# The State of AI

## INNOVATIONS, APPLICATIONS, AND IMPACTS IN PLANNING

J. Langley
Founder, Huntsville AI
CTO, CohesionForce, Inc

Josh Phillips
Data Group Working Lead, Open Model Initiative
CDO, CohesionForce, Inc

HSV.AI

## INTRODUCTION TO HUNTSVILLE AI

- Our vision is a group of individuals and organizations in the metro Huntsville area who collaboratively advance the knowledge and application of artificial intelligence in ways that make it available to everyone and improve our quality of life.

https://hsv.ai/subscribe/

HSV.AI

# How did we get here?

## INSANELY FAST HISTORY LESSON

**We're going to move fast through the next 25 slides, so here's a link to download if you want to circle back later :)**

# POP QUIZ

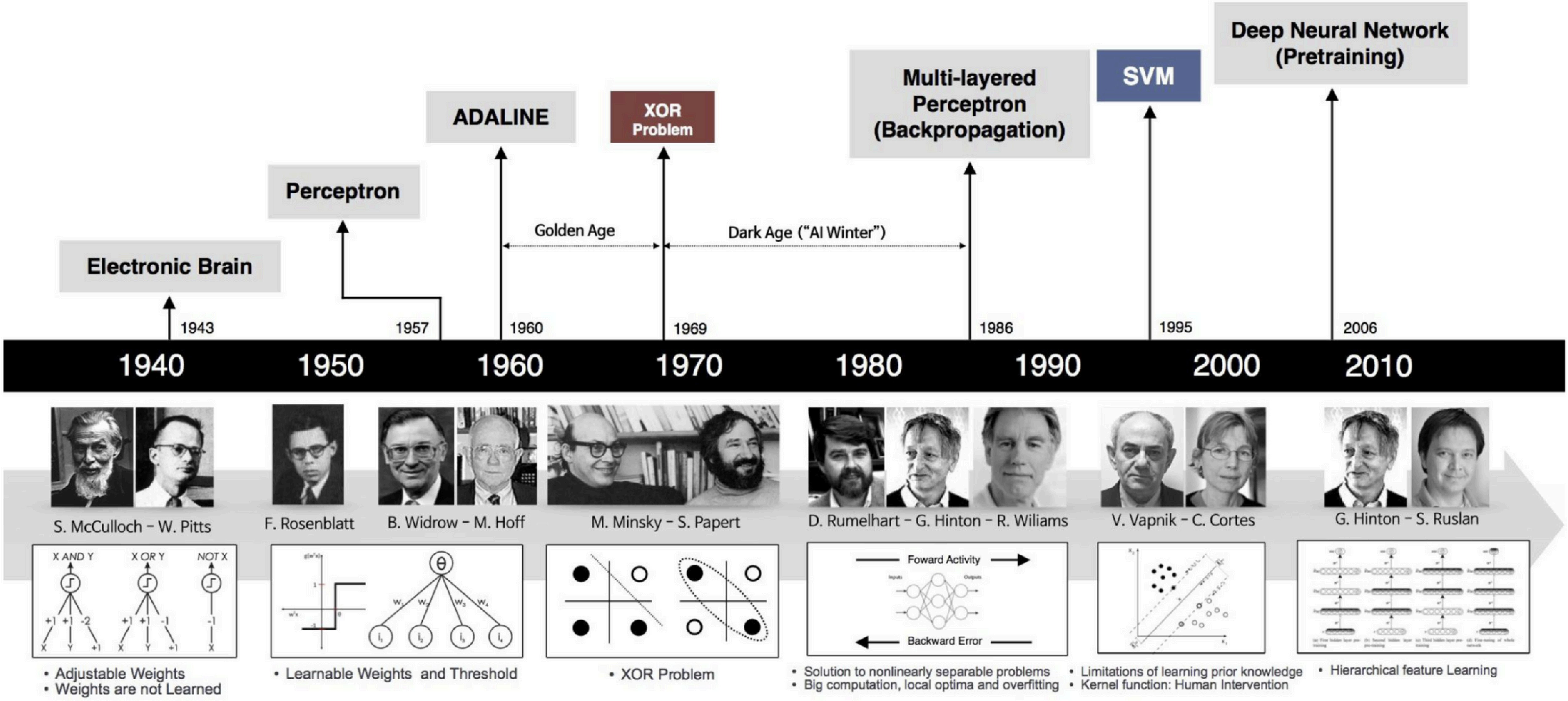
What is the first mention in writing of a “Thinking Machine”?

- **1950**
- **1945**
- **1936**
- **1920**
- **1863**

## POP QUIZ

What is the first mention in writing of a “Thinking Machine”?

- **1950 - Computing Machinery and Intelligence”, Alan Turing**
- **1945 - "As We May Think", Vannevar Bush**
- **1936 - “On Computable Numbers, with an Application to the Entscheidungsproblem”, Alan Turing**
- **1920 - “R.U.R”, Karel Capek (Rossum's Universal Robots)**
- **1863 - “Darwin Among the Machines”, Samuel Butler**



Deep Neural Network (Pretraining)
SVM
Multi-layered Perceptron (Backpropagation)
ADALINE
XOR Problem
Perceptron
Golden Age
Dark Age ("AI Winter")
Electronic Brain
1943
1957
1960
1969
1986
1995
2006
1940
1950
1960
1970
1980
1990
2000
2010
S. McCulloch – W. Pitts
F. Rosenblatt
B. Widrow – M. Hoff
M. Minsky – S. Papert
D. Rumelhart – G. Hinton – R. Wiliams
V. Vapnik – C. Cortes
G. Hinton – S. Ruslan
X AND Y
X OR Y
NOT X
Foward Activity
Backward Error
• Adjustable Weights
• Weights are not Learned
• Learnable Weights and Threshold
• XOR Problem
• Solution to nonlinearly separable problems
• Big computation, local optima and overfitting
• Limitations of learning prior knowledge
• Kernel function: Human Intervention
• Hierarchical feature Learning

HSV.AI

# ALAN TURING (1950)

Alan Turing's "Computing Machinery and Intelligence," published in 1950 in the journal Mind.

