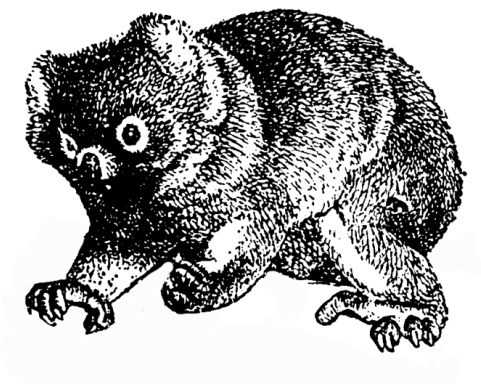
A Soft Introduction to zkML

Lucas A. Vittor

lucasvittor.com, @lvvittor

August 26, 2023



Warm Up

1. Who works in Web3 as a developer/researcher?
2. Who works in a data-first company?
3. Who works with Machine Learning (ML)?
4. Who knows what Zero Knowledge Proofs (ZKPs) are?
5. What is the name of the set of problems that can be solved in the $\mathcal{O}(n^k)$ time complexity type?

What are Proofs?

Def. Mathematical Proof

A deductive argument for a proposition, showing that the stated assumptions logically guarantee the conclusion.

Common methods of proof:

1. Direct
2. Contradiction
3. Contraposition
4. Induction

Proofs as Static Objects

Verifier

Given the proposition $\mathbb{X}$

$axioms \implies \ldots$

$\implies conclusion$

Proof (Π)

Prover

Example of a Proof

$\sqrt{2}$ is irrational

Assume $\sqrt{2}$ is rational, so $\sqrt{2} = \frac{a}{b}$ in simplest form.

Squaring both sides yields $2 = \frac{a^2}{b^2}$, implying $a^2 = 2b^2$.

Since a^2 is even, a must be even (odd² is odd).

Let $a = 2k$, then $4k^2 = 2b^2$, giving $2k^2 = b^2$.

Now, both a and b are even, contradicting their coprimality.

Thus, our assumption is false, and $\sqrt{2}$ is irrational. ■

How about efficiency?

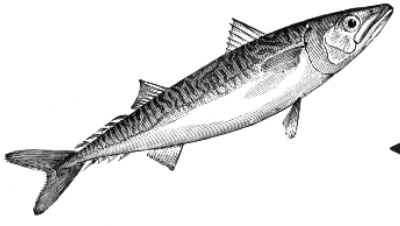
Given a valid proof Π about a proposition $\mathbb{X}$,
how *"easy"* is to convince me it's true?

An Intuition About Interactive Proofs (IPs)

1. ***True*** *propositions are* ***Provable***
2. ***False*** *propositions* ***NOT***

Visualizing IPs

Verifier
(Lazy)



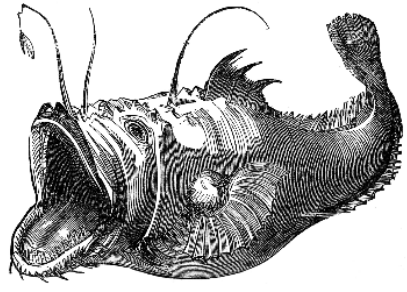
$\mathcal{O}(n^k)$

Given the proposition $\mathbb{X}$

Messages exchange
(several rounds)

Probabilistic Proof (Π')

Prover
(Hard Working)



∞

Defining Knowledge

Q: *How much knowledge to verify a proof?*

Q: *What is knowledge?*

Zero Knowledge Proofs (ZKPs)

”Nothing but the truth”

Formal Definition of ZKPs

Def. Zero Knowledge Proof

Suppose $\mathcal{L}$ is a language. A zero-knowledge protocol is an interaction between two algorithms P and a probabilistic polynomial time (PPT) algorithm V with P trying to convince V that $x \in \mathcal{L}$ and the satisfying properties:

(i) *Completeness*

(ii) *Soundness*

(iii) *Zero-knowledge*

ZK Properties Definition

Completeness

If the proposition $\mathbb{X}$ is **true** and V, P are honest $\implies$ V will be convinced.

Soundness

If the proposition $\mathbb{X}$ is **false** $\implies$ $\nexists$ cheating P who can convince an honest V that $\mathbb{X}$ is true (except with some small probability).

Zero Knowledge

$\forall x \in \mathcal{L}, z \in \{0,1\}^*, View_V[P(x) \longleftrightarrow V(x,z)] = S(x,z)$

The Two Balls Problem

Proposition $\mathbb{X}$:

I have two balls of different colors.

Soundness Score

k	$1/2^k$
1	0.5
2	0.25
3	0.125
4	0.0625
5	0.03125
6	0.015625
$\vdots$	$\vdots$

Where We Can Use ZKPs?

