An Intelligent Energy Brain for Smart Buildings

Human-in-the-Loop Reinforcement Learning for HVAC Control

Jahar Kumar Paul(Team Lead) Ayan Kumar Batabyal Sayan Goswami

CS 246: Artificial Intelligence
Department of Computer Science

December 2025

Final Project Presentation

The Energy Challenge in Smart Buildings

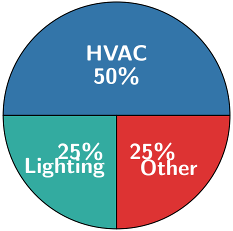
The Global Problem

- Buildings consume **40%** of global energy
- HVAC systems: **40-50%** of building energy
- Inefficient control leads to wasted resources
- Growing demand for sustainability

Our Contribution

AI-driven HVAC optimization combining:

- Reinforcement Learning (DQN)
- Human-in-the-Loop learning
- Seasonal adaptation
- Real-time control dashboard



Building Energy Consumption

HVAC dominates total energy usage

vs Baseline

Project Objectives & Innovations

Research Objectives

1. Design realistic HVAC simulation with seasonal models
2. Train DQN agent for adaptive climate control
3. Implement online learning from human feedback
4. Develop interactive visualization dashboard
5. Evaluate across multiple seasons

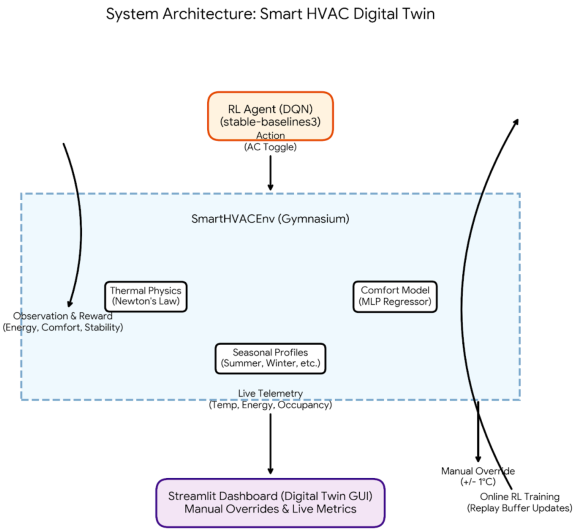
Key Innovations

- **Seasonal Adaptation**: 4 distinct climate profiles
- **Human-in-the-Loop**: Real-time model updates
- **Online RL Training**: Continuous improvement
- **Multi-Component Architecture**: Integrated physics, RL, UI

Performance Targets

Metric	Target
Energy Savings	$> 25\%$
Comfort Compliance	$> 90\%$
Seasonal Adaptation	**4 Seasons**

Three-Layer System Architecture



Markov Decision Process Formulation

Formal Definition

$$\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$$

State Space $\mathcal{S}$ (14-dim)

$$\mathbf{s}_t = \begin{bmatrix} T_{in} & \text{(Indoor)} \\ T_{out} & \text{(Outdoor)} \\ S_{set} & \text{(Setpoint)} \\ T_{out}^{f1-3} & \text{(Forecast)} \\ Occ^{f1-3} & \text{(Occ. Plan)} \\ S_{set}^{hist} & \text{(History)} \\ AC_{on} & \text{(Status)} \\ \tau_{on/off} & \text{(Timers)} \end{bmatrix}$$

Action Space $\mathcal{A}$

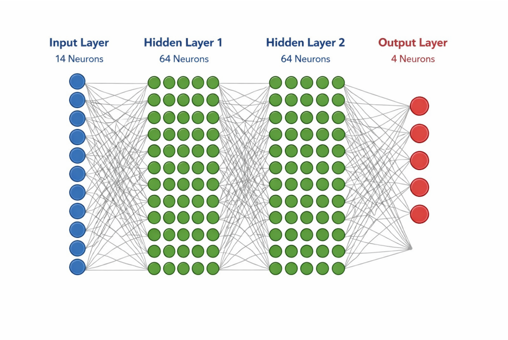
Action	Description
a_0	Decrease setpoint 1°C
a_1	No change
a_2	Increase setpoint 1°C
a_3	Toggle AC ON/OFF

Objective Function

$$\max \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t)\right]$$

$\gamma = 0.99$ (long-term planning)