The Imitation Game (Turing Test):
Turing proposed the "Imitation Game" (later known as the Turing Test) as a way to operationalize the question, "Can machines think?" The test involves a human evaluator engaging in a text-based conversation with a machine and another human, without knowing which is which. If the evaluator cannot reliably distinguish the machine from the human, the machine is said to exhibit intelligent behavior.

The Turing Test has been a central topic in debates about AI's capabilities and the nature of intelligence.

# JOHN MCCARTHY (1956)

The title "Father of Artificial Intelligence" is most commonly attributed to John McCarthy, an American computer scientist. He is credited with coining the term "Artificial Intelligence" in 1956 during the Dartmouth Conference, which is widely regarded as the founding event of AI as a field of study.

We'll run into Mr McCarthy again in a few slides :)

# ROSENBLATT’S PERCEPTRON (1958)

The perceptron algorithm was invented in 1957 at the Cornell Aeronautical Laboratory by Frank Rosenblatt, funded by the United States Office of Naval Research.

The perceptron was intended to be a machine, rather than a program, and while its first implementation was in software for the IBM 704, it was subsequently implemented in custom-built hardware as the "Mark 1 perceptron".

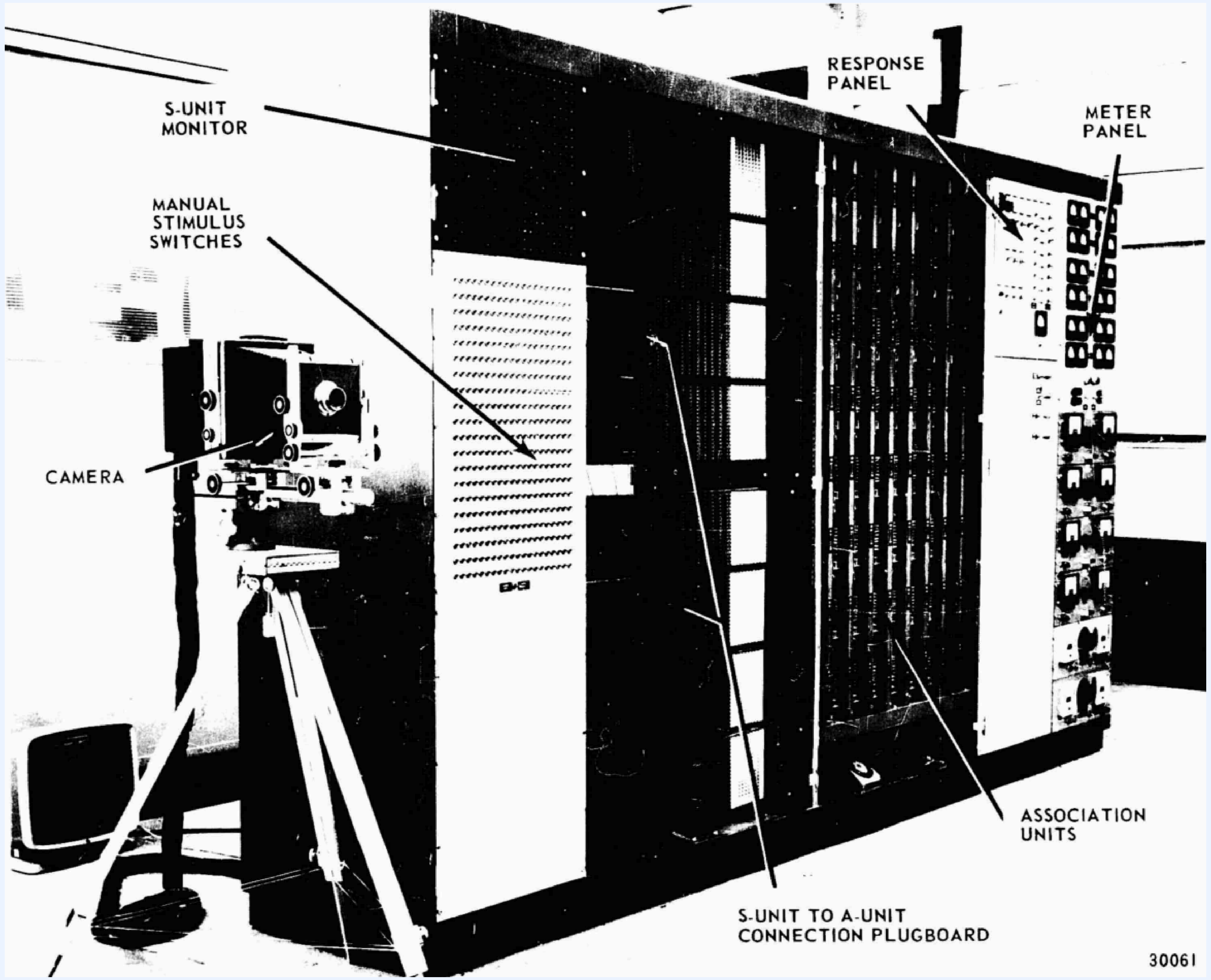
This machine was designed for image recognition: it had an array of 400 photocells, randomly connected to the "neurons". Weights were encoded in potentiometers, and weight updates during learning were performed by electric motors.

https://homepages.math.uic.edu/~lreyzin/papers/rosenblatt58.pdf

# ROSENBLATT'S PERCEPTRON (1958)

It utilized vacuum tubes as well as resistors, capacitors, and other electronic components.

The perceptron was built with 512 adjustable weights, implemented using motorized potentiometers, which allowed the system to "learn" by adjusting these weights based on input-output relationships.




HSV.AI

# ROSENBLATT'S PERCEPTRON (1958)

The design of the perceptron was based on how the brain was thought to work.
This is the first instance of what would later become neural networks.

