# Deep Learning

Pournami P N

**Associate Professor**

Department of Computer Science and Engineering

National Institute of Technology Calicut

NIT Campus Post, Kerala

pournamipn@nitc.ac.in

Deep Learning - An MIT Press book - Ian Goodfellow and Yoshua Bengio and Aaron Courville

# Introduction

- Inventors have long dreamed of creating machines that think - dates back to at least the time of ancient Greece.

- When programmable computers were first conceived - such machines might become intelligent.

- Today Artificial intelligence (AI) is a thriving field
  - intelligent software to automate routine labor
  - understand speech or images
  - make diagnoses in medicine
  - support basic scientific research

# Challenges

- Early days of artificial intelligence
  - rapidly tackled and solved problems that are intellectually difficult for human beings but relatively straight-forward for computers
  - problems that can be described by a list of formal, mathematical rules.
  - *IBM's Deep Blue chess-playing system defeated world champion Garry Kasparov in 1997*
  - Chess can be completely described by a very brief list of completely formal rules, easily provided ahead of time by the programmer.
- The t**rue challenge to artificial intelligence**
- **Solving the tasks that are easy for people to perform but hard for people to describe formally**
  - **problems that we solve intuitively, that feel automatic, like recognizing spoken words or faces in image**

# Challenges

- Abstract and formal tasks that are among the most difficult mental undertakings for a human being are among the easiest for a computer.
- Computers defeat even the best human chess player - but only recently matching some of the abilities of average human beings to recognize objects or speech.
- A person's everyday life requires an immense amount of knowledge about the world, which is subjective and intuitive, and therefore difficult to articulate in a formal way.
- Computers need to capture this same knowledge in order to behave in an intelligent way.
- *One of the key challenges in artificial intelligence is how to get this informal knowledge into a computer.*

# Knowledge base Approach to AI

- Hard-code knowledge about the world in formal languages

- A computer can reason automatically about statements in these formal languages using logical inference rules.
  - Cyc (Lenat and Guha, 1989) failed to understand a story about a person named Fred shaving in the morning (Linde, 1992)

# Solution

- Allow computers to learn from experience

- Understand the world in terms of a hierarchy of concepts, with each concept defined through its relation to simpler concepts.

- This approach avoids the need for human operators to formally specify all the knowledge that the computer needs.

- The hierarchy of concepts enables the computer to learn complicated concepts by building them out of simpler ones.

- This is done by extracting patterns from raw data - known as **_MACHINE LEARNING_** (subjective decisions by computer programs)

# Learning Algorithm

## Learning

**What is the Learning Problem?**

Learning = Improving with experience at some task
   Improve over task T ,
   with respect to performance measure P ,
   based on experience E.
E.g., Learn to play checkers
   T : Play checkers
   P : % of games won in world tournament
   E: opportunity to play against self

Lecture slides for textbook Machine Learning, T. Mitchell, McGraw Hill, 1997

- A simple machine learning algorithm called logistic regression can determine whether to recommend cesarean delivery (Mor-Yosef et al., 1990).
- A simple machine learning algorithm called naive Bayes can separate legitimate e-mail from spam e-mail

# Representation of Data

- The **performance of the machine learning algorithms** depends on the **representation of the data** they are given.
  - When logistic regression is used to recommend cesarean delivery, the AI system does not examine the patient directly.
  - Instead, the doctor tells the system several pieces of relevant information, such as the presence or absence of a uterine scar.
  - Each piece of information included in the representation of the patient is known as a **feature**.
  - Logistic regression learns how each of these features of the patient correlates with various outcomes

# Representation of Data

- The choice of representation has an enormous effect on the performance of machine learning algorithms



Representations Matter

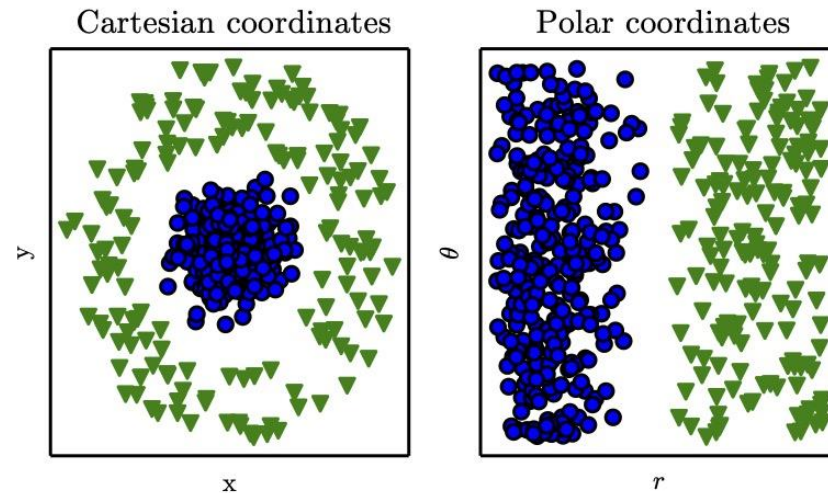Cartesian coordinates    Polar coordinates

Figure 1.1

(Goodfellow 2016)

- Many artificial intelligence tasks - designing the right set of features to extract for that task, providing these features to a simple machine learning algorithm.
  - Speaker identification from sound - an estimate of the size of the speaker's vocal tract.
  - This feature gives a strong clue - the speaker is a man, woman, or child.

# Representation of Data

- **What features should be extracted?**
  - Write a program to detect cars in photographs.

- **Solution - Representation Learning**
  - Use machine learning to discover not only the mapping from representation to output but also the representation itself.
  - Enable AI systems to rapidly adapt to new tasks, with minimal human intervention.

# Factors of Variation

- When designing features or algorithms for learning features, our goal is usually to *separate the factors of variation* that explain the observed data.

- Many of the factors of variation influence every single piece of data we are able to observe.

- Many of these factors of variation, such as a speaker's accent, can be identified only using sophisticated, nearly human-level understanding of the data.

- **Deep learning solves this central problem in representation learning by introducing representations that are expressed in terms of other, simpler representations.**
  - **Multi-Layer Perceptron (MLP)**

