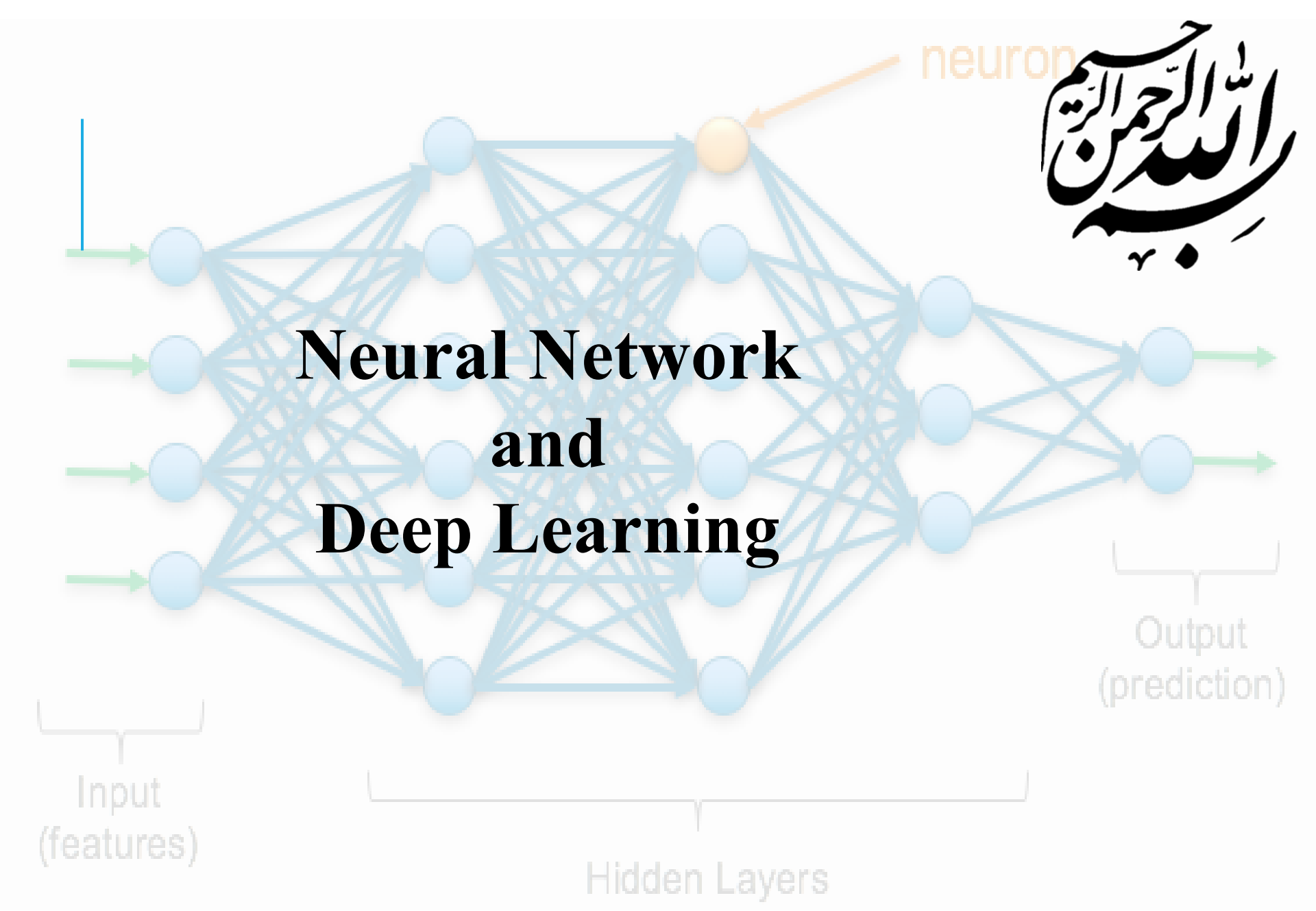


بسم الله الرحمن الرحیم

Neural Network and Deep Learning

Lecture1:

Introduction

OUTLINE

- Course outline
- Evaluation
- Introduction
- History
- Applications

COURSE OUTLINE

- Basics of Neural Network programming
  - Biological VS. Artificial Neural Networks
  - Neural model
  - Applications of artificial neural network
  - Activation Functions
  - Perceptron
  - Perceptron (learning rule)
  - Multilayer Perceptrons
  - ADALINE Network
  - Recurrent Neural Network (RNN)
  - LMS algorithm
  - Backpropagation Algorithm

COURSE OUTLINE

- Artificial Neural Network life cycle
  - Data
  - Features
  - Train the model
    - Loss functions
    - Batch size
    - Optimizers
    - Underfitting, overfitting, and generalization
    - Regularization in Neural Networks
  - Test the model
    - Evaluation metrics

COURSE OUTLINE

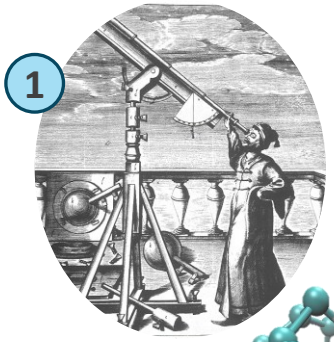
- CNN Architectures
- Recurrent Neural Networks
- Attention and Transformers
- Self-supervised Learning
- Generative Models
- …