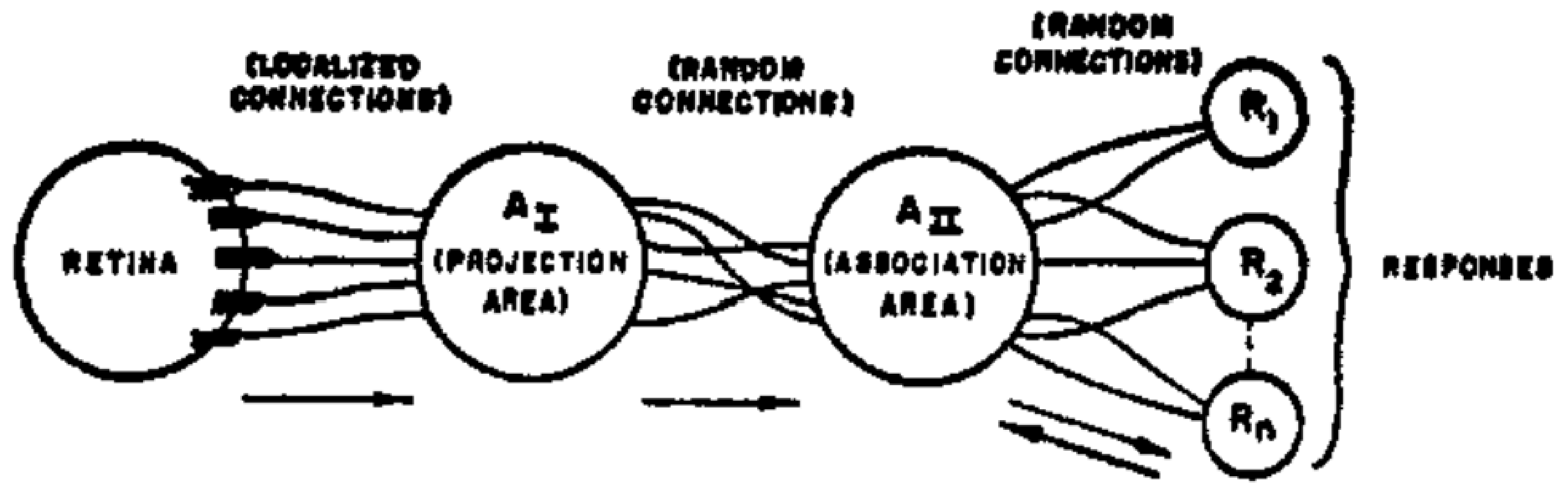



FIG. 1. Organization of a perceptron.

## GOLDEN AGE

- General Problem Solver (1959):
  - Developed by Herbert Simon and Allen Newell, it was designed to solve a wide range of problems by breaking them into smaller sub-problems.
- A* Algorithm (1968):
  - Created by Peter Hart, Nils Nilsson, and Bertram Raphael, A* became a foundational algorithm for pathfinding and graph traversal.
- Lisp (1958):
  - Created by John McCarthy, Lisp became the first programming language specifically designed for AI.
- Prolog (1969):
  - Prolog, developed by Alain Colmerauer and Robert Kowalski, became the foundation for logic programming and expert systems.

# GOLDEN AGE

- ELIZA (1966):
  - Developed by Joseph Weizenbaum, ELIZA simulated a Rogerian psychotherapist, marking one of the first successful attempts at conversational AI.
- SHRDLU (1968):
  - Terry Winograd created SHRDLU, which could understand and execute commands in a simulated blocks world, advancing NLP and contextual reasoning.
- Automata Theory:
  - Research into automata and formal languages, led by Noam Chomsky and others, influenced AI by formalizing computational processes.
- Logic and Reasoning:
  - Advancements in formal logic, particularly predicate calculus, became integral to AI's ability to model knowledge and reasoning.

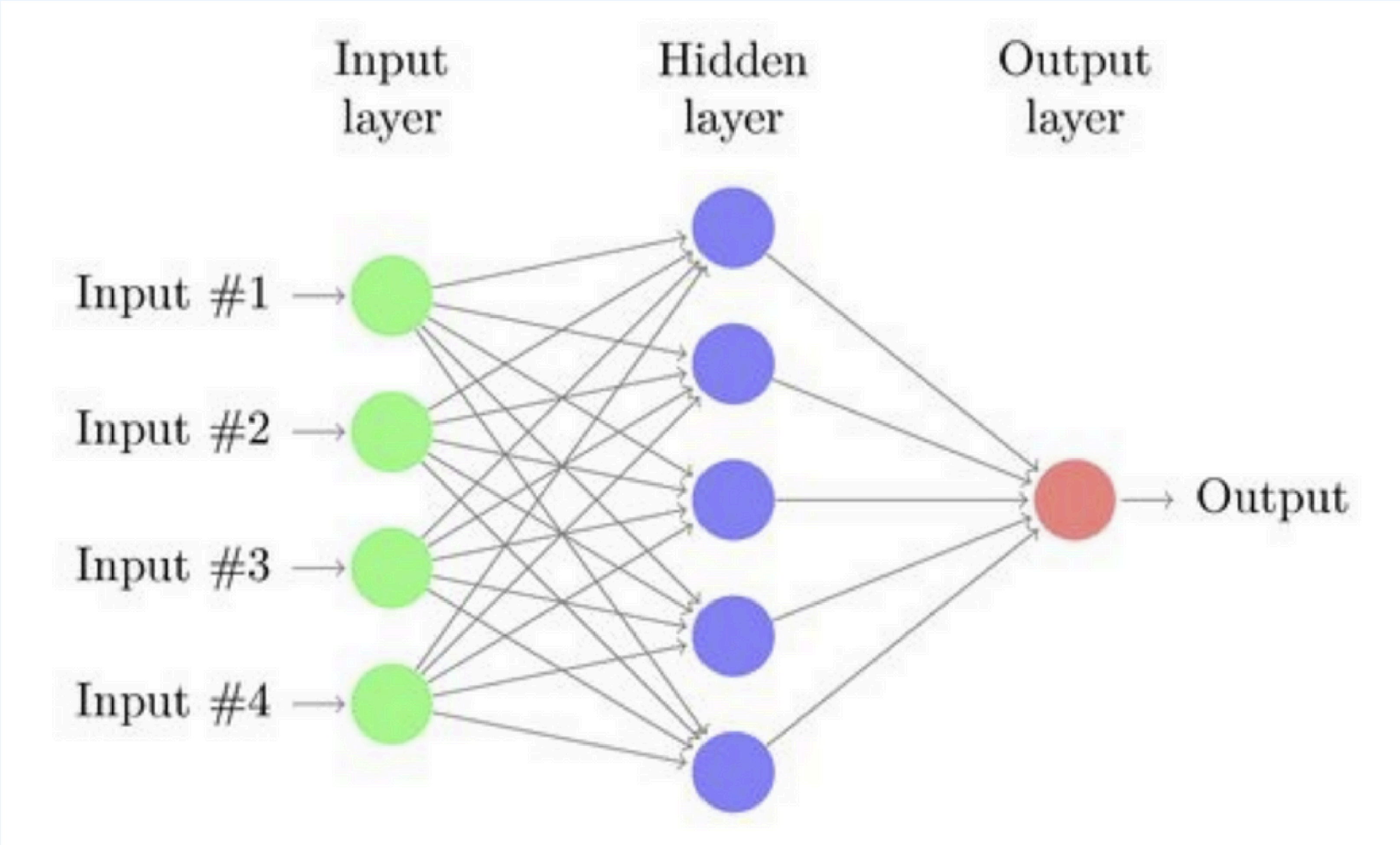
# AI WINTER (1969)