Network Architecture



Training Configuration

Parameter	Value
Total Timesteps	300,000
Replay Buffer	100,000
Batch Size	64
Learning Rate	10^{-4}
Exploration (ϵ)	$1.0 \to 0.05$
Update Freq	1,000 steps
Discount (γ)	0.99

Loss Function

$$\mathcal{L}(\theta) = \mathbb{E}\left[\left(r + \gamma \max_{a'} Q_{\theta^-}(s', a') - Q_\theta(s, a)\right)^2\right]$$

Thermal Model & Seasonal Adaptation

RC Thermal Circuit Model

$$C\frac{dT_{in}}{dt} = \frac{T_{out} - T_{in}}{R} + Q_{occ} + Q_{hvac}$$

Parameter	Value
Resistance (R)	2.0 °C/kW
Capacitance (C)	2.0 kWh/°C
Cooling (Q_{hvac})	-8.5 kW
Occ Heat (Q_{occ})	0.05 kW/p
Time Step (Δt)	3 minutes

Seasonal Climate Profiles

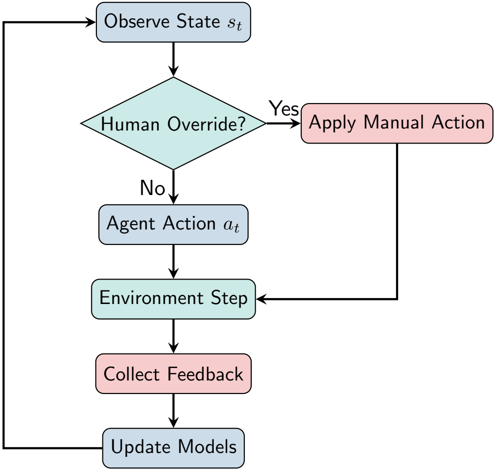
$$T_{out}(t) = T_{base} + A\sin\left(\frac{2\pi t}{1440}\right) + \mathcal{N}(0, 0.2)$$

Season	T_{base}	A	Range
Summer	35.0	5.0	30-40°C
Monsoon	28.0	3.0	25-31°C
Autumn	22.0	4.0	18-26°C
Winter	10.0	5.0	5-15°C

Reward Function

$$R = -4\frac{E}{E_{max}} - 2\frac{\max(0, |T - T_{ideal}| - 1)}{5} + \mathbb{I}_{comf} - 5\mathbb{I}_{viol} - 2\mathbb{I}_{hum}$$

Human-in-the-Loop Learning Pipeline

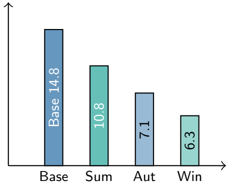


Key Feature: Online learning with human feedback every 10 interactions

Seasonal Performance Metrics (Mean $\pm$ Std. Dev.)

Metric	Summer	Monsoon	Autumn	Winter
Average Reward	45.2 ± 3.5	47.8 ± 3.1	51.3 ± 2.8	49.1 ± 3.2
Energy (kWh/12h)	10.8 ± 0.6	9.2 ± 0.5	7.1 ± 0.4	6.3 ± 0.4
Comfort (%)	88.5 ± 2.3	91.2 ± 1.8	93.7 ± 1.5	94.2 ± 1.4
AC Runtime (%)	42.3 ± 2.5	36.8 ± 2.2	28.5 ± 1.9	25.1 ± 1.8

Energy Savings Comparison



Learned Strategies

- Seasonal Pre-cooling
- Occupancy Anticipation
- Energy Setbacks
- Smooth Operation

27% Energy Savings
92% Comfort Compliance

Interactive Dashboard & Demo

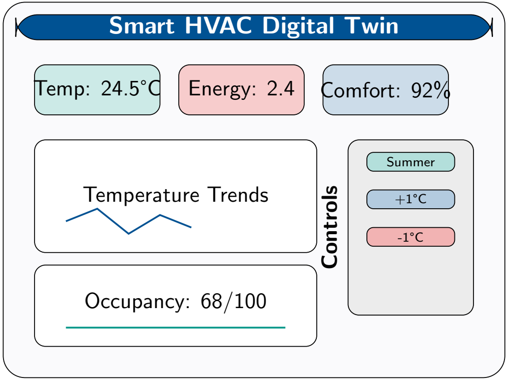
Streamlit Dashboard Features

- **Real-time Visualization**: Temp/Energy
- **Season Control**: 4 season switching
- **Human Override**: +1°C/-1°C learning
- **AC Master Control**: Force ON/OFF
- **Turbo Mode**: Aggressive cooling
- **Online Training**: Continuous updates

