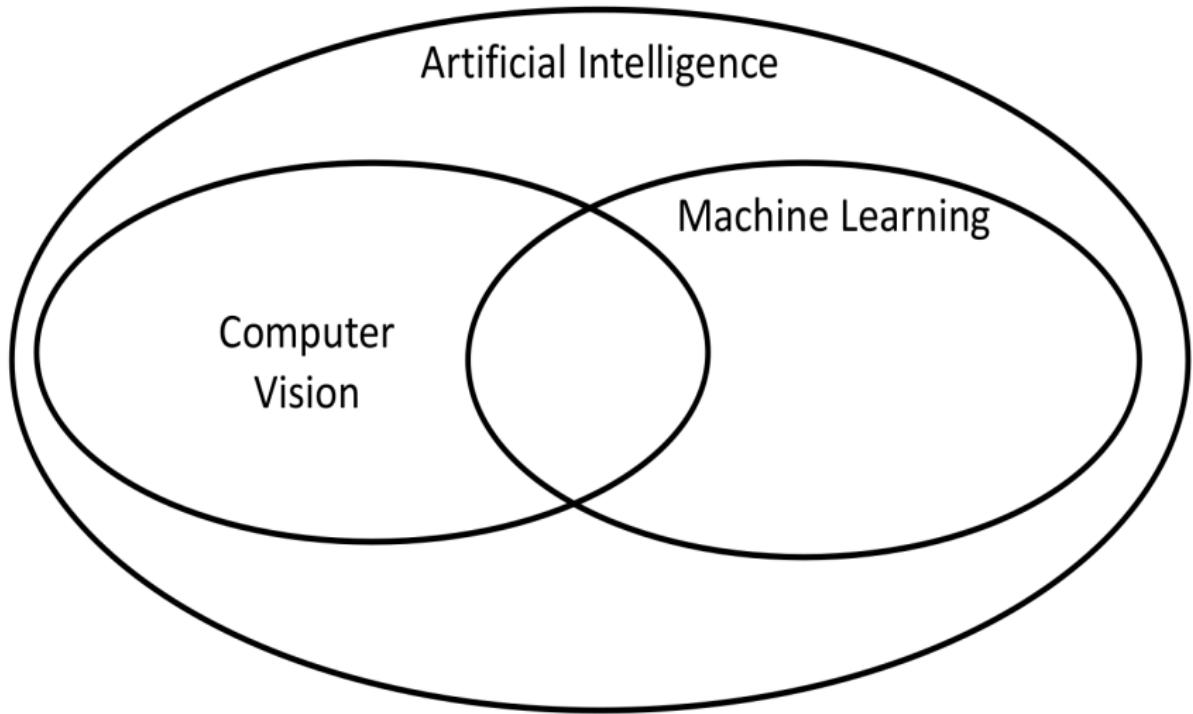


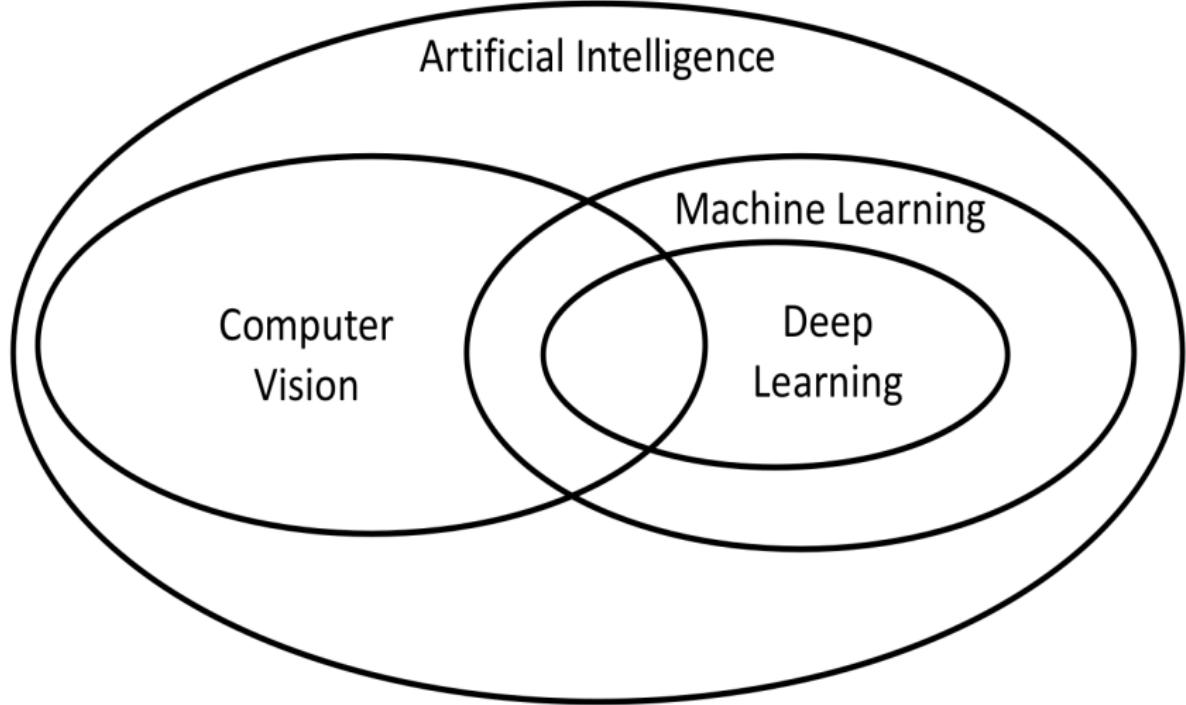
# Deep Learning

Ranjeet Ranjan Jha

Mathematics Department  
Indian Institute of Technology Patna

# Artificial Intelligence





# Deep Learning Basics

- Image Classification: A core task in Computer Vision



→ cat

This image by Nikita is  
licensed under CC BY 2.0

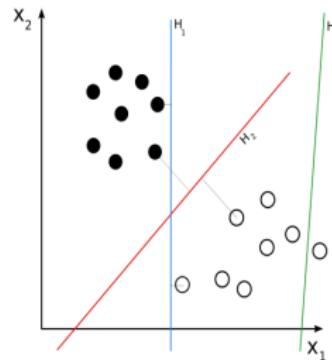
# Deep Learning Basics

- Image Classification: A core task in Computer Vision



This image by Nikita is  
licensed under CC-BY 2.0

→ cat



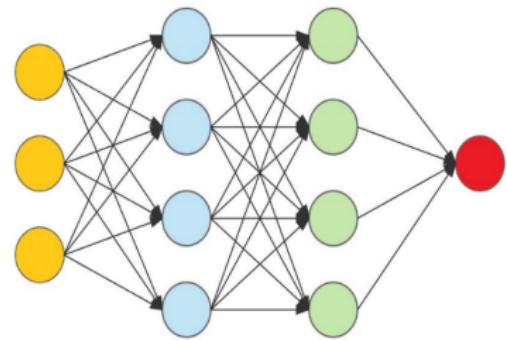
Linear Classifier

# Deep Learning Basics

- Image Classification: A core task in Computer Vision



→ cat



Neural Networks

This image by [blikita](#) is  
licensed under CC-BY 2.0

# Tasks Beyond Image Classification

Classification



CAT

No spatial extent

Semantic Segmentation



GRASS, CAT, TREE,  
SKY

No objects, just pixels

Object Detection



DOG, DOG, CAT

Multiple Object

This image is CC0 public domain

Instance Segmentation



DOG, DOG, CAT

# Tasks Beyond Image Classification

Video  
Classification

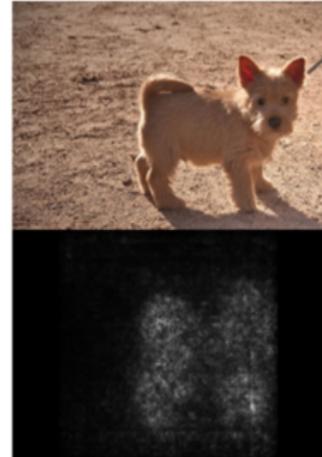


Running? Jumping?