Perceptrons: an introduction to computational geometry by Marvin Minsky and Seymour Papert - published in 1969.

It offered a mathematical proof that the perceptron could not approximate an XOR function given an infinite training set.

# MULTI-LAYER PERCEPTRONS

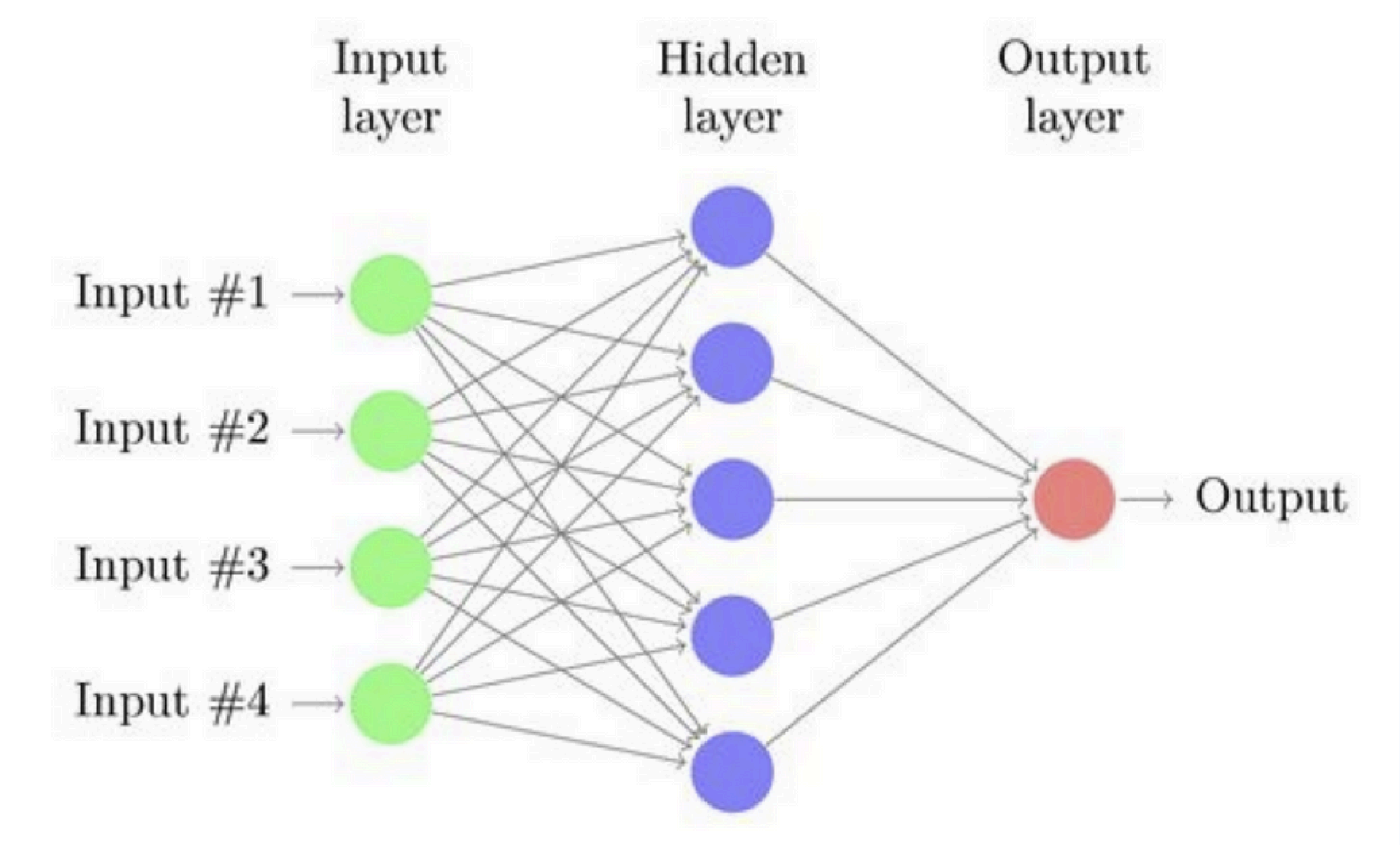
By stacking several layers of perceptrons, researchers were able to overcome the XOR problem.




But it was nearly impossible to train!

# BACKPROPAGATION (1986)

Geoff Hinton, along with David Rumelhart and Ronald Williams, published a paper entitled "Learning representations by back-propagating errors"




Loss is computed at the output based on difference between it and the known answer.

Weights are updated moving backward through the neural network.