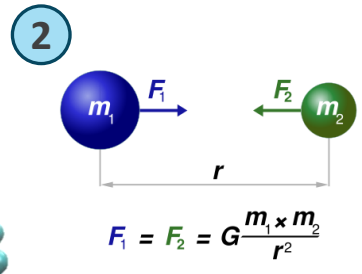
EVALUATION

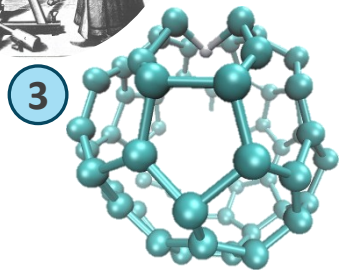
- Assignments: 4 Point
- Quizzes: 1 Point
- Midterm: 6 Point
- Final exam: 9 Point
- Final project: 1 Point

INTRODUCTION

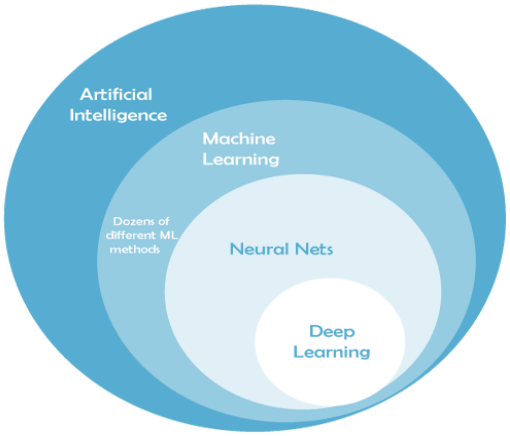
- **Science Paradigms**
  1. Thousand years ago:
     science was **empirical**
     describing natural phenomena
  2. Last few hundred years:
     theoretical branch
     using models, generalizations
  3. Last few decades:
     a **computational** branch
     simulating complex phenomena
  4. Today: **data exploration** (eScience)
     unify theory, experiment, and simulation
     – Data captured by instruments or generated by simulator
     – Processed by software
     – Information/knowledge stored in computer
     – Scientist analyzes database / files using data management and statistics





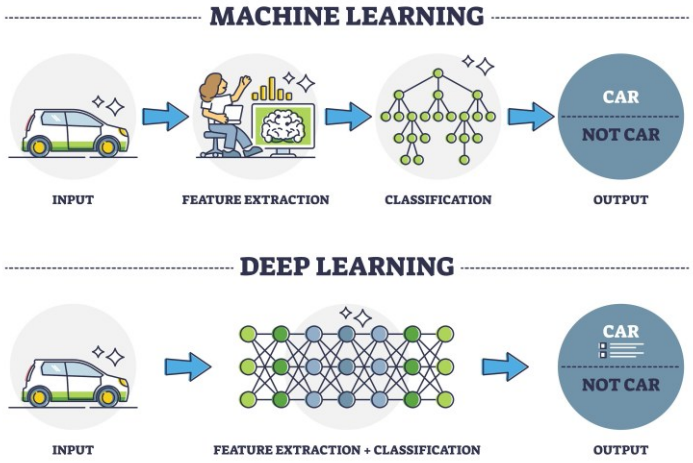


INTRODUCTION



- **Artificial intelligence** refers to the general ability of computers to emulate human thought and perform tasks in real-world environments.

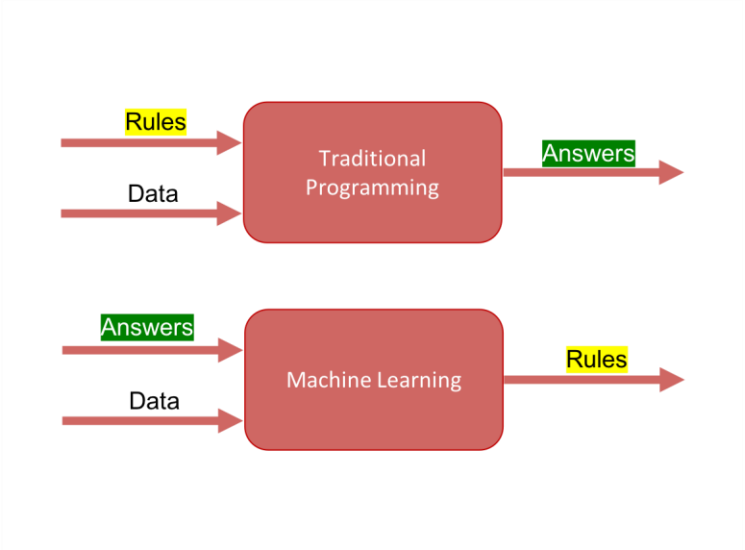
- **Machine learning** refers to the technologies and algorithms that enable systems to identify patterns, make decisions, and improve themselves through experience and data.
- **Deep learning** is a subset of machine learning. The primary difference between machine learning and deep learning is how each algorithm learns and how much data each type of algorithm uses.
  - Deep learning automates much of the feature extraction piece of the process, eliminating some of the manual human intervention required.
  - It also enables the use of large data sets. That capability is exciting as we explore the use of unstructured data further, particularly since over 80% of an organization's data is estimated to be unstructured



- **Artificial neural networks** or simulated neural networks, are a subset of machine learning and are the backbone of deep learning algorithms. They are called "neural" because they mimic how neurons in the brain signal one another.

