I found inspiration during my Anatomy and Physiology class, learning about homeostasis, how systems regulate themselves under changing conditions. As I learned about the energy demands of AI data centers, I noticed that operators often lack visibility in how energy is wasted. EcoEyes represents my first step toward exploring adaptive, learning-based system to support more sustainable AI infrastructure. I built a prototype in Replit using AI-assisted prompts to create a simulation framework. Constraints included energy balance, fan power proportional to speed^3, and Power Usage Effectiveness (PUE $\geq$ 1.0). Cooling was modeled through adjustments to fan speed within the simulation rather than through physical hardware. To reflect real-world uncertainty, Gaussian noise was added to temperature readings and fault injection was used to simulate disturbances. Over 30 randomized runs, EcoEyes used adapted to cooling behavior based on past outcomes rather than fixed thresholds. While a simple rule-based system could increase cooling when temperature rises at a certain rate, Q-learning, allows the system to learn which actions reduce overheating over time (Sutton & Barto, 2015). EcoEyes also used a lightweight spiking neural network to identify temperature changes. Fixed-rule control resulted in approximately 80 overheating incidents, while EcoEyes maintained lower peak temperatures without any violations. PID control also avoided violations, but EcoEyes provided a larger safety margin, with 1% higher energy use. Overheating was evaluated based on violations of thermal constraints rather than a single absolute temperature value. Future work would explore using a Leaky Integrate-and-Fire spiking neural network for more robust temporal change detection, suitable for data centers because they can be implemented on existing monitoring systems and their integrate-and-threshold behavior reflects the slow, cumulative nature of server-rack heating (Gerstner, 2014).

EcoEyes – Optimize AI Data Centers Energy Usage | Helena Bahramirad infrastructure.

I appreciate open-source resources like Replit, Python libraries such as NumPy, and International Energy Agency reports, which helped me develop, test, and validate EcoEyes.

References:

Boi, F., Rössert, C., & Ekanayake, D. (2024). Spike-based learning enables energy-efficient and scalable edge AI. Nature Communications, 15, Article 5797. https://www.nature.com/articles/s41467-024-51110-5

DeepMind. (2016, July 20). DeepMind AI reduces Google data centre cooling bill by 40%. https://deepmind.google/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-by-40/

Gao, Y., Wang, G., & Guan, X. (2016). Using model predictive control in data centers for dynamic server provisioning. arXiv. https://arxiv.org/abs/1611.00522

Gerstner, W. (2014). *1.3 integrate-and-fire models*. 1.3 Integrate-And-Fire Models | Neuronal Dynamics online book. https://neuronaldynamics.epfl.ch/online/Ch1.S3.html

Kansal, A., Zhao, F., Liu, J., Kothari, N., & Bhattacharya, A. A. (2008). Sense and sensitivity: Characterizing temperature and power management in data centers (Technical Report MSR-TR-2008-107).

Microsoft Research. https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/dcsensefinal.pdf

Lohrmann, C., & Pitakrat, T. (2016). Using model predictive control in data centers for dynamic server provisioning. ResearchGate. https://www.researchgate.net/publication/301435822

Orvis, R. (2020, March 17). How much energy do data centers really use? Energy Innovation. https://energyinnovation.org/expert-voice/how-much-energy-do-data-centers-really-use/

Sutton, R. S., & Barto, A. G. (2015). Reinforcement Learning: An Introduction. Cambridge, Massachusetts; The MIT Press Cambridge, Massachusetts London, England. https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf

Zanatta, L., Barchi, F., Manoni, S., Tolu, S., Bartolini, A., & Acquaviva, A. (2024, December 28). *Exploring spiking neural networks for deep reinforcement learning in robotic tasks*. Nature News. https://www.nature.com/articles/s41598-024-77779-8

Bibliography:

Array programming with NumPy. Nature, 585, 357–362. https://numpy.org/doc

Åström, K. J., & Murray, R. M. (2010). Feedback systems: An introduction for scientists and engineers. Princeton University Press. https://fbswiki.org/wiki/index.php/Main_Page

Summer Learning Program. (2024). Anatomy and physiology worksheets [Unpublished educational material]. North Vancouver School District Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., … Oliphant, T. E. (2020).

MIT Open Learning. (2021). AI 101: An introduction to artificial intelligence [Course]. MIT OpenCourseWare / MIT Open Learning. https://ocw.mit.edu/courses/res-6-013-ai-101-fall-2021/