Caution - there's a lot of math here

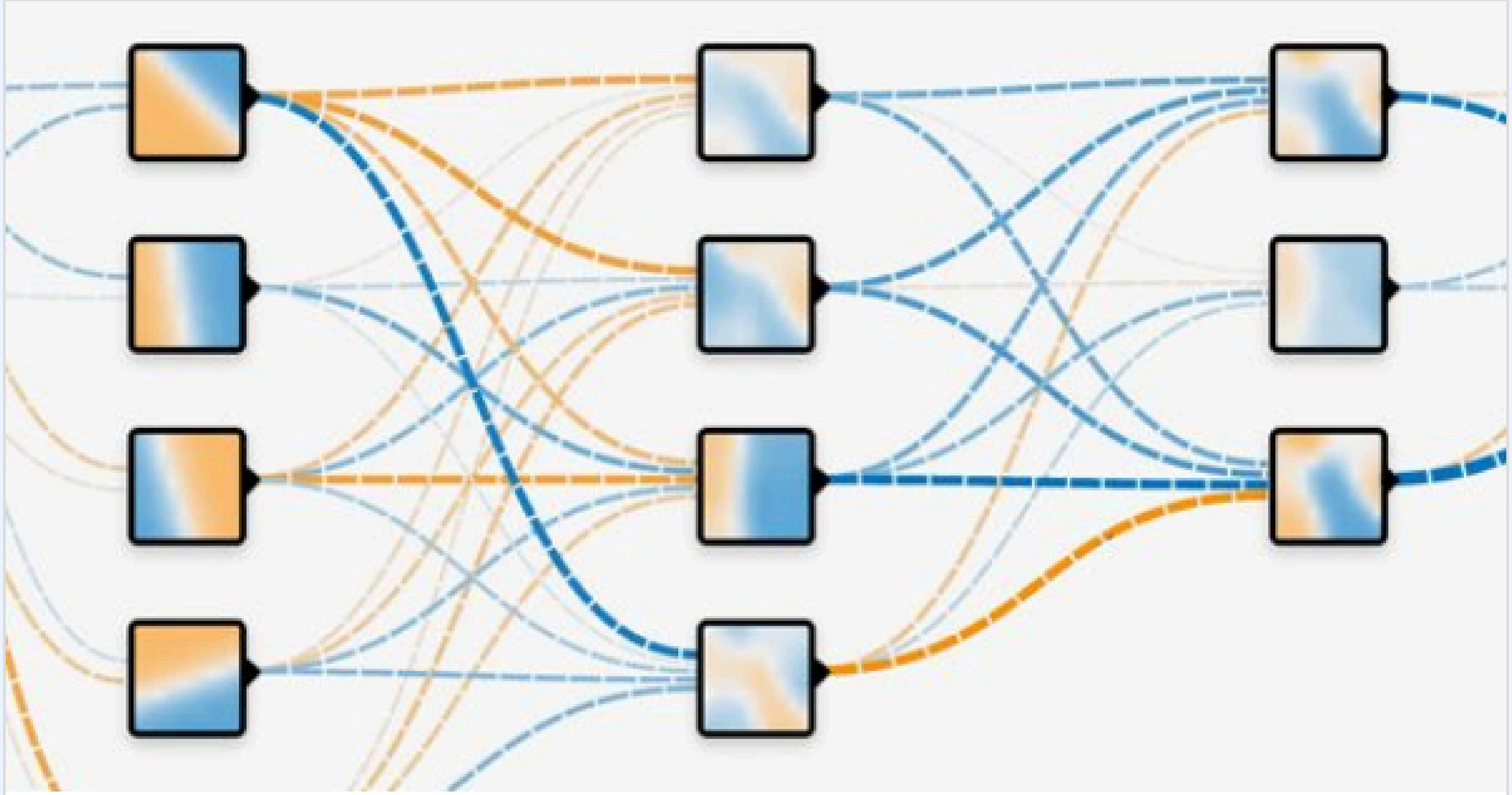
HSV.AI

# UNIVERSAL APPROXIMATION THEOREM (1989)

The **Universal Approximation Theorem** is a fundamental result in the theory of neural networks. It states that a feedforward neural network with at least one hidden layer, a sufficient number of neurons, and a suitable activation function can approximate any continuous function to any desired degree of accuracy, given appropriate weights and biases.

Generally credited to papers by George Cybenko (1989) and Kurt Hornik et al. (1989)

# NEURAL NETWORKS



**Tensorflow — Neural Network Playground**

Tinker with a real neural network right here in your browser.

tensorflow.org /

[Tensorflow Playground]

# LEARNING BEFORE “DEEP” (1998)

Gradient Based Learning Applied to Document Recognition - Yann Lecun -
CNN from Yann Lecun (AT&T Bell Labs) could recognize handwritten digits.

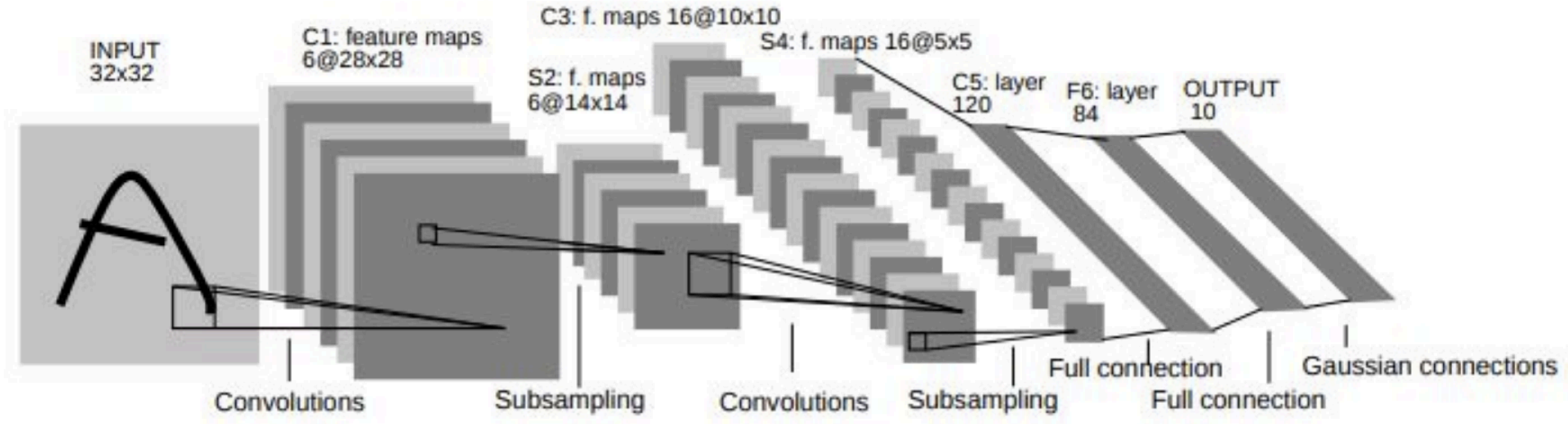



Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

# ROLLING IN THE DEEP (2006)

Deep Learning (2006) - Again with Geoff Hinton. The idea was to train a simple 2-layer unsupervised model like a restricted boltzman machine, freeze all the parameters, stick on a new layer on top and train just the parameters for the new layer. Using this strategy, people were able to train networks that were deeper than previous attempts, prompting a rebranding of ‘neural networks’ to ‘deep learning’.

