# Energy Consumption of Training Large Models

Quacky has been enjoying his adventures in the pond, training models to predict swimming speeds, classify feathers, and even detect unusual behavior. But one day, the wise owl lands beside him with a serious expression.

"Quacky, there is something every machine learning enthusiast must consider, the energy it takes to train models."

Quacky tilts his head in curiosity

"Energy? Like the energy I use to swim around the pond?"

The owl chuckles softly

"Yes, but imagine the pond as a gigantic hamster wheel. Every time we train a machine learning model, that wheel spins. And bigger models spin it much faster, consuming vastly more electricity."

Suddenly, Quacky understands: training a model is not just about intelligence, it is about power and resources.

## Why Training Consumes So Much Energy

The owl explains the key factors:

### 1. Large Models = More Computation

Deep learning models, especially those with many layers, can contain millions or even billions of parameters. Each parameter needs to be updated repeatedly during training.The more parameters a model has, the more calculations it performs, and each calculation consumes electricity

**Quacky imagines: "It's like teaching thousands of ducks to swim at the same time, all flapping as hard as they can. That takes a lot of energy!"**

### 2. Big Datasets = More Work

Large datasets require multiple passes through the data, called iterations or epochs. Each pass involves complex mathematical operations.More data means more passes, more computation, and therefore, more energy

### 3. Hardware Usage

Training is not done by ordinary pond ducks, it happens on supercharged duck brains: GPUs and TPUs. These devices crunch massive amounts of numbers constantly. They draw enormous amounts of electricity, generate heat, and require cooling.

**Quacky pictures it vividly: a room full of super-ducks flapping non-stop, each powered by its own tiny electricity generator. The energy consumption is enormous.**

## Real-Life Scale of Energy Use

The owl gives Quacky some eye-opening examples:

- Training a single large language model can consume as much electricity as several households use in a month.
- The carbon footprint of one training run can be equivalent to the emissions from multiple airplane flights.
- Even a single model training session can burn hundreds or thousands of liters of fuel if powered by conventional electricity sources.


**Quacky gasps: "Just to teach a machine, humans can use enough energy to power my entire pond for weeks!"**

## Why Energy Consumption Matters

The owl explains two major reasons why it matters:

### 1. Operational Costs

Running large models is not cheap. Data centers spend huge amounts of money on:

- Electricity to power the machines
- Cooling systems to prevent overheating
- Maintenance of the hardware

### 2. Environmental Impact

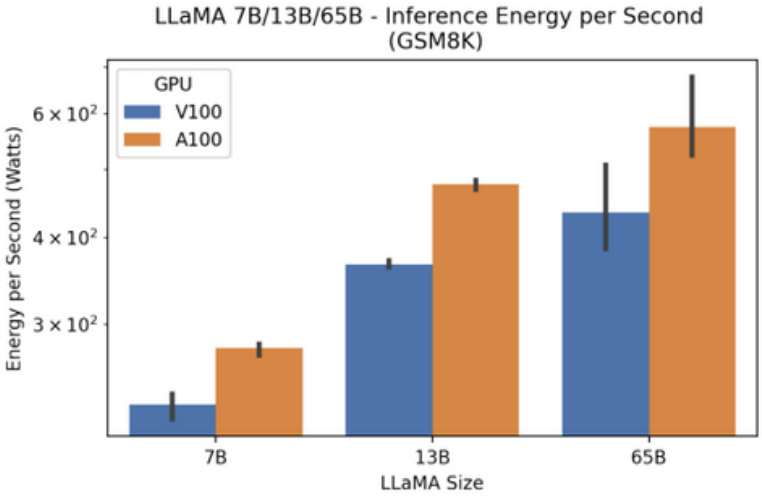
If the electricity powering these machines comes from fossil fuels, greenhouse gases are released into the atmosphere. Large-scale training can contribute significantly to climate change.

**Quacky reflects thoughtfully: "Even as a duck trying to build smart models, I need to consider the energy footprint. Every prediction, every training run, has a cost beyond just time and effort."**

## Ways to Visualize Energy Usage

The owl shares ideas Quacky can imagine:

**Bar Charts:** Comparing energy used by small models versus large models clearly shows the massive difference.




- **Duck World:**
  - A tiny model is like a single duck on a small hamster wheel, slow and low energy.
  - A huge model is like a turbo-charged duck on a giant wheel spinning uncontrollably, consuming enormous power.
- **Relatable Comparisons:**
  - Kilowatt-hours used are equivalent to powering a household for a month
  - Carbon emissions are equivalent to a flight or multiple cars driving the same distance

**These comparisons make the abstract numbers concrete for Quacky's duck brain.**

## Quacky's Takeaway

Quacky fluffs his feathers and nods:

> **Smart machines are amazing, but teaching them comes at a cost. Bigger models need more energy, which impacts both the planet and my pond budget. Thinking about energy efficiency is just as important as thinking about accuracy or intelligence. Responsible machine learning means being aware of the power it consumes.**

With this understanding, Quacky now knows that building intelligent models is not just about brains, it's also about power, resources, and responsibility.

## References

S. Samsi, D. Zhao, J. McDonald, B. Li, A. Michaleas, M. Jones, W. Bergeron, J. Kepner, D. Tiwari, and V. Gadepally, "From Words to Watts: Benchmarking the Energy Costs of Large Language Model Inference," arXiv:2310.03003, Oct. 2023.