ALGORITHMS AND PERSONALITIES

Anonymous authors
Paper under double-blind review

ABSTRACT

Most Deep Learning based Model Compression methods draw their inspiration from some kind of natural phenomenon. This simple paper aims to draw attention toward a similar mapping between algorithms and personalities hoping to benefit the discipline of one from the other.

1 INTRODUCTION

In vain, many disciplines have tried to define how exactly the human mind works. On the other hand, the world of algorithms is also thriving at an exponential rate. Thousands of research papers push the limits of the state of the art every year. Having a complexity that can potentially match that of our mind, algorithms can assist in uncovering the mysteries of the mind. This article is an attempt to bridge the two areas. It starts with simple ideas but also discusses future endeavors.

2 MOTIVATION

Neural Network based models are made up of Artificial Neurons that serve as abstractions to mimic the mind. Since the exact number and arrangement of these Artificial Neurons is still a hyper-parameter, a decision of the designer, we have them large first and then try to reduce their size with Model Compression, a procedure to make them small yet performant. The three major Model Compression methods are Pruning Quantization, and Knowledge Distillation. The interesting point is that they are somehow connected to some natural phenomena outside of their discipline: Pruning is related to how we humans do not need most information at graduate school that we needed earlierFrankle & Carbin (2019), Quantization is similar to how we human brain encode and store information Gholami et al. (2022), and Knowledge Distillation as well. These ideas are only from one sub-field of Artificial Intelligence. It begs the question if we can find other similar ideas. Furthermore, it triggers other questions if the other way round could work: whether or not it is possible to use concepts in the space of algorithms in real-world applications.
The other motivation is the advent of Large Language Models. With the capability to promptly convert any idea expressed in natural languages into code or experiment, our next space of experimentation and computation is in the space of ideas.

3 PROPOSED STATEMENT

We humans, run algorithms in the space of ideas in our heads. And, there is a connection between the nature of humans and the type of algorithms we run.

4 DISCLAIMER

These are merely personal observations and are not ignorant of the works of Neuroscientists, Psychologists, or any other group working on the human mind.

5 DEMONSTRATION

Certain anecdotal marriages between two ideas one from the space of personalities and the other from the space of algorithms are presented here as a way of communicating the concept. For the

sake of simplicity, they are based on the famous introvert/extrovert classification of people which is highlighted below.

5.1 INTROVERT AND EXTROVERT

The Introvert and Extrovert classification of people is a simple tool to crudely classify people into two categories Freyd (1924). Surely, it is vague and not accurate. But it simplifies the classification down to binary. Basically, it says that people who are introverts tend to be inward-looking, inward-thinking, shy, number-crunching, and silent. And, extroverts are sociable, public speakers, and essentially the opposites of introverts.

5.2 SPECIFIC EXAMPLES

The following are specific examples of how the introvert and extrovert classification can be mapped to the world of algorithms.

- **Introverts think depth-first and extroverts think breadth-first**. A fundamental concept in algorithms is depth-first, where a task is done depth-wise and then progresses to the rest, and breadth-first approaches, which are the opposite. So the idea is that introverts think depth first and extroverts think breadth-first. Introverts prefer explanations that are deep. An extroverted person prefers a breadth-first explanation. They are looking for names, and might not even care about details. This is probably because they did too many conversations, unlike their counterparts, and they want to be efficient. This thinking pattern can explain why most introverts are builders, and extroverts are sellers. Conversely, we can run depth and breadth algorithms to mimic their behaviors to create better seller and better builder algorithms.
- **Machine Learning models are introverts**. So Machine Learning models do some kind of introspection and adjust themselves by iteratively looking at their internal parameters, termed called weights and biases, and adjusting them. This is typical of an introvert's behavior.
- **Introverts are exploiters and extroverts are explorers**. The explores/exploit tradeoff in Reinforcement Learning is where an AI agent decides to explore uncharted areas for optimal gains or explore the ones it already knows it will get some. This is much more relatable in real life with specific personalities. Some people stay in their comfort zone for too long while others are okay with jumping around and changing careers, courses, and even life partners.
- **Overthinking and Overfitting.** Ideally, a Machine Learning model is expected to generalize, not overfit. Overfitting is when a model remembers the training data so much that it doesn't care for what's new. Injecting noise, and removing some memory structures have been used as cures for overfitting. The idea is quite similar to overthinking where a person thinks about something so much about something that it becomes harmful to her/himself.

6 CHALLENGES

Clearly, the examples discussed above are simple conjectures or opinions. The concept is required to be substantiated by solid research. The challenges to do that are in general computational and psychological. Computational challenges include formulating computational units, representations, experimentation, benchmarks, etc. Psychological challenges include the formulation of stable personality classification frameworks, identifying basic units of ideas, etc.

7 BENEFITS

There is a lot of usefulness to the concept if it is established well. In general, one can benefit the other: the world of algorithms helps with problems related to personality, and we might come up with better algorithms based on personality. But there are specific benefits that can be mentioned here. The following are some of them.

- **Better self-awareness.** The world is eventually recognizing the importance of emotional intelligence and self-awareness. There are thousands of books written about self-help, self-development, and related topics. Approaching the self-awareness problem algorithmically will be a valuable step forward.
- **Algorithmic Psychotherapy.**A far-fetched but worthy-of-discussing idea is Algorithmic Psychotherapy, an idea that tries to help therapists use algorithms to heal trauma. For example, one can probably address overthinking through the solutions suggested for overfitting models. Another example is giving a depth-first problem to an extroverted person to get him out of his thinking pattern and heal him of trauma. This can potentially be a less pervasive alternative for Nueralink.
- **Understand others better**. This will also help us understand others better which will improve our communication and relationship with others to potentially solve problems. For example, schools might use it to identify a student's strengths and weaknesses and help them accordingly.
- **Innovation**. Running human-like generative algorithms can discover things that wouldn't be discovered otherwise. For example, the stories generated by an algorithm with such awareness can be more humane. We can also get a better sales algorithm.
- **Meta**. This is the ultimate meta approach towards anything.

8 PITFALL

Manipulation (Mass Manipulation) is an issue that can be considered from start.

9 CONCLUSION

Entirely out of personal observation, a simplified link between the world of algorithms and personalities has been introduced in an attempt to benefit the study of the mind from the thriving discipline of algorithms. But also, the other way around. The idea, which doesn't sound absurd considering the advent of current progress in the space of algorithms and Large Language Models, can serve as a motivation for a well-established framework that can potentially lead to the unification of the two areas.

REFERENCES

Jonathan Frankle and Michael Carbin. The lottery ticket hypothesis: Finding sparse, trainable neural networks. *arXiv: Learning*, 2019.

Max Freyd. Introverts and extroverts. *Psychological review*, 31(1):74, 1924.

Amir Gholami, Sehoon Kim, Zhen Dong, Zhewei Yao, Michael W. Mahoney, and Kurt Keutzer. A survey of quantization methods for efficient neural network inference. *ArXiv*, abs/2103.13630, 2022.