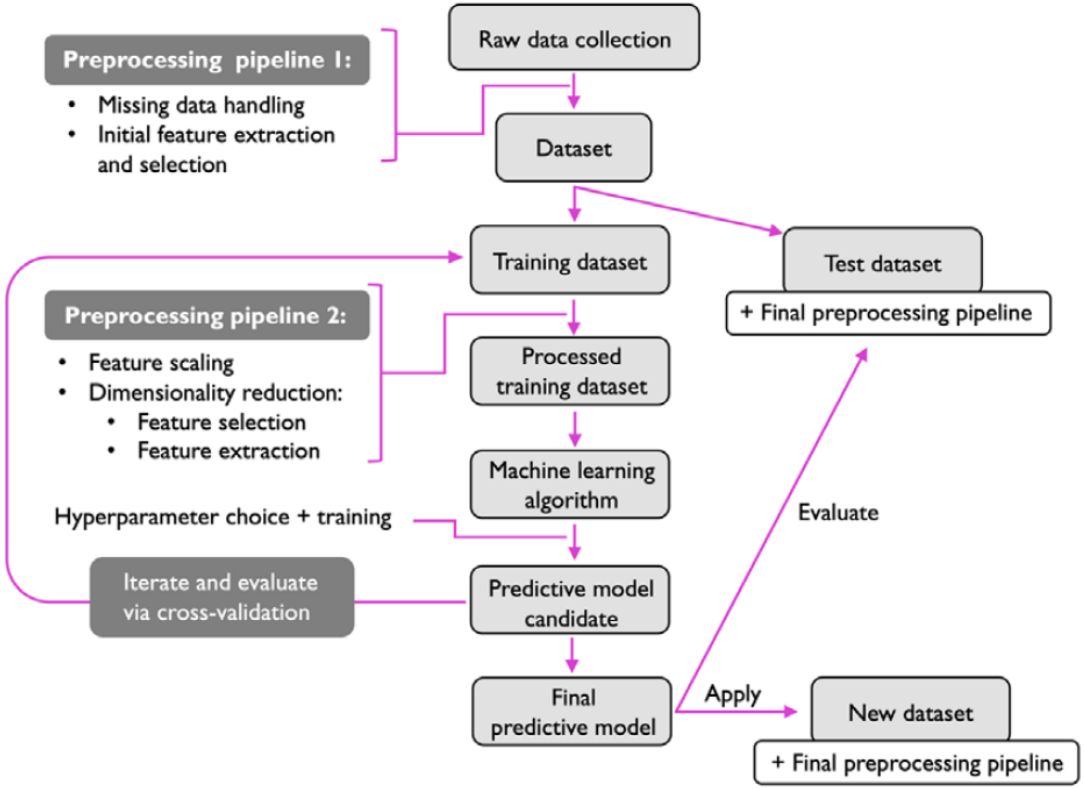
INTRODUCTION

- **Building intelligent machines to transform data into knowledge**
  - Instead of requiring humans to manually derive rules and build models from analyzing large amounts of data, machine learning offers a more efficient alternative for capturing the knowledge in data to gradually improve the performance of predictive models and make data-driven decisions.



