Empirical AIs

Experience

Aristotle and John Stuart Mill

Empirical Approach

- AI Systems designed to learn through experience to improve outputs (predictions).

- Prediction
  - Process by which an AI Model, trained on **historical** data, uses learned patterns to make an informed estimate about a future or unknown outcome based on **new** input data.

Empirical Learning

- The process where a system improves its knowledge or performance by analyzing past experiences or observed data, identifying patterns, and generalizing them to make future predictions.

Just a few years ago…

AlphaGo – March 2016 -- Seoul

First AI to defeat a world champion of the game GO

- More possible moves than Chess
- 19 by 19 Board
- ~ 10 ^ 1048

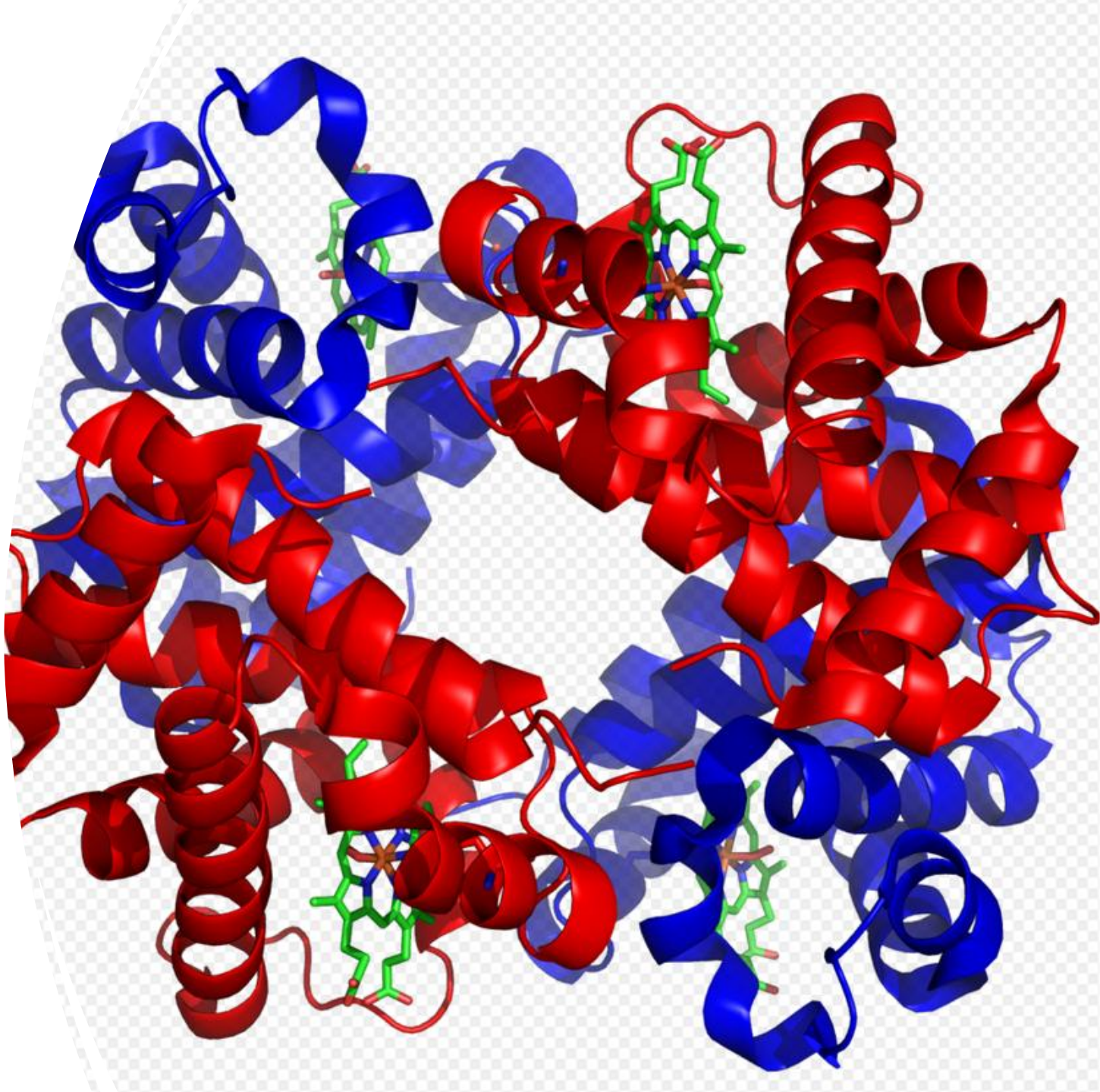
Learned by analyzing millions of human played games