## ACCELERATED BY HARDWARE AND DATA

It takes a lot of data to train a deep neural network:

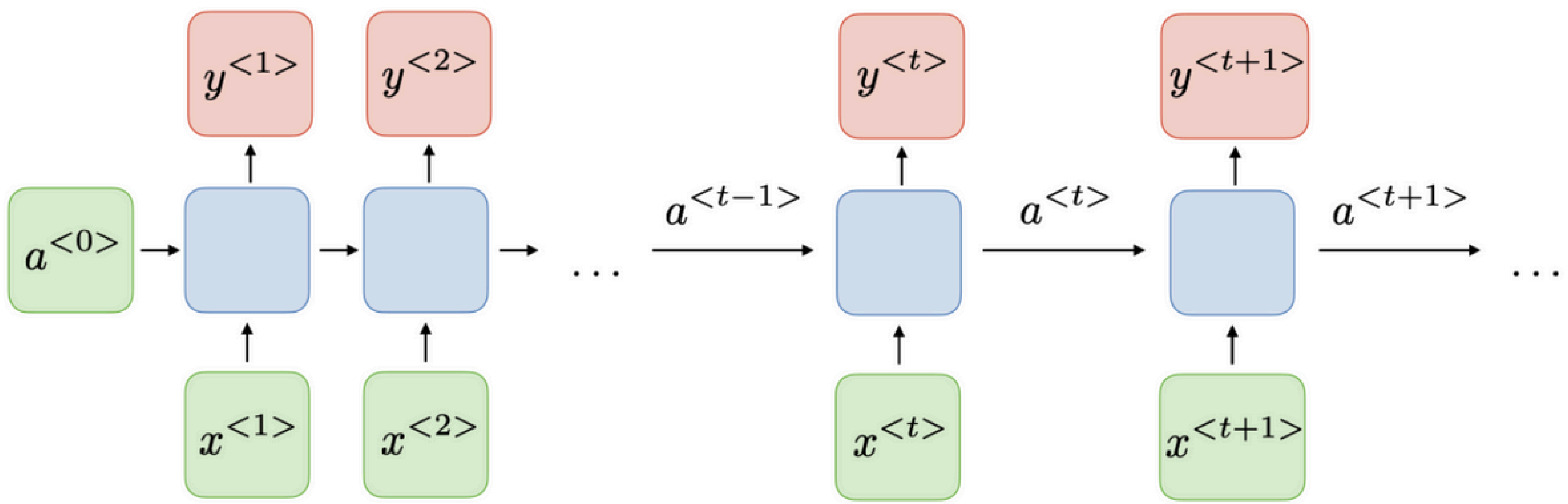
- Imagenet (2009) - millions of labeled images created and published by Fei-Fei Li at Stanford
- MNIST (1999) - Handwritten digits
- Google House Numbers from street view (2014) - created by an intern at Google (Ian Goodfellow)
- Flickr 30k Image dataset (2014)

You also need some good hardware at doing simultaneous calculation:

- GPU - used for multi-core floating point calculation

# RECURRENT NEURAL NETWORKS (2010 - 2017)

Recurrent neural networks, also known as RNNs, are a class of neural networks that allow previous outputs to be used as inputs while having hidden states. They are typically as follows:




For each timestep $t$, the activation $a^{<t>}$ and the output $y^{<t>}$ are expressed as follows:

$$\boxed{a^{<t>} = g_1(W_{aa}a^{<t-1>} + W_{ax}x^{<t>} + b_a)} \quad \text{and} \quad \boxed{y^{<t>} = g_2(W_{ya}a^{<t>} + b_y)}$$

where $W_{ax}, W_{aa}, W_{ya}, b_a, b_y$ are coefficients that are shared temporally and $g_1, g_2$ activation functions.

Credit - Stanford

# ADVANCEMENTS OF DEEP LEARNING

Major leaps began to happen in the 2010's as the capability of these networks began to reach human level.

Alexnet (2012) - Won the Large Scale Visual Recognition Challenge(LSVRC) with an error rate 10% lower than the previous year. Used dropout to reduce overfitting and a rectified linear activation unit (ReLU)

Generative Adversarial Networks (2014) - Ian Goodfellow -

Gated Recurrent Units (2014) - this is when Siri, Alexa, Google Voice began to actually understand my accent.

"Langley Theory of AI" - they tried something and it somehow worked. Then they had to write a paper to figure out why.

# THE TRANSFORMER (2017 - PRESENT)

Introduced by: Vaswani et al. in the paper "Attention is All You Need"

- **Self-Attention Mechanism**: Models relationships between all tokens in a sequence simultaneously, eliminating the need for sequential processing.
- **Positional Encoding**: Adds positional information to input embeddings since Transformers do not process data sequentially.
- **Feedforward Networks**: Applied after the self-attention mechanism within each layer.
- **Multi-Head Attention**: Allows the model to focus on different parts of the input simultaneously.
- **Parallelism**: Processes entire sequences at once, significantly speeding up training compared to RNNs.

# THE TRANSFORMER (2017 - PRESENT)

Advantages:

- Handles long-range dependencies effectively.
- Scales well with increased data and compute.
- Allows for parallel processing, leading to faster training.

BERT (2018)

- Bidirectional Encoder Representations from Transformers.
- Pretrained on massive text corpora using a masked language modeling objective.
- Fine-tuned for downstream tasks like sentiment analysis, question answering, and named entity recognition.

GPT (2018–2023+)

- Generative Pre-trained Transformer.
- Introduced by OpenAI, focused on autoregressive language modeling (predicting the next token).
- GPT models (GPT-2, GPT-3, GPT-4) scaled up significantly, achieving remarkable results in text generation.

