

.....

## Aristotle the Ethical Agent

Aayush Bajaj

February 26, 2025

.....

I think I could do it. I think I could build on the discovery of *Deep Learning* and produce an artificial agent that behaves ethically.

I think Kant is historical and he never knew Bayesian Probability. I also think many of the papers that I studied for Essay 1 — whilst they were written after the seminal Attention is All You Need paper by Google — they never realised the impact of *truly* deep learning.

All of this is very speculative at the moment, but given a month to train on UNSW's HPC Katana, and time within which to study Virtue Ethics, Utilitarianism, Deontological ethics and the mathematical unions of their moral behaviours, I believe our Neural Network can learn the arbitrarily complex decision surface that would compose *ethical* decisions.

Furthermore, I intend to study a Master's in Statistics, which is accompanied by research work towards the end of the post-graduate degree. I wonder if creating "Aristotle, the Ethical Agent" would be something my school would approve. And if so, combining my understanding of Probability with my aptitude in Pytorch should yield an Agent that is in fact Kantian — by way of leveraging the *likelihood* of immoral actions to be low, and thus technically allowing the agent to "pick" in the same way a human would — thus enabling autonomy and the existence of a Kantian AMA!

In conclusion, it is only week 3 and I am likely talking out of my ass;

Que sais-je?
--------------

— Michel de Montaigne
-----------------------

Aayush