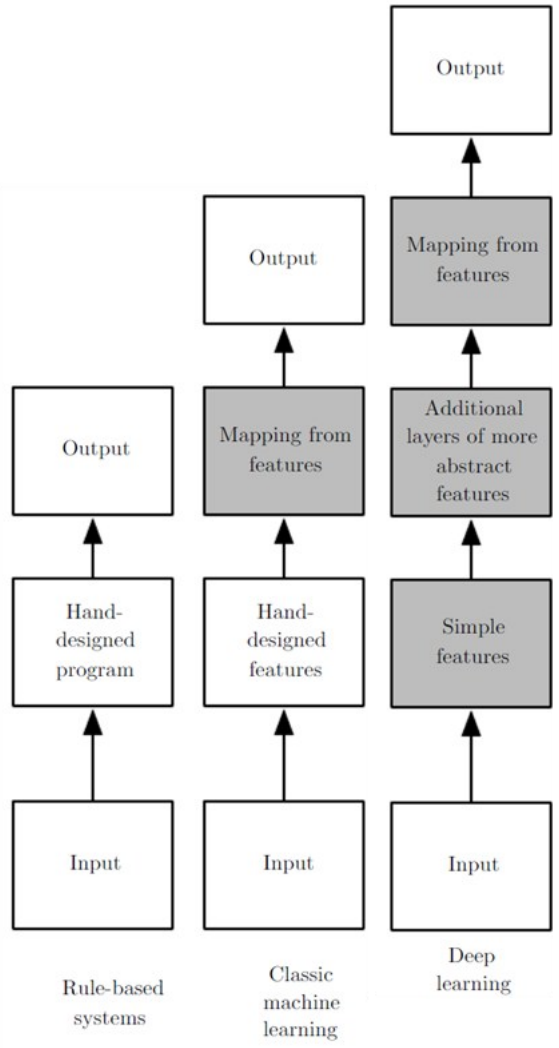
INTRODUCTION

- **A roadmap for building machine learning systems**



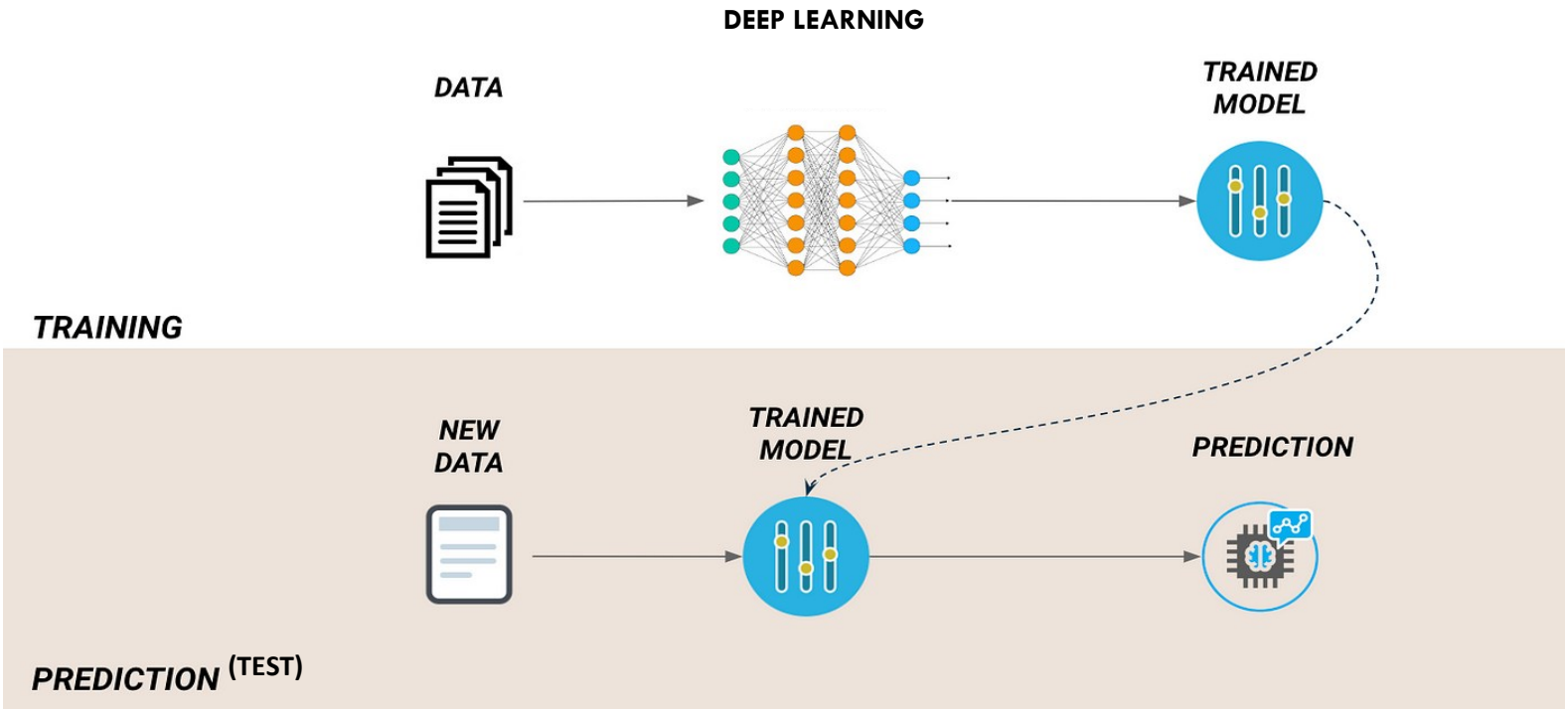
INTRODUCTION

- **Rule-based, classic machine learning and deep learning**



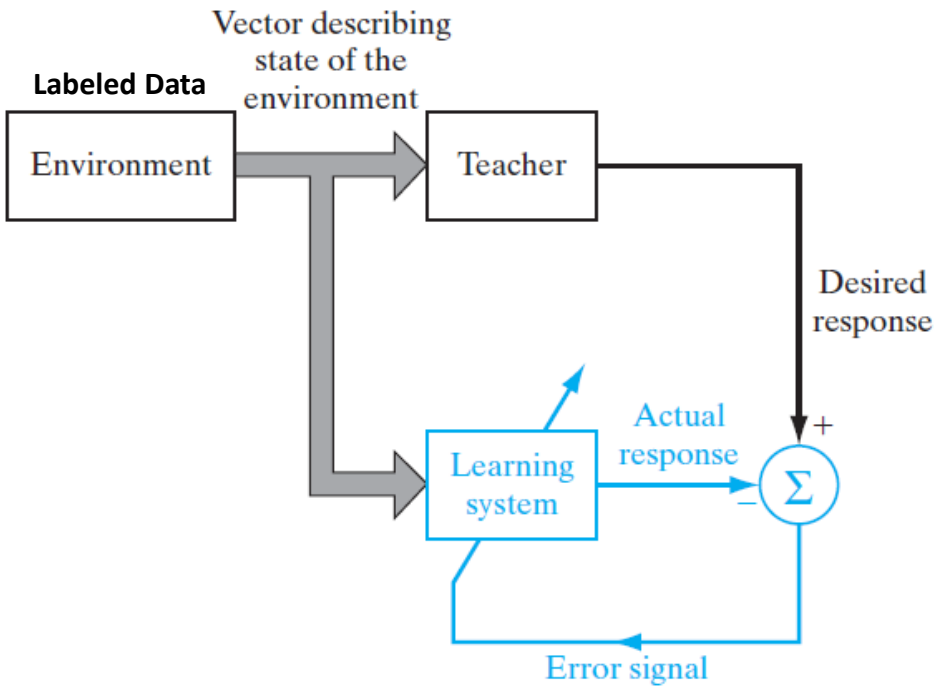
INTRODUCTION

- **Deep learning process steps**



INTRODUCTION

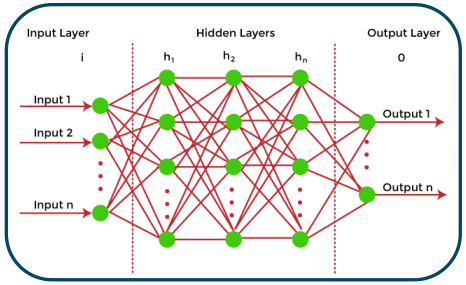
- **Training a deep mode**
  - Labeled data
  - Model (involve of weights)
  - Loss function
  - Optimizer



