A Proposal to Reduce Energy Consumption at OpenAI

Prepared for:

Mira Murati

CTO of OpenAI

By

Galal Qasas and Mark Samuel

8/5/2023

Table of contents

PROPOSED SOLUTION

To solve the problem of unsustainable AI that the training of a model costs a lot memory and energy-wise, OpenAI needs to take approaches: (1) using quantum computing in their future models (2) start using transfer learning.

Phase 1: Using a Quantum Computer

Traditional computers rely on traditional physics which is based on the single state of a bit either 0 or 1. On the other hand, quantum computers give more energy-efficient opportunities. They rely on the rules of quantum mechanics to solve complex problems. The superposition property states that qubit the basic unit of quantum information can be in different states simultaneously, meaning the computer can do multiple calculations simultaneously. This accelerates the computing to 10^{10} bits/s (Ezhov & Ventura, 2000) with less energy consumption (Elsayed et al., 2019).

Entanglement, another fundamental property of quantum mechanics, when qubits entangles they change the states of all qubits together so they can compute multiple calculations simultaneously. so a lot of qubits compute complex problems (Jeswal & Chakraverty, 2018) much faster than traditional computers.

Implementation of a future AI model on a quantum computer would need hiring a quantum circuits expert(s) depending on the model complexity and size to design the quantum circuit for the model to run on. In Figure 1 the average energy consumption for a supercomputer compared to a quantum computer.

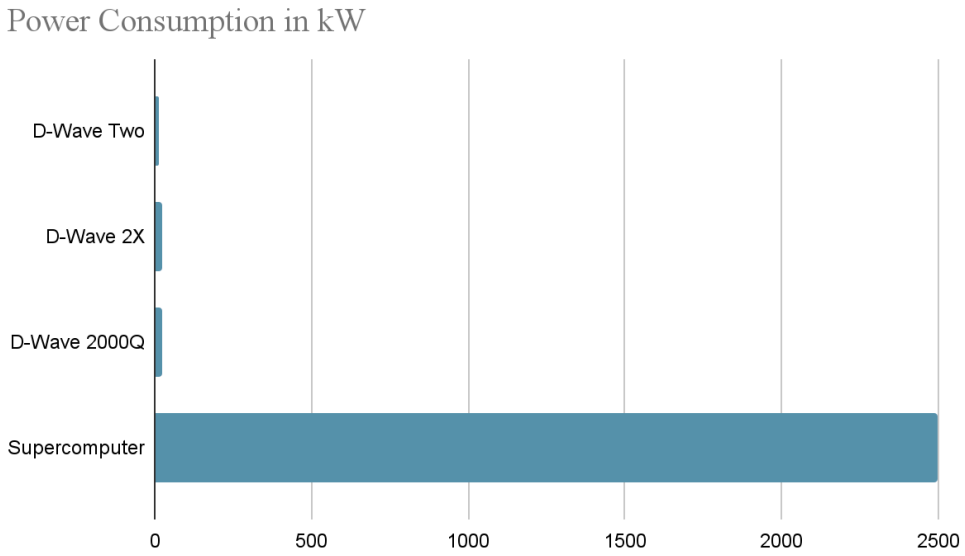

Figure 1: Adapted from Review paper "A Review of Quantum Computer Energy Efficiency" by (Elsayed N et al.,2019). Without adding the cooling cost.

Phase 2: Using Transfer Learning (TL)

Transfer learning is a machine learning approach. It is training a model on general data and then reusing the model for a different purpose than the original data. The model relates the new tasks with the old tasks. In this approach, The engineer should separate the searching and training algorithms saving money, hardware and reducing CO_2 emissions (Cai et al., 2020).

"The total CO2 emissions of OFA is 16x fewer than ProxylessNAS, 19x fewer than FBNet, and 1,300x fewer than MnasNet" (Cai et al., 2020) after the optimization of the model and comparing it to other models. Figure 2 is a statistic about the models and it's carbon emissions.