Reinforced by playing millions more against itself

It’s All Fun and Games….

AlphaFold 2020

- AI that predicts 3D protein structures from amino acid sequences.
- Protein structure determines its function and interactions
- Traditional methods identified about 227,000 by 2025
- AlphaFold ~ 214 million predicted



“Unfold” the Definition Further

Empirical Learning is the process where algorithms improve performance by analyzing input-output pairs or data patterns.

The system extracts relationships directly from data to generalize beyond specific cases.

Empirical because knowledge comes from experience not abstract reasoning alone

Input / Output Pairs

Spam Filtering

- Input: the email's text and metadata
- Output: Label = Spam or Not Spam

Type Ahead

- The word "Coffee" is frequently followed by 'cup', 'shop', 'drink'
- AI use statistical co-occurrence patterns to predict

Errors

Difference between the model's prediction and actual truth (label)

Image labeled "Dog" when it is a "Cat"

Classification error

Training on a dataset of labeled photos tunes the model to reduce errors

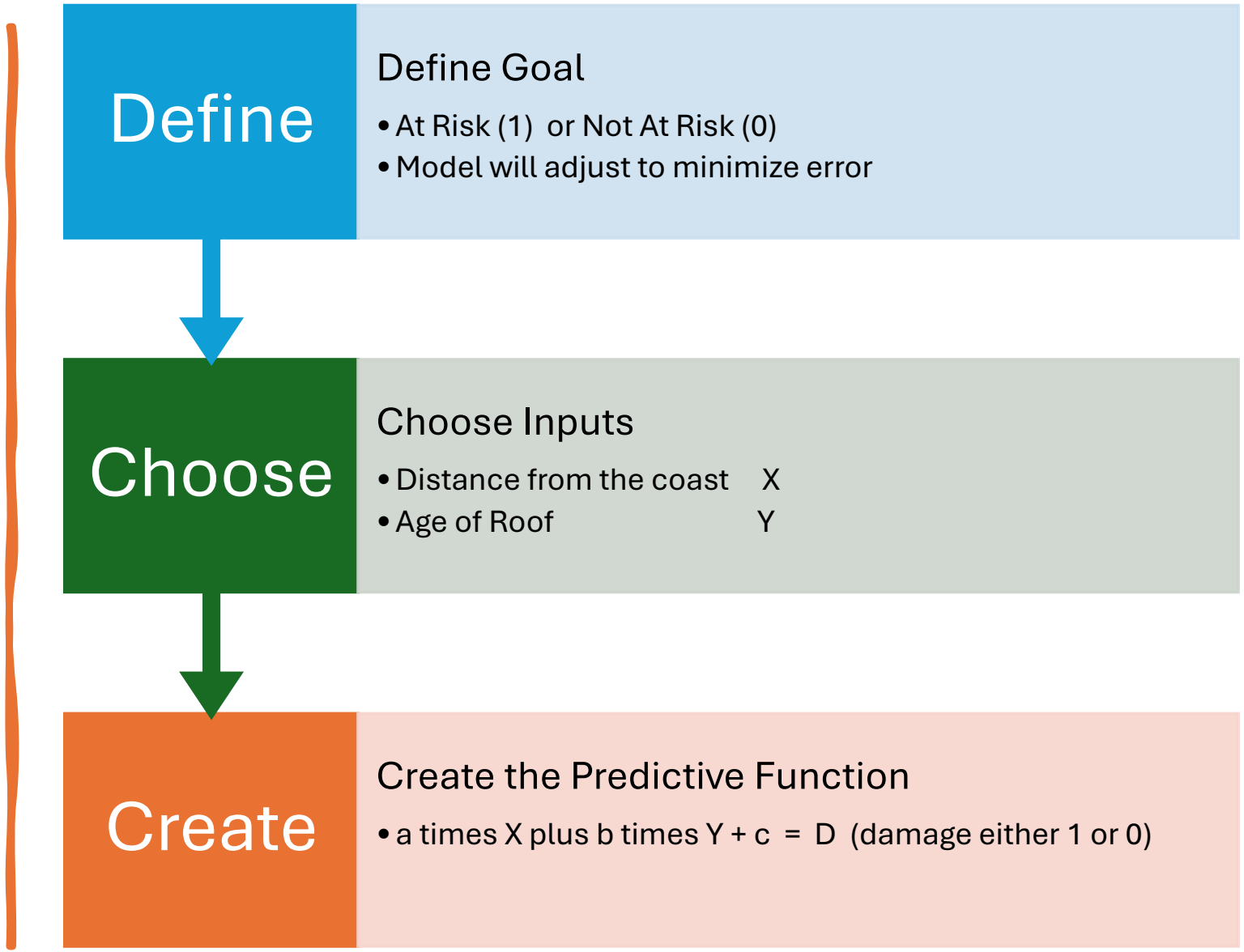
Let’s Try It Out – Closer to Home

Output desired:

How likely is it that a house in Brunswick area will be damaged by high winds?

Identify Two Factors that would help you predict damage

The Process


Define
Define Goal
• At Risk (1) or Not At Risk (0)
• Model will adjust to minimize error
Choose
Choose Inputs
• Distance from the coast X
• Age of Roof Y
Create
Create the Predictive Function
• a times X plus b times Y + c = D (damage either 1 or 0)

Enter the Training Data – Minimize the Error

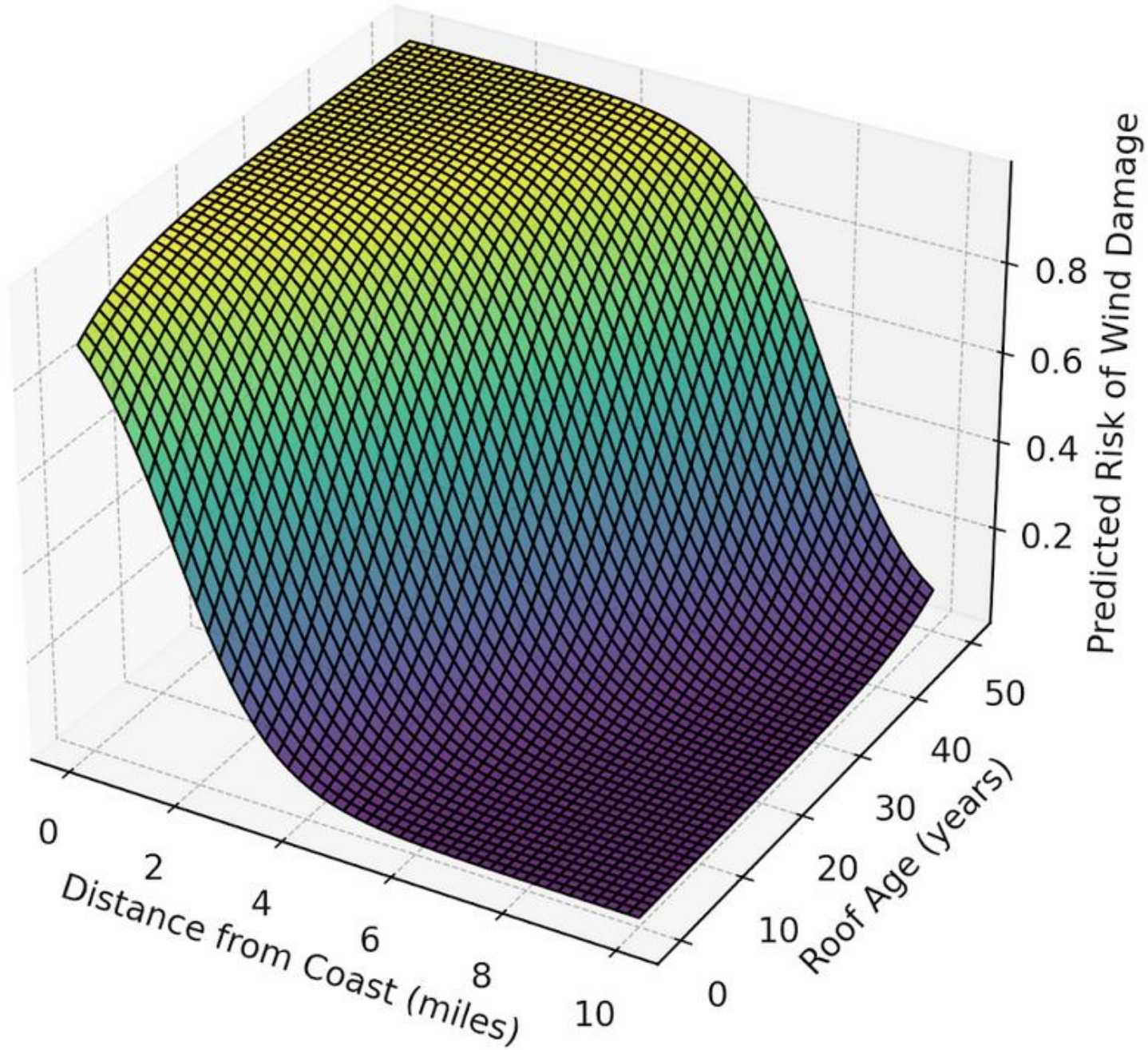
Data is houses in Brunswick with known storm outcomes