INTRODUCTION

- **Training a deep mode**
  - Labeled data
  - Model (involve of weights)
  - Loss function
  - Optimizer

Database:
input and label



Model: $f(\mathbf{x}, w_1, w_2, \ldots)$

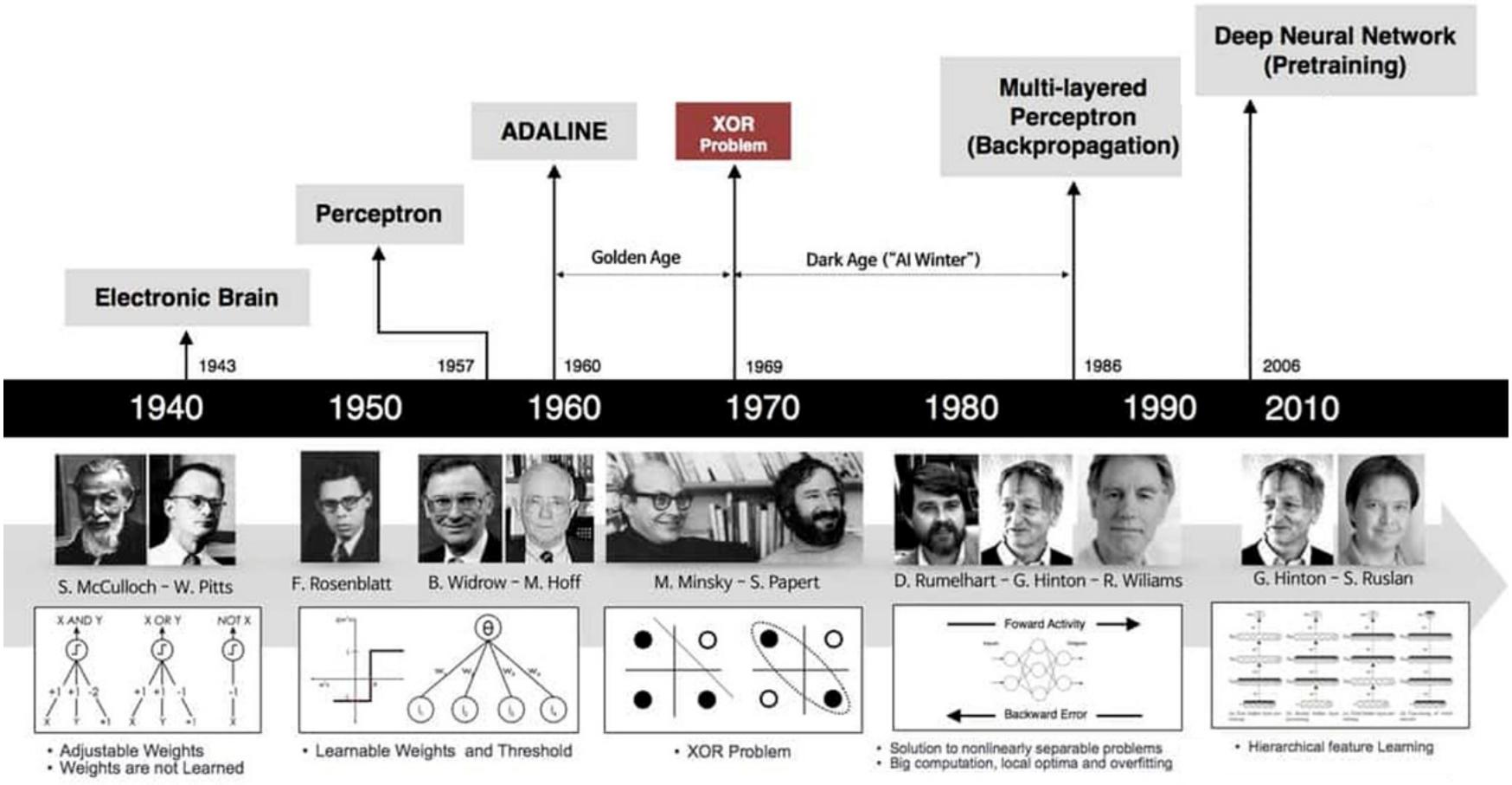
Input data **Weights**

Loss function

$$MSE = \frac{1}{N}\sum_{i}^{N}(Yi - \hat{Y}i)^2$$

Optimizer for updating of weights to reduction of loss function

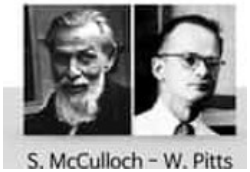
HISTORY


Deep Neural Network (Pretraining)
Multi-layered Perceptron (Backpropagation)
ADALINE
XOR Problem
Perceptron
Golden Age
Dark Age ("AI Winter")
Electronic Brain
1943
1957
1960
1969
1986
2006
1940
1950
1960
1970
1980
1990
2010
S. McCulloch - W. Pitts
F. Rosenblatt
B. Widrow - M. Hoff
M. Minsky - S. Papert
D. Rumelhart - G. Hinton - R. Wiliams
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
• Hierarchical feature Learning

HISTORY

1. In 1943 Warren McCulloch, a neurophysiologist, and Walter Pitts, a logician, teamed up to develop a mathematical **model of an artificial neuron**. In their paper "A Logical Calculus of the Ideas Immanent in Nervous Activity," they realized that a simplified model of a real neuron could be represented using <u>simple addition and thresholding</u>.