HSV.AI

# FINE TUNING

Fine-tuning a large language model (LLM) involves adapting a pretrained model (e.g., GPT, BERT) to perform specific tasks or align its behavior with particular requirements. This process leverages the general knowledge learned during pretraining and refines it using task-specific data.

Fine-tuning can also be applied to other deep learning networks, such as CNNs to be able to classify specific images.

# REINFORCEMENT LEARNING WITH HUMAN FEEDBACK (RLHF)

Description:

- Combines reinforcement learning with human-labeled data to guide the model's behavior.

Pretraining:

- The model is pretrained on a large corpus of data using standard objectives (e.g., predicting the next word).

Fine-tuning:

- The model is fine-tuned using human feedback to optimize for specific goals.

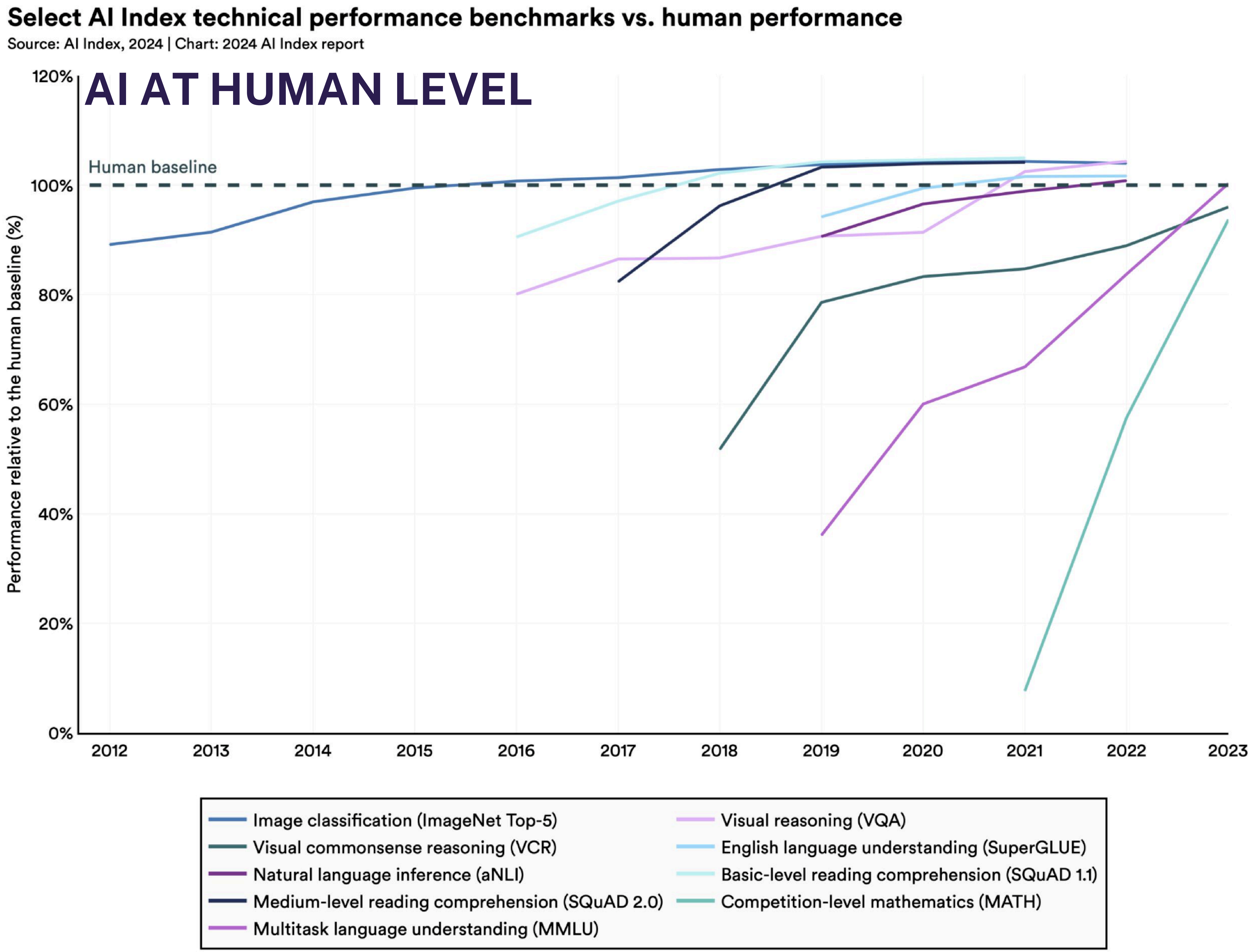
Reward Modeling:

- Human preferences are collected by showing multiple outputs to humans and asking them to rank the responses.
- A reward model is trained on this data to predict the quality of a response.

Policy Optimization:

- The model is fine-tuned using reinforcement learning, optimizing for the reward model's scores.

# AI AT HUMAN LEVEL



# WHAT DOES HUMAN LEVEL MEAN ANYWAY?

Aoccdrnig to a rscheearch at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe.

Factually incorrect, but fun meme

# AI Challenges for Planning

**How are advances and adoption of AI by the community affecting urban planning?**

**How can the challenges be addressed?**

# AUTONOMOUS VEHICLES

- Ghost Riders
  - Camera-covered cars with nobody inside are causing traffic jams, angering residents and amazing tourists
- Parking
  - They don't actually have to park - it's often cheaper to cruise city streets at low speeds. Or park in front of private houses.
  - Waymo cars honk at each other throughout the night, disturbing SF neighbors
- Driving Behavior
  - Waymo passenger nearly misses his flight after car drives in circles
  - 'Complete meltdown': Driverless cars in San Francisco stall causing a traffic jam
  - Armed with traffic cones, protesters are immobilizing driverless cars

HSV.AI

# ROUTING & RECOMENDATION

Nagivation

- Your Navigation App Is Making Traffic Unmanageable
  - Individual rather than Community focused
- Automated navigation systems are still wreaking havoc on small towns' streets
  - Apps like Google and Waze are redirecting traffic to secondary roads that are not equipped to handle the traffic, disrupting their infrastructure.