- Distance is known
- Damage is known

Adjust a, b, c to the predicted risk is as close as possible to the observed outcome.

- Nerd Alert: Gradient Descent is the method

Model learns the best a, b, c (weights)


Predicted Risk of Wind Damage
0.2
0.4
0.6
0.8
Roof Age (years)
0
10
20
30
40
50
Distance from Coast (miles)
0
2
4
6
8
10

Improving

Brunswick House Example = 2 parameters and 3 weights

GPT-4 and variants range from between 100 billion and 300 billion parameters

Research suggests surpassing 1 trillion

LLM Terms

- Token
  - The smallest unit of data that a model processes. Typically representing words, sub words, characters, or symbols in text. Input

- Embeddings
  - Numerical vectors that map high dimensional data in vector space.

- Vector Space
  - Arrays of numbers allowing operations like addition and scalar multiplication. Similar items are positioned closer together to permit analysis

LLM Terms - 2

- Parameters or Weights
  - Internal variables with the model that are adjusted during training to capture how tokens relate to each other.
  - The internal tunable elements to allow model to analyze, predict, and generate outputs from tokens

Consider

What are the Advantages of this method?

What are the Disadvantages?

How does this inform your perception of the public policy issues?

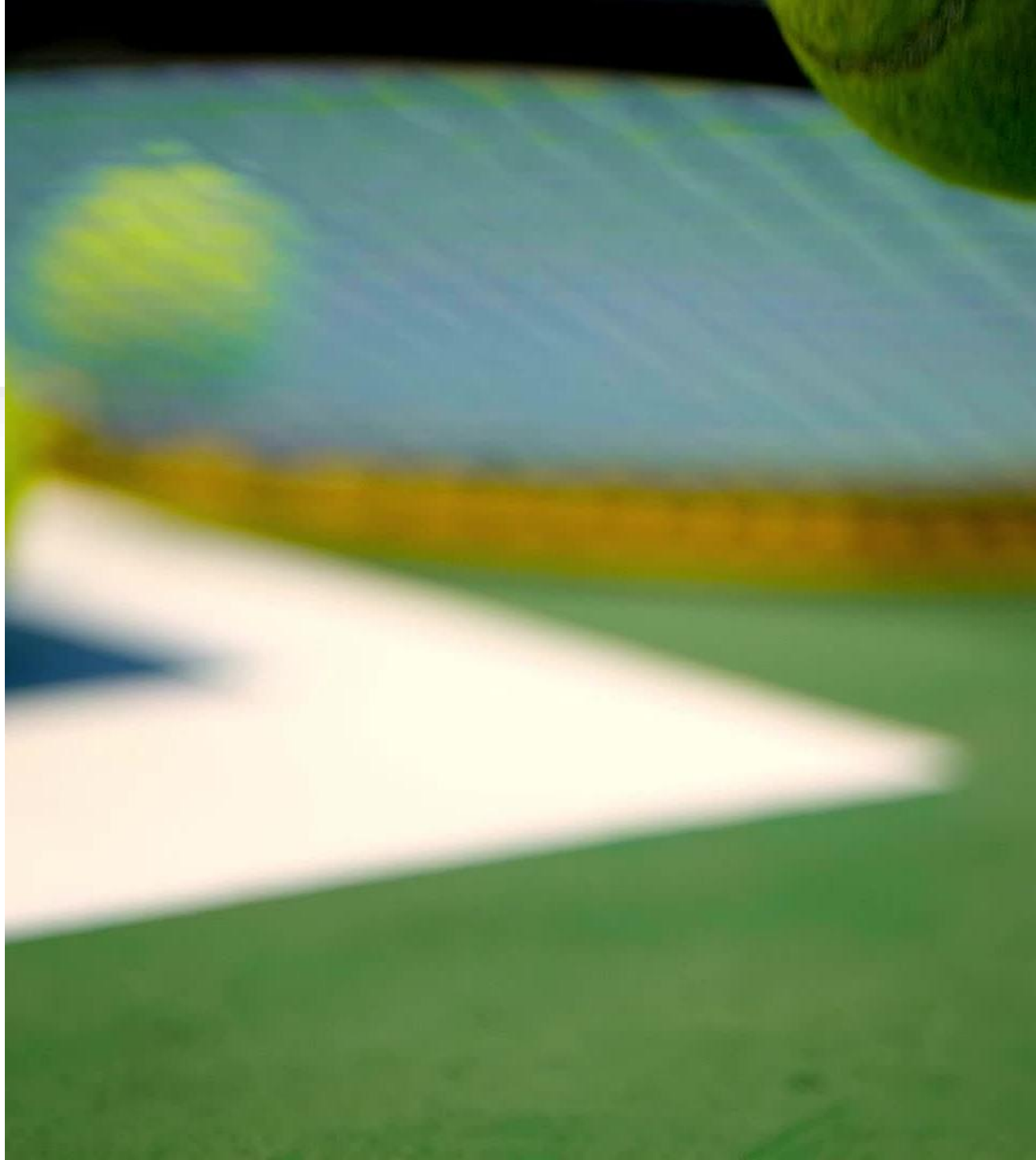
A Note of Caution (Again)

Thinking Fast: This is the brain's fast, intuitive, and automatic mode of thinking. It operates quickly, with little or no effort, and without voluntary control. System 1 is responsible for immediate reactions, pattern recognition, and gut feelings.

Thinking Slow: This is the brain's slower, more deliberate, and analytical mode of thought. It involves conscious reasoning, effortful mental activities, and complex computations. System 2 is used for problem-solving, logical analysis, and tasks requiring focus and attention.

Illustrate

A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?



Check Your Math

The cost of the ball = x

The cost of the bat = x plus $ 1

The total cost is x plus (x plus $ 1) = $ 1.10

Combine Like Terms 2x plus $ 1 = $ 1.10

Subtract $ 1.00 from each side 2x = $ 0.10

Why

RATIONAL MODELS HAVE THEIR PLACE.

EMPIRICAL MODELS HAVE THEIR PLACE.

IMPORTANT TO APPRECIATE HOW OUR MINDS WORK AS WELL AS THESE ALIEN MINDS.

Next Up – Session 4

- Neural Networks
- Training
- Alignment
- World Models