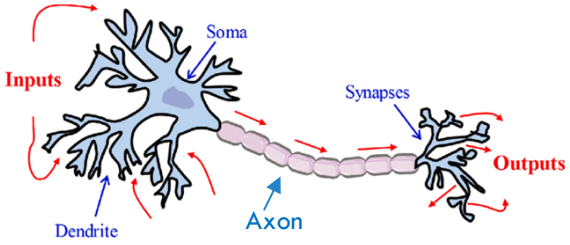
BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 5, 1943



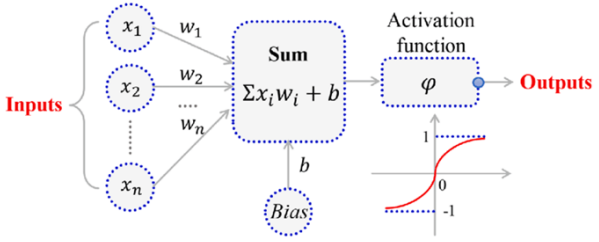
A LOGICAL CALCULUS OF THE
IDEAS IMMANENT IN NERVOUS ACTIVITY

WARREN S. MCCULLOCH AND WALTER PITTS

FROM THE UNIVERSITY OF ILLINOIS, COLLEGE OF MEDICINE,
DEPARTMENT OF PSYCHIATRY AT THE ILLINOIS NEUROPSYCHIATRIC INSTITUTE,
AND THE UNIVERSITY OF CHICAGO



(a) Biological neuron



(b) Artificial neuron

HISTORY

2. In the late 1950s (**1958**), the world witnessed a significant advancement in artificial intelligence with the invention of the Perceptron by Frank Rosenblatt. This innovation, based on the concept of artificial neurons, laid the groundwork for the development of complex neural networks that drive modern AI.

   Inspired by the way neurons work together in the brain, the perceptron is a single-layer neural network – an algorithm that classifies input into two possible categories. The neural network makes a prediction – say, right or left; or dog or cat – and if it's wrong, tweaks itself to make a more informed prediction next time. It becomes more accurate over thousands or millions of iterations.

HISTORY

3. **ADALINE** (Adaptive Linear Neuron or later Adaptive Linear Element) is an early single-layer artificial neural network. It was developed by professor Bernard Widrow and his doctoral student Hoff at Stanford University in **1960**. It is based on the perceptron. It consists of a weight, a bias and a summation function.

ADALINE: ADAptive LINear Element

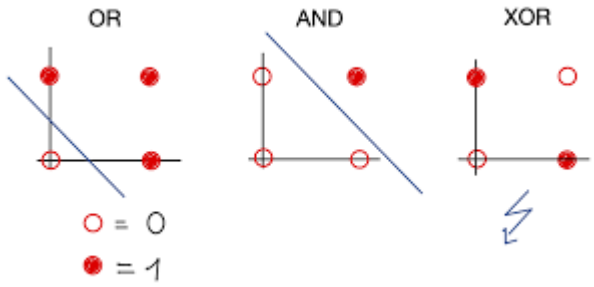
$$O(x) = \sum_{i=1}^{n} x_i . w_i + w_0$$

HISTORY

4. An MIT professor named Marvin Minsky (who was a grade behind Rosenblatt at the same high school!), along with Seymour Papert, wrote a book called Perceptrons (MIT Press) about Rosenblatt's invention. They showed that a single layer of these devices was unable to learn some simple but critical mathematical functions (such as **XOR**).

   In the same book, they also showed that using multiple layers of the devices would allow these limitations to be addressed.


M. Minsky – S. Papert

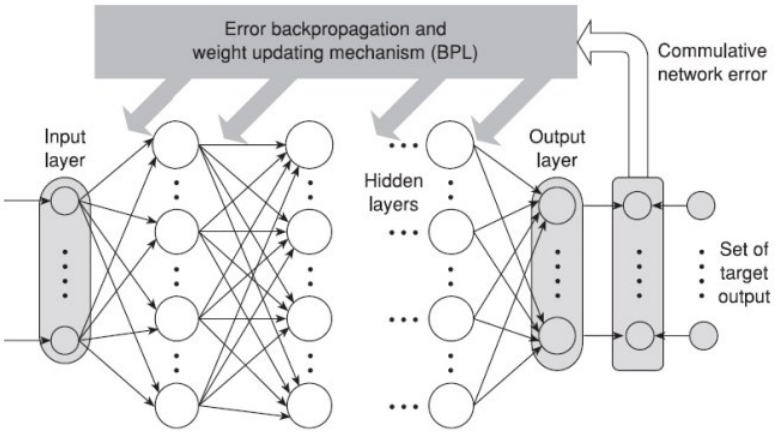


HISTORY

5. Rumelhart was the first author of a highly cited paper from **1985** (co-authored by Geoffrey Hinton and Ronald J. Williams) that applied the **back-propagation** algorithm to **multi-layer neural networks**. This work showed through experiments that such networks can learn useful internal representations of data.