One proposition?

Some propositions?

Special propositions?

Machine Learning (ML)

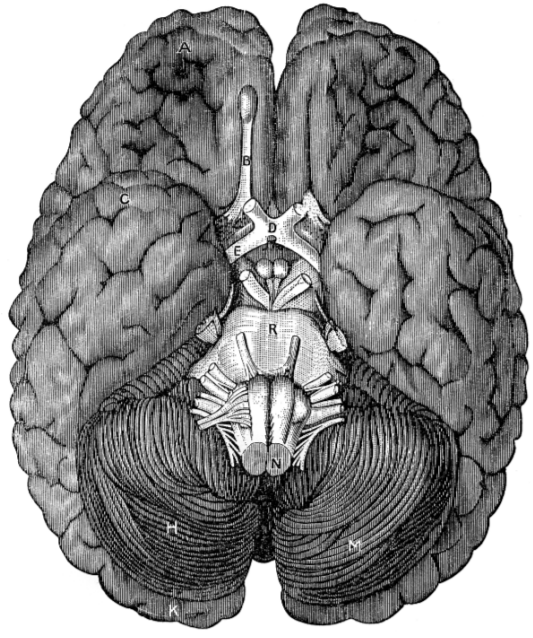
Which option is correct?

1) $\mathbb{AI} = \mathbb{ML}$

2) $\mathbb{AI} \neq \mathbb{ML}$

What is AI?

AI $\geq$

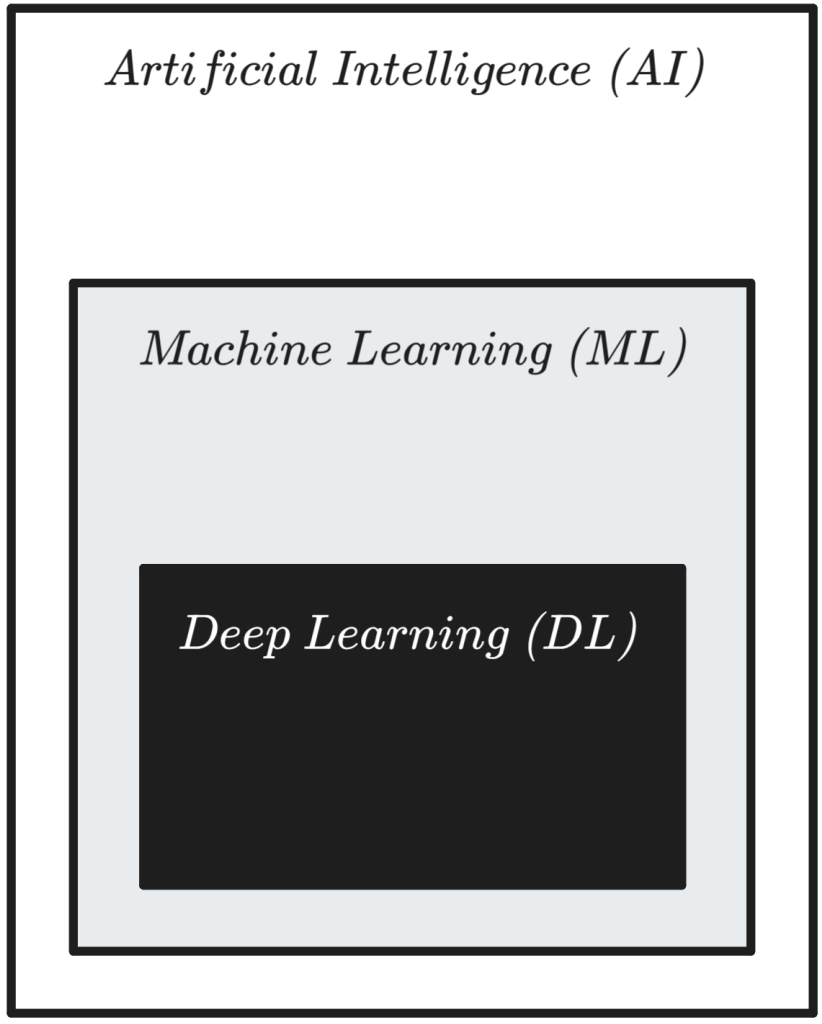


What is ML?