Communication

- Friction with social media recommendation algorithms

Real Estate Investment

- Tons of AI based investment advisers - may not match city planning forecasts
- Zillow, Redfin, Mashvisor, Reonomy (Commercial Real Estate)
- I lost a lot more time looking at property than I should have

# AI Advancements in Planning

**What types of things can AI do?**

**How are advances and adoption of AI by urban planners affecting the community?**

# AI TECHNIQUES

1. Natural Language Processing - Classification, Summarization, Generation
2. Computer Vision - Classification, Segmentation, Enhancement, Generation
3. Audio Processing - Recognition, Classification, Generation
4. Multimodal Tasks - Text-to-Image, Text-to-Sound, Text-to-Video
5. Reinforcement Learning - Autonomous Vehicles, Robotics, Optimization
6. Recommendation Systems - Collaborative Filtering, Content Based, Implicit
7. Generative AI - Text, Images, Audio, Video
8. Agentic AI - Tool Calling, Decision Making
9. Robotics - Humanoid Robots, Autonomous Drones, Self Driving Cars

# NATURAL LANGUAGE PROCESSING

- Classification
- Named Entity Recognition
- Sentiment Analysis
- Summarization
- Semantic Understanding
- Language Translation

# NATURAL LANGUAGE PROCESSING

## Limitations

- Language Bias
- Gender Bias
- Racial / Ethnic Bias
- Detecting sarcasm or dark humor
- Words change meaning over time
- Implied context
- Model context size
- Emerging language (generational problems)

## COMPUTER VISION

- Classification
- Segmentation
- Enhancement
- Object Recognition
- Facial Recognition
- Pose Recognition

# COMPUTER VISION

## Limitations

- Gender Bias
- Racial / Ethnic Bias
- Images without context
- Difficulty translating to real world implementation
- Adversarial vulnerabilities by changing small parts of images
- Lack of physical world understanding

# MULTIMODAL TASKS

Image-to-Text

- Generating descriptive textual captions for a given image.
- Answering questions based on the content of an image.
- Retrieving the most relevant image given a textual query or vice versa.

Text-to-Image

- Generating images based on textual descriptions.

Audio-to-Text

- Generating textual descriptions of audio events.
- Speech Recognition

Text-to-Audio

- Generating speech from textual input.
- Generating music from textual input.

Audio/Video-to-Text

- Improved speech recognition by combining video and audio

# MULTIMODAL TASKS

## Limitations

- Data Alignment (time alignment), scarcity for some domains, data diversity.
- Domain transfer problems (movie subtitles vs instructional videos)
- Interpretability or Explainability with troubleshooting problems
- Difficult to evaluate correctness - lack of benchmarks and metrics
- Imbalanced datasets (more text than audio than video)

# GENERATIVE AI

- Generating Text - lots and lots of text.
  - Can be formatted in specific ways
  - Can match a given tone
- Images
  - Can create images from text descriptions.
  - Can alter images based on another provided image
- Audio
  - Can create audio from text
  - Can create music from text
- Video
  - Can create video from text
  - Short clips are fairly easy at the moment, with longer running videos coming soon

# GENERATIVE AI

## Technical Limitations

- Plagiarism and Copyright Issues - Generated content may inadvertently copy from the training data, raising intellectual property concerns.
- Misinformation and Deepfakes - Generative AI can create highly convincing fake text, images, videos, and audio, making it easier to spread misinformation.
- Coherence in Long Outputs - Struggles to maintain coherence in long or complex narratives.

## Cultural/Ethical Limitations

- Alignment to cultural norms
- Misuse by bad actors

## URBAN AI OCTOBER 2024

2nd ACM SIGSPATIAL International Workshop on Advances in Urban AI

Proceedings are available here: https://urbanai.ornl.gov/urbanai2024/wp-content/uploads/sites/5/2024/11/Urban-AI-2024-proceedings.pdf

Topics:

- A Graph Deep Learning Model for Station Ridership Prediction in Expanding Metro Networks
- Smart Route: A GIS-Based Solution for Mass Transit Design and Optimization
- Generative-AI based Map Representation and Localization
- Encryption Techniques for Privacy-Preserving CNN Models Performance and Practicality in Urban AI Applications
- SurfaceAI: Automated creation of cohesive road surface quality datasets based on open street-level

HSV.AI

# AAAI 2025'S WORKSHOP ON AI FOR URBAN PLANNING

The 1st Workshop on AI for Urban Planning aims to bring together researchers, practitioners, and policymakers to explore innovative AI-driven solutions for the multifaceted challenges in urban planning.

Website - https://ai-for-urban-planning.github.io/AAAI25-workshop

Selected Presentations:

- Integrating Human Dynamics into AI-Driven GeoDesign for Sustainable Urban Futures
- Mining E-scooter Safety Policies and Plans with GPT-4o and Latent Dirichlet Allocation (LDA) Topic Modeling

# AAAI 2025'S WORKSHOP ON AI FOR URBAN PLANNING

Integrating Human Dynamics into AI-Driven Urban Science for Symbiotic Futures

Applications of AI:

- Image detection of doors viewed from street. Triangulation to estimate elevation of the door. Data used to evaluate risk of flood damage.
- Combining multiple data sources like building footprints, LiDAR, property records, and street view images helps overcome individual dataset constraints
- Smart Phone APP to provide an effective way to collect the public's' design demands and ideas - using AR to let the public view a proposed structure in place
- Demonstrate a workflow that efficiently generates detailed and visually coherent landscape designs, including natural parks, city plazas, and courtyard gardens

# AVAILABLE PRODUCTS

Stanford University InVEST® is a suite of free, open-source software models used to map and value the goods and services from nature that sustain and fulfill human life.
https://naturalcapitalproject.stanford.edu/software/invest

Autodesk Offerings
https://www.autodesk.com/solutions/urban-design-planning
https://www.autodesk.com/eu/campaigns/spacemaker

Simulation of Urban MObility
https://eclipse.dev/sumo/

# Questions?

## COMMENTS, OR CRIES OF HERESY?

J. Langley
Founder, Huntsville AI
CTO, CohesionForce, Inc

Josh Phillips
Data Group Working Lead, Open Model Initiative
CDO, CohesionForce, Inc