(1986)
David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams

Learning representations by back-propagating errors
Nature 323:533–536



HISTORY

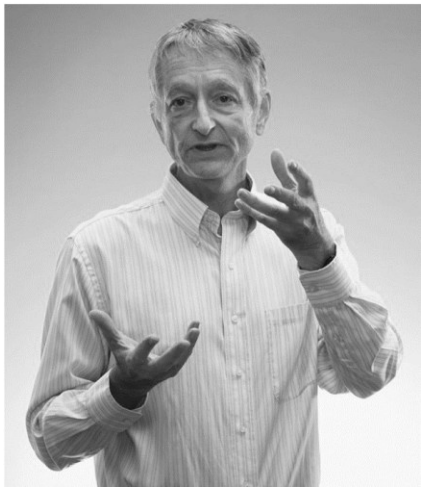
6. Perhaps the most pivotal work in neural networks in the last 50 years was the multivolume Parallel Distributed Processing (PDP) by David Rumelhart, James McClelland, and the PDP Research Group, released in **1986** by MIT Press.

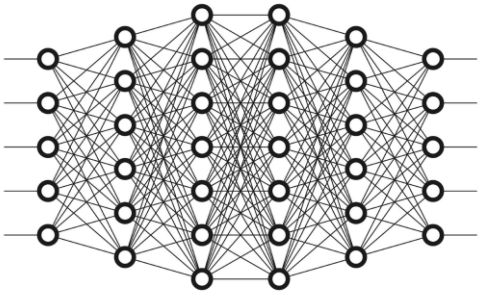
The Appeal of Parallel Distributed Processing

J. L. McCLELLAND, D. E. RUMELHART, and G. E. HINTON

HISTORY

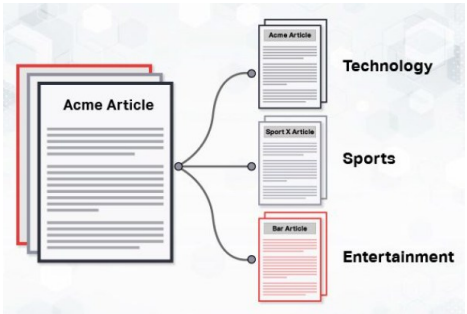
7. Computer scientist **Geoffrey Hinton** — now a Google researcher and long known as the "Godfather of Deep Learning" Ruslan's doctoral advisor was Geoffrey Hinton. Ruslan was considering quitting the field of artificial intelligence when he met Hinton in 2004, but changed his mind after Hinton asked him to take part in a project focused on a new way to train artificial neural networks.





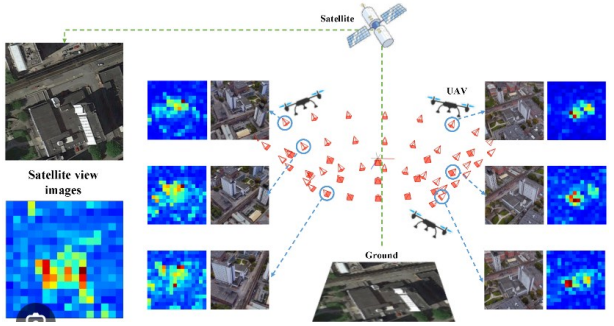
APPLICATIONS

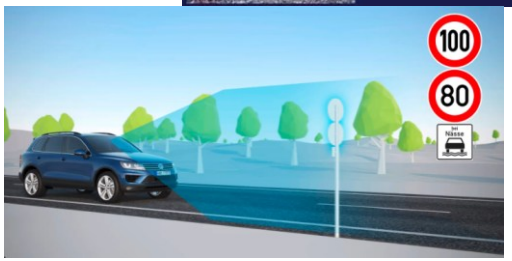
- **Natural language processing (NLP)**
  - Answering questions
  - Speech recognition
  - Summarizing documents
  - Classifying documents
  - Finding names, dates, etc. in documents
  - Searching for articles mentioning a concept

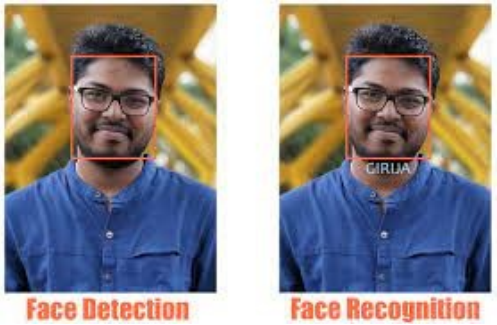


APPLICATIONS

- **Computer vision**
  - Satellite and drone imagery interpretation (e.g., for disaster resilience),
  - Face detection,
  - Face recognition,
  - Image captioning,
  - Reading traffic signs,
  - Locating pedestrians and
  - Vehicles in autonomous vehicles