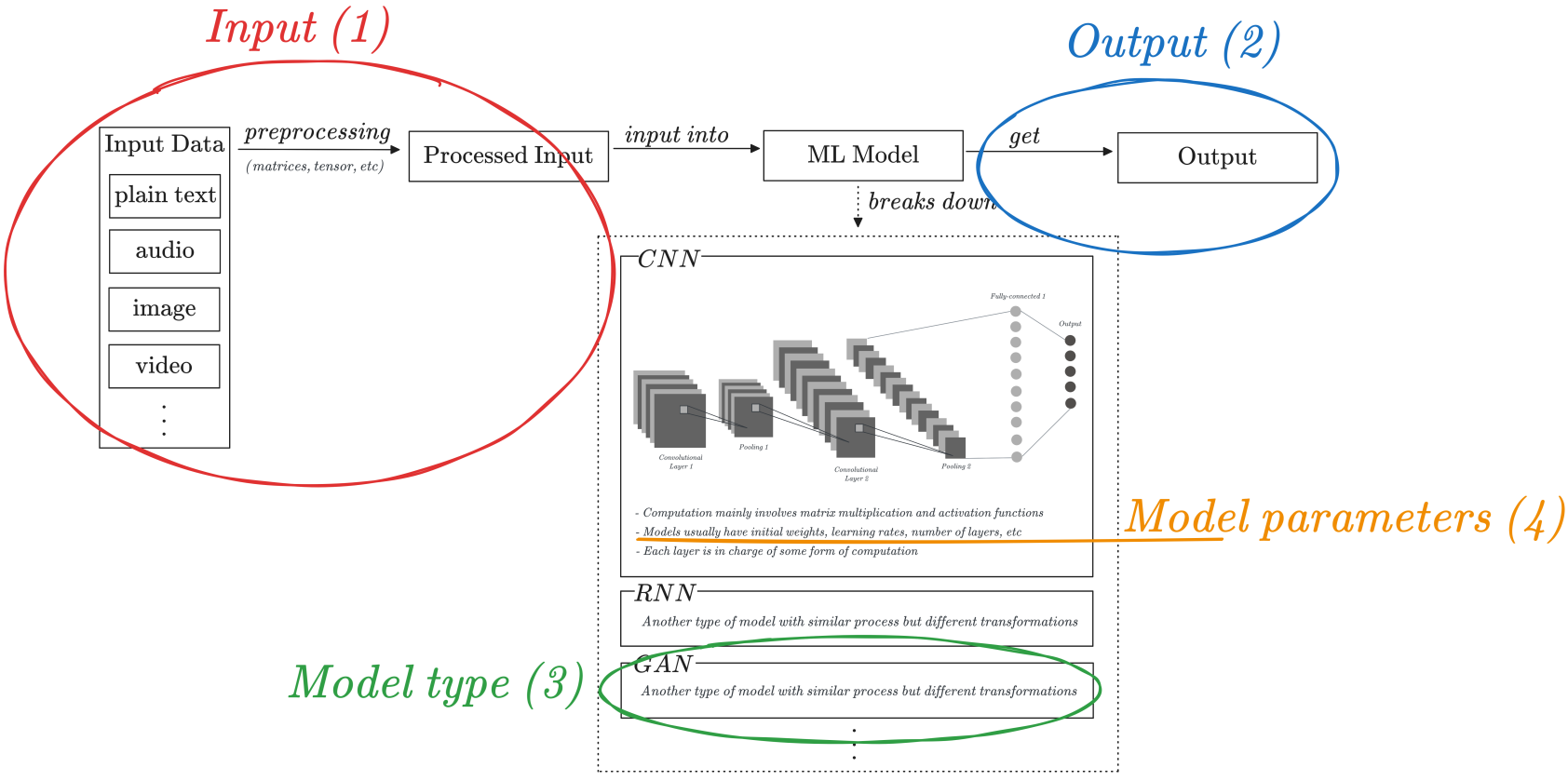
A computer program is said to learn from experience ***E*** *with respect to some task* ***T*** *and some performance measure* ***P****, if its performance on* ***T****, as measured by* ***P****, improves with experience* ***E****. — Tom Mitchell, 1997*

It is the adaptation of the free parameters of a system in order to satisfy an object function. — J. M. Santos

Venn Diagram



ELI5 ML


Input (1)
Output (2)
Input Data
plain text
audio
image
video
preprocessing
(matrices, tensor, etc)
Processed Input
input into
ML Model
get
Output
breaks down
CNN
Fully-connected 1
Output
Convolutional Layer 1
Pooling 1
Convolutional Layer 2
Pooling 2
- Computation mainly involves matrix multiplication and activation functions
- Models usually have initial weights, learning rates, number of layers, etc
- Each layer is in charge of some form of computation
Model parameters (4)
RNN
Another type of model with similar process but different transformations
GAN
Another type of model with similar process but different transformations
Model type (3)

Issues with State Of the Art ML

1. *Black-box nature of NNs*

2. *Privacy*

$$zk + ML = \heartsuit$$

Web3 Tales

Making web3 uncool again @ web3tales.com