal ruangan dengan baik, ditunjukkan oleh akurasi model dengan MAE suhu kurang dari 0,5°C sehingga layak digunakan sebagai environment simulasi. Dibandingkan metode kontrol manual, SAC mampu menjaga kestabilan suhu dengan error kenyamanan yang lebih kecil, yaitu rata-rata 0,25°C, sementara metode manual mencapai 1,46°C. Selain meningkatkan kenyamanan termal, SAC juga menunjukkan efisiensi energi yang lebih baik dengan penghematan konsumsi listrik sebesar 10–17% selama simulasi 10 jam. Kinerja tersebut dicapai melalui strategi penyesuaian setpoint yang adaptif dan seimbang antara kenyamanan dan efisiensi energi. Hasil simulasi juga menunjukkan bahwa sistem SAC beroperasi secara stabil dan aman, sehingga memiliki potensi untuk diimplementasikan pada sistem AC nyata berbasis perangkat IoT.

REFERENSI

[1] Kementerian Energi dan Sumber Daya Mineral, Statistik Ketenagalistrikan Indonesia 2023, Jakarta, Indonesia, 2023.

[2] W. Yang et al., "Thermal comfort and productivity in office buildings," *Energy and Buildings*, vol. 49, pp. 15–22, 2012.

[3] Kementerian ESDM RI, *Handbook of Energy & Economic Statistics of Indonesia*, 2023.

- [4] Z. Yu, B. Fung, F. Haghghat, H. Yoshino, and E. Morofsky, "A systematic procedure to study the influence of occupant behavior on building energy consumption," *Energy and Buildings*, vol. 43, no. 6, pp. 1409–1417, 2011.
- [5] A. Kusiak, M. Li, and F. Tang, "Modeling and optimization of HVAC energy consumption," *Applied Energy*, vol. 87, no. 10, pp. 3092–3102, 2010.
- [6] Y. Wei et al., "Deep reinforcement learning for building HVAC control: A review," *Applied Energy*, vol. 268, 2020.