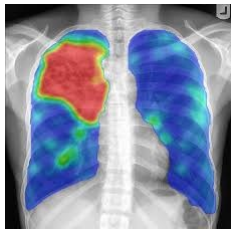


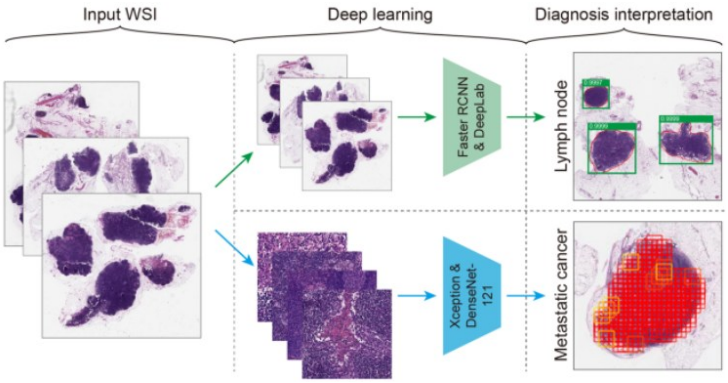


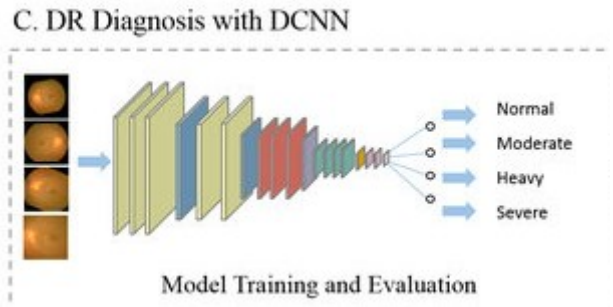
APPLICATIONS

Medicine

- Finding anomalies in radiology images, including CT, MRI, and X-ray images
- Counting features in pathology slides
- Diagnosing diabetic retinopathy

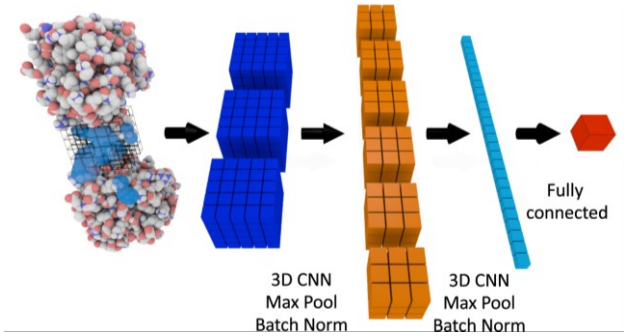


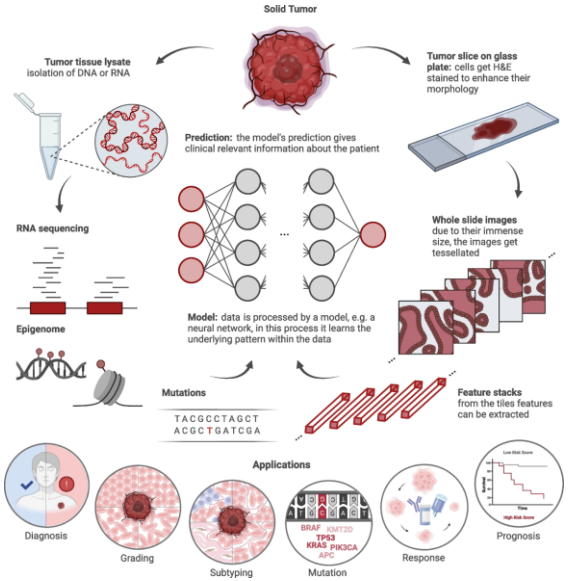




APPLICATIONS

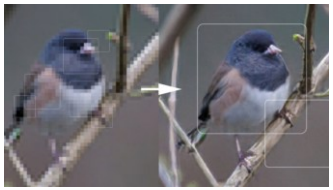
- **Biology**
  - Classifying proteins
  - Genomics tasks such as:
    - tumor-normal sequencing
    - classifying clinically actionable genetic mutations
  - Cell classification
  - Analyzing protein/protein interactions





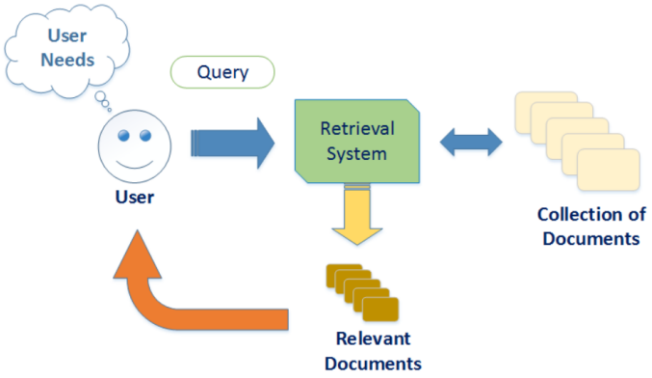
APPLICATIONS

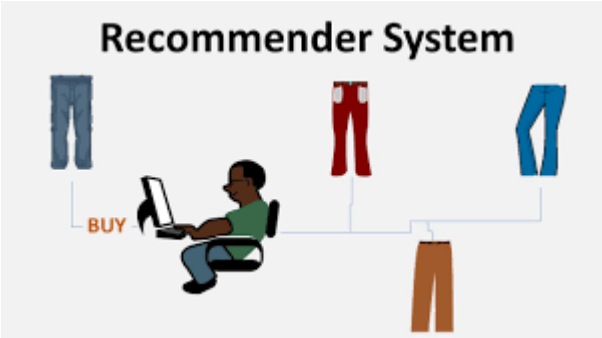
- **Image generation**
  - Colorizing images
  - Increasing image resolution
  - removing noise from images,
  - converting images to art in the style of famous artists



APPLICATIONS

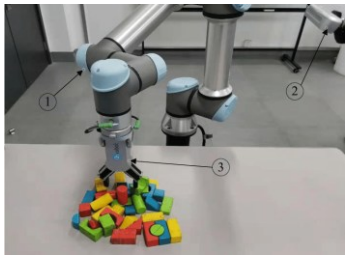
- **Recommendation systems**
  - Web search
  - product recommendations





APPLICATIONS

- **Playing games**
  - Chess, Go, most Atari video games, and many real-time strategy games
- **Robotics**
  - Handling objects that are challenging to locate or hard to pick up



- **Other applications**
  - Financial and logistical forecasting, text to speech, and much, much more…