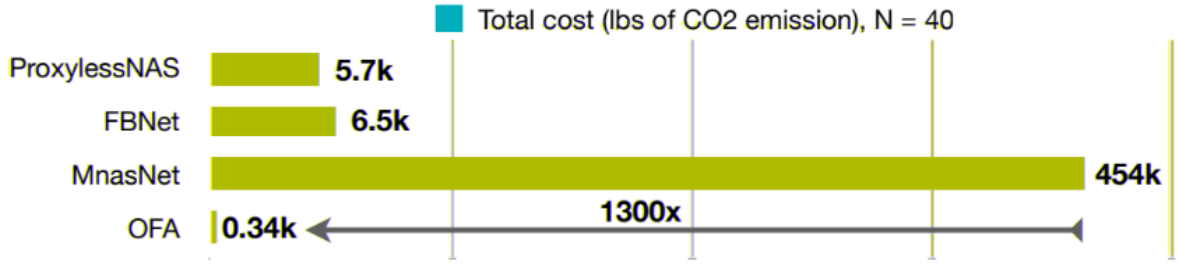

Figure 2: Adapted from conference paper " Once-for-all: train one network and specialize it for efficient deployment" by (Cai et al.,2019)

Adopting the transfer learning model would let OpenAI create new models without retraining the model for the data repeatedly for editing the training parameters repeatedly.

Costs:

- The majority of the costs are from the quantum computer approach.
- Hiring quantum computing expert(s) to design the quantum circuits will cost annually $80k to $130k per expert. (

  , n.d.).
- The Amazon pricing would vary according to the model training cycles for a single cycle it will cost $0.3 * the iteration times (, n.d.).
- The Transfer learning training approach would cost significantly less than what training a model from scratch will cost, but it varies according to the model size and complexity. (Hosna, 2022)

Benefits:

- Implementation of these phases will reduce OpenAI's environmental impact by reducing its carbon footprint.

- Using the transfer learning approach in future models will reduce the cost of creating new models by using already trained models in new models in addition to reducing the energy cost for the new models.
- Using quantum computing would reduce the energy usage of the model and increase its efficiency.
- Enhance the name of OpenAI as a company that pioneers in using efficient and environmentally friendly ways of creating AI models and helps them in reaching their goal of implementing artificial general intelligence efficiently and accurately.

Glossary:

Qubit: corresponds to the bit in the traditional computer.

ProxylessNAS, FBNet, MnasNet: AI models.

Transfer learning (TL): another approach in machine learning instead of training the model from scratch

Quantum computing: it is another tepe of computing which happens on quantum computers

References

Cai, H., Gan, C., Wang, T., Zhang, Z., & Han, S. (2020). Once for All: Train One Network and Specialize it for Efficient Deployment.

. https://openreview.net/pdf?id=HylxE1HKwS

Cai, H., Gan, C., Zhu, L., & Han, S. (2020). TinyTL: Reduce Activations, Not Trainable Parameters for Efficient On-Device Learning. .

http://export.arxiv.org/pdf/2007.11622

Elsayed, N. S., S. Maida, A. S. M., & Bayoumi, M. B. (2019, April 1).

. IEEE Conference Publication | IEEE Xplore.

https://ieeexplore.ieee.org/document/8767125

Ezhov, A. A., & Ventura, D. (2000). Quantum Neural Networks. , .

https://doi.org/10.1007/978-3-7908-1856-7_11

Hosna, A. (2022). Transfer learning: a friendly introduction. .

https://doi.org/10.1186/s40537-022-00652-w

Jeswal, S. K., & Chakraverty, S. (2018). Recent Developments and Applications in Quantum Neural Network: A Review. . https://doi.org/10.1007/s11831-018-9269-0

. (n.d.). Amazon Web Services, Inc. https://aws.amazon.com/braket/pricing/

. (n.d.). ZipRecruiter.

https://www.ziprecruiter.com/Salaries/Quantum-Computing-Salary