Multimodal Video  
Understanding



Visualization &  
Understanding



# Models Beyond Multi-Layer Perceptron

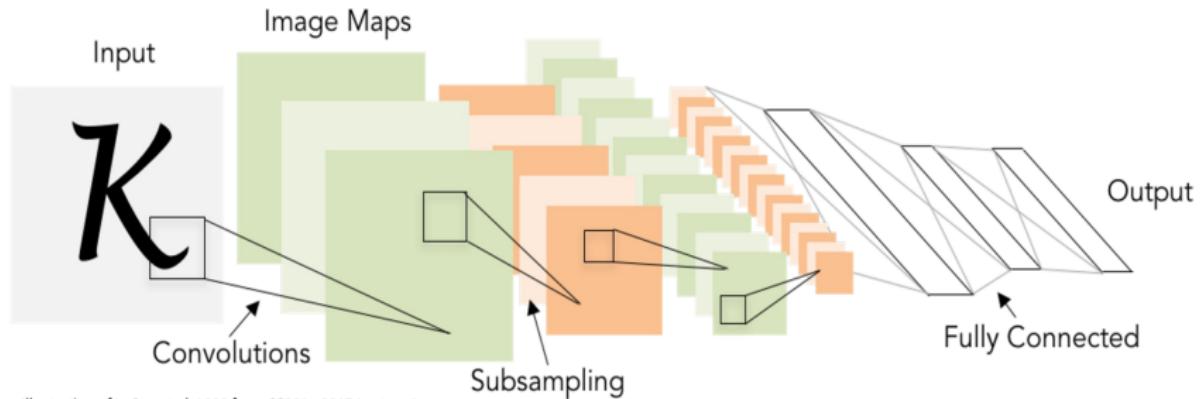
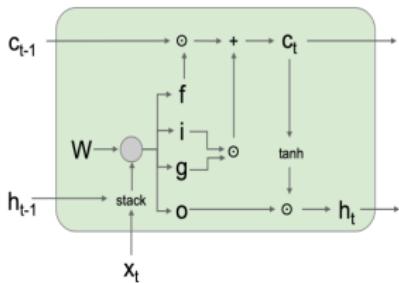
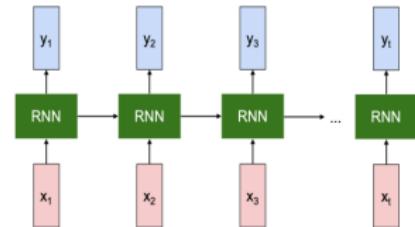


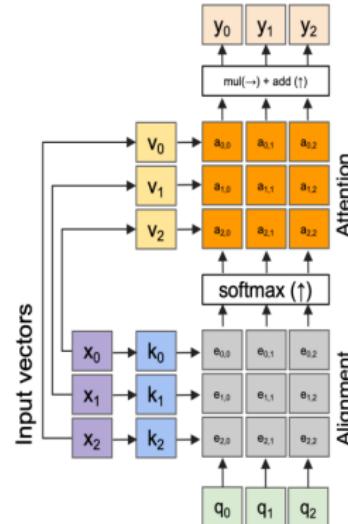
Illustration of LeCun et al. 1998 from CS231n 2017 Lecture 1

Convolutional neural network

# Models Beyond Multi-Layer Perceptron



Recurrent neural network



Attention mechanism / Transformers

## Beyond 2D Recognition

# Beyond 2D Recognition: Generative Modeling



Style Transfer

# Beyond 2D Recognition: Generative Modeling



“Teddy bears working on new  
AI research underwater with  
1990s technology”

DALL-E 2

# DALL-E 2



"Teddy bears working on new AI research on the moon in the 1980s."



"Rabbits attending a college seminar on human anatomy."



"A wise cat meditating in the Himalayas searching for enlightenment."

Image source: Sam Altman, <https://openai.com/dall-e-2/>, <https://twitter.com/sama/status/1511724264629678084>

# DALL-E 3

An illustration from a graphic novel. A bustling city street under the shine of a full moon. The sidewalks bustling with pedestrians enjoying the nightlife. At the corner stall, a young woman with fiery red hair, dressed in a signature velvet cloak, is haggling with the grumpy old vendor. The grumpy vendor, a tall, sophisticated man wearing a sharp suit, who sports a noteworthy mustache is animatedly conversing on his steampunk telephone.