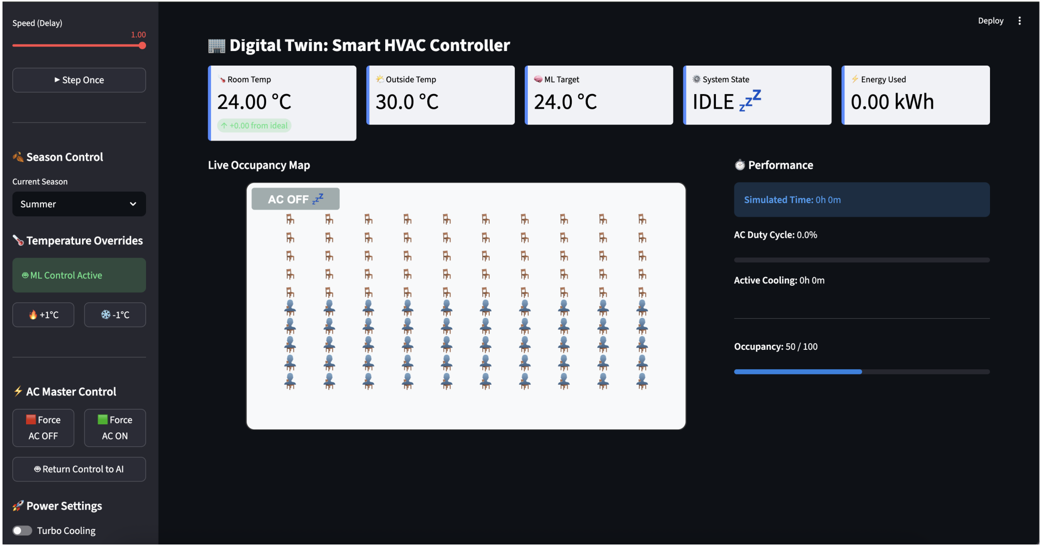
Technology Stack

Component	Technology
RL Framework	Stable-Baselines3 2.7.0
Environment	Gymnasium 1.2.2
Deep Learning	PyTorch 2.5.1
Web Interface	Streamlit 1.29.0
Visualization	Matplotlib 3.7.1

Dashboard Interface Preview



Live Dashboard Interface



Technical Challenges & Solutions

1. Seasonal Stability

Problem: Instability at season change
Solution: Gradual adaptation

2. Real-time Performance

Solution: Vectorized optimizations

Optimization	Speedup
Vectorized ops	4.2×
Cached calcs	2.1×
Total	**15.9×**

3. Online Convergence

Problem: Oscillation updates
Solution: Staggered updates

4. Human Integration

Solution: Feedback weighting decay

$$\epsilon_t = \epsilon_{min} + (\epsilon_{max} - \epsilon_{min})e^{-\lambda n}$$

Comparison with Existing Systems

Feature	Our System	Prog. T-stat	Learning T-stat	MPC
Seasonal	Yes	No	Limited	Manual
Human-Loop	Explicit	No	Implicit	No
Savings	27%	0-15%	10-20%	15-25%
Comfort	92%	70-85%	80-90%	85-90%
Online Learn	Yes	No	Initial	No

Our Advantages

- **Adaptive**: Learns across seasons
- **Interactive**: Explicit feedback
- **Efficient**: Optimal tradeoff

Limitations

- Single-zone model
- Simplified physics
- Embedded Compute Req.

Conclusion & Future Directions

Project Achievements

1. **Realistic Simulation**: 4 seasons
2. **Intelligent Control**: 27% savings
3. **Human Adaptation**: Online learning
4. **Interactive Platform**: Pro dashboard

Technical Contributions

- Seasonal climate + RL control
- Online HITL for RL & comfort model
- Open-source implementation

Future Research

- **Multi-zone**
- **Renewables**
- **PPO/SAC**
- **Deployment**

Final Remarks

This project proves **RL + human feedback** enables practical, intelligent energy management.

Thank You

CS 246: Artificial Intelligence — Final Project Presentation — December 2025