Betker, James, et al. "Improving image generation with better captions." *Computer Science*. <https://cdn.openai.com/papers/dall-e-3.pdf> (2023).

# GPT-4

User What is unusual about this image?



Source: [Bumorama](#)

GPT-4 The unusual thing about this image is that a man is ironing clothes on an ironing board attached to the roof of a moving taxi.

User Can you explain this meme?

Sometimes I just look at pictures of the earth from space and I marvel at how beautiful it all is.



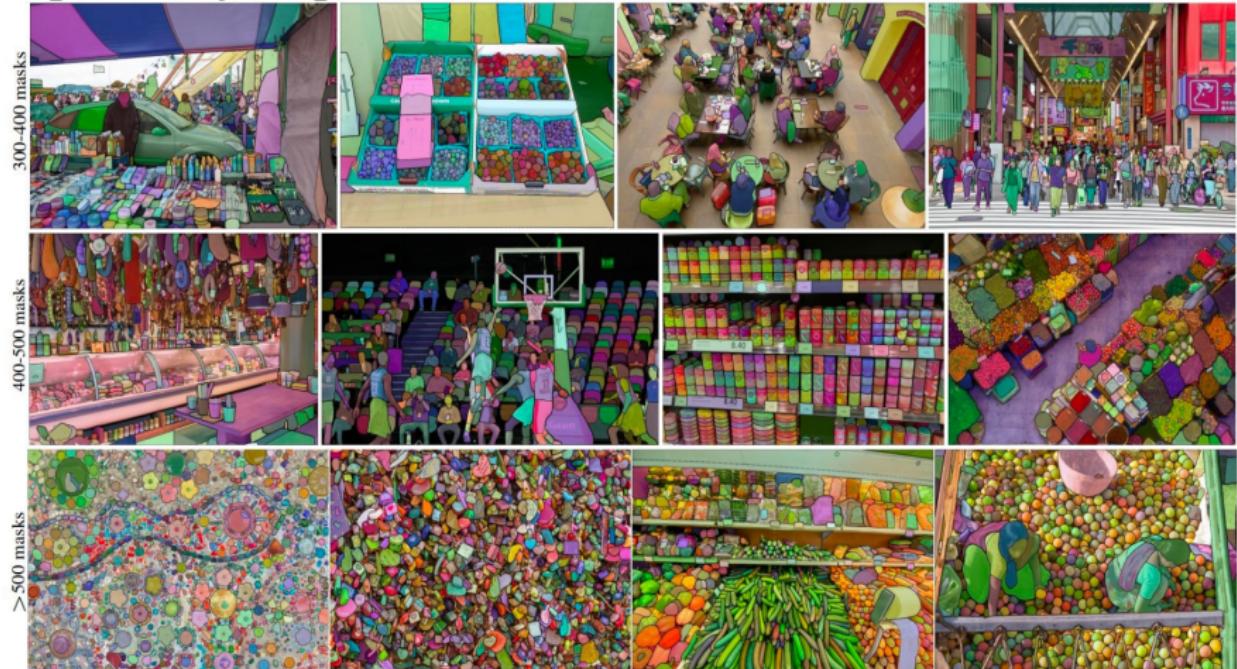
GPT-4 This meme is a joke that combines two unrelated things: pictures of the earth from space and chicken nuggets.

The text of the meme suggests that the image below is a beautiful picture of the earth from space. However, the image is actually of chicken nuggets arranged to vaguely resemble a map of the world.

The humor in this meme comes from the unexpected juxtaposition of the text and the image. The text sets up an expectation of a majestic image of the earth, but the image is actually something mundane and silly.

Image source: <https://openai.com/research/gpt-4>

# Segment Anything Model (SAM)





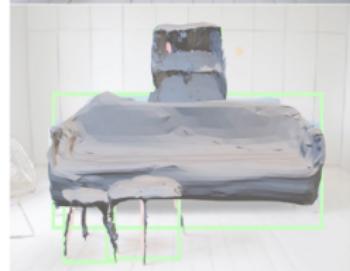
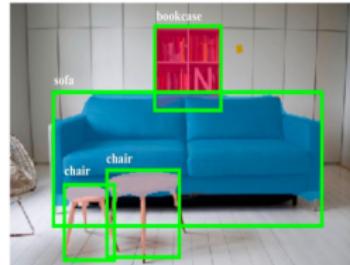
# Beyond 2D Recognition: 3D Vision



Choy et al., 3D-R2N2: Recurrent Reconstruction Neural Network (2016)



Zhou et al., 3D Shape Generation and Completion through Point-Voxel Diffusion (2021)



Gkioxari et al., "Mesh R-CNN", ICCV 2019

# 2018 Turing Award for deep learning

most prestigious technical award, is given for major contributions of lasting importance to computing.



[This image is CC0 public domain](#)

[This image is CC0 public domain](#)

[This image is CC0 public domain](#)

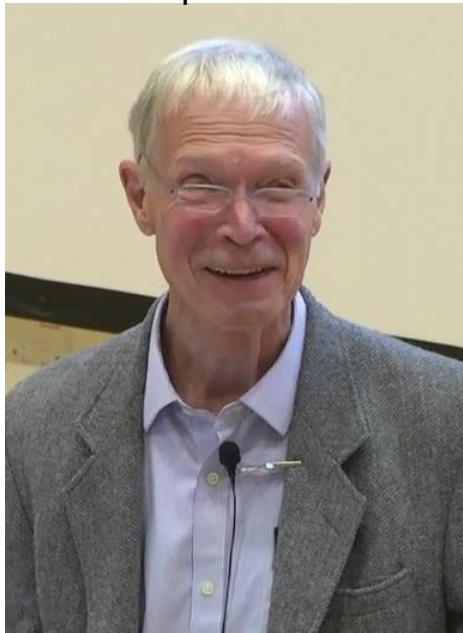
# 2024 Nobel Prize in Physics

Hinton speaking at the Nobel Prize  
Lectures in Stockholm in 2024



[This image is CC0 public domain](#)

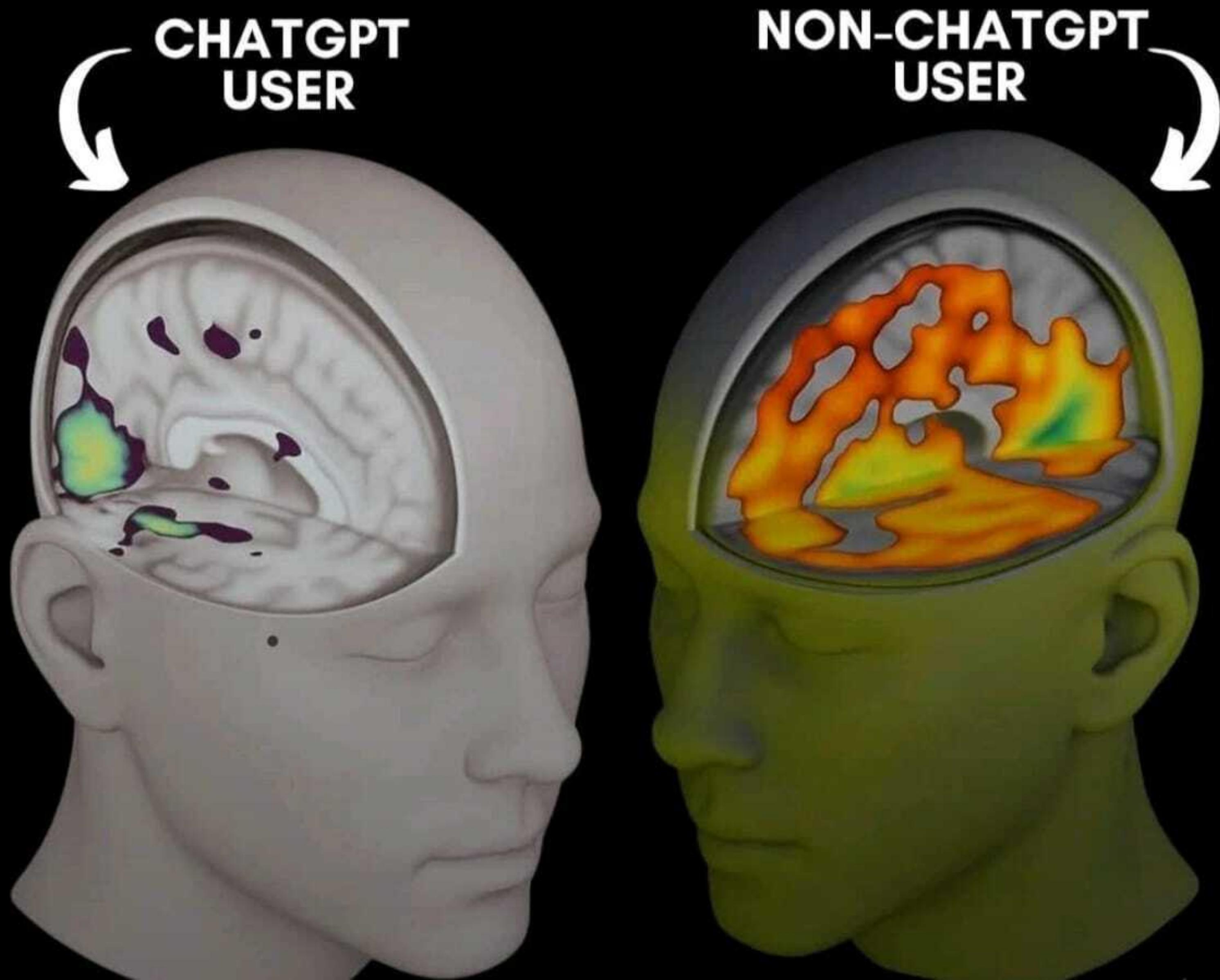
John Hopfield in 2016



[This image is CC0 public domain](#)

In 2024, he was jointly awarded the [Nobel Prize in Physics](#) with [John Hopfield](#) “for foundational discoveries and inventions that enable machine learning with artificial neural networks.”

Scientists just completed the first **brain scan study** of **ChatGPT users**. The results are terrifying.



<b>Course Number</b>	MA4104 (DE)
<b>Course Credit (L-T-P-C)</b>	2 – 0 – 2 – 3
<b>Course Title</b>	Deep Learning
<b>Learning Mode</b>	Lectures and Labs
<b>Learning Objectives</b>	Gain expertise in artificial neural networks, covering fundamentals, feedforward and deep neural networks, convolutional networks, recurrent neural networks, and popular deep learning architectures, to proficiently tackle pattern recognition tasks and real-world challenges.
<b>Course Description</b>	Explore the foundations and applications of deep learning, covering artificial neural networks, convolutional networks, recurrent neural networks, and popular architectures like VAE, and GANs for solving real-world tasks.
<b>Course Outline</b>	Basics of artificial neural networks (ANN); Feedforward neural networks: Pattern classification using perceptron, Multilayer feedforward neural networks, Backpropagation learning, Normalization; Deep neural networks (DNNs): Difficulty of training DNNs, Optimization for training DNNs, Optimization methods for neural networks (AdaGrad, RMSProp, Adam etc.), Regularization methods. Convolutional Networks (CNNs): Introduction to CNNs – convolution, pooling, Deep CNNs, Deep CNN architectures (AlexNet, VGG, GoogLeNet, ResNet), Other Recent CNN architectures. Recurrent neural networks (RNNs), Long Short Term Memory (LSTM), Other Recent Sequential Networks; Some popular Architectures/concepts in Deep Learning: Object Detection and Localization, Siamese Networks, Autoencoders & VAE, Generative Adversarial Networks (GANs), Other Recent Topics.
<b>Learning Outcome</b>	Students will acquire a thorough grasp of both foundational and advanced concepts in deep learning, along with practical proficiency in utilizing these methods to address real-world challenges.
<b>Assessment Method</b>	Quiz /Assignment/ Project / MSE / ESE

**Text Books:**

1. Deep learning, Ian Goodfellow, Yoshua Bengio and Aaron Courville, MIT Press, 2016.

**Reference Books:**

1. S. Haykin, Neural Networks and Learning Machines , Prentice Hall of India, 2010
2. Satish Kumar, Neural Networks - A Class Room Approach, Second Edition, Tata McGraw-Hill, 2013
3. C.M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006

Thank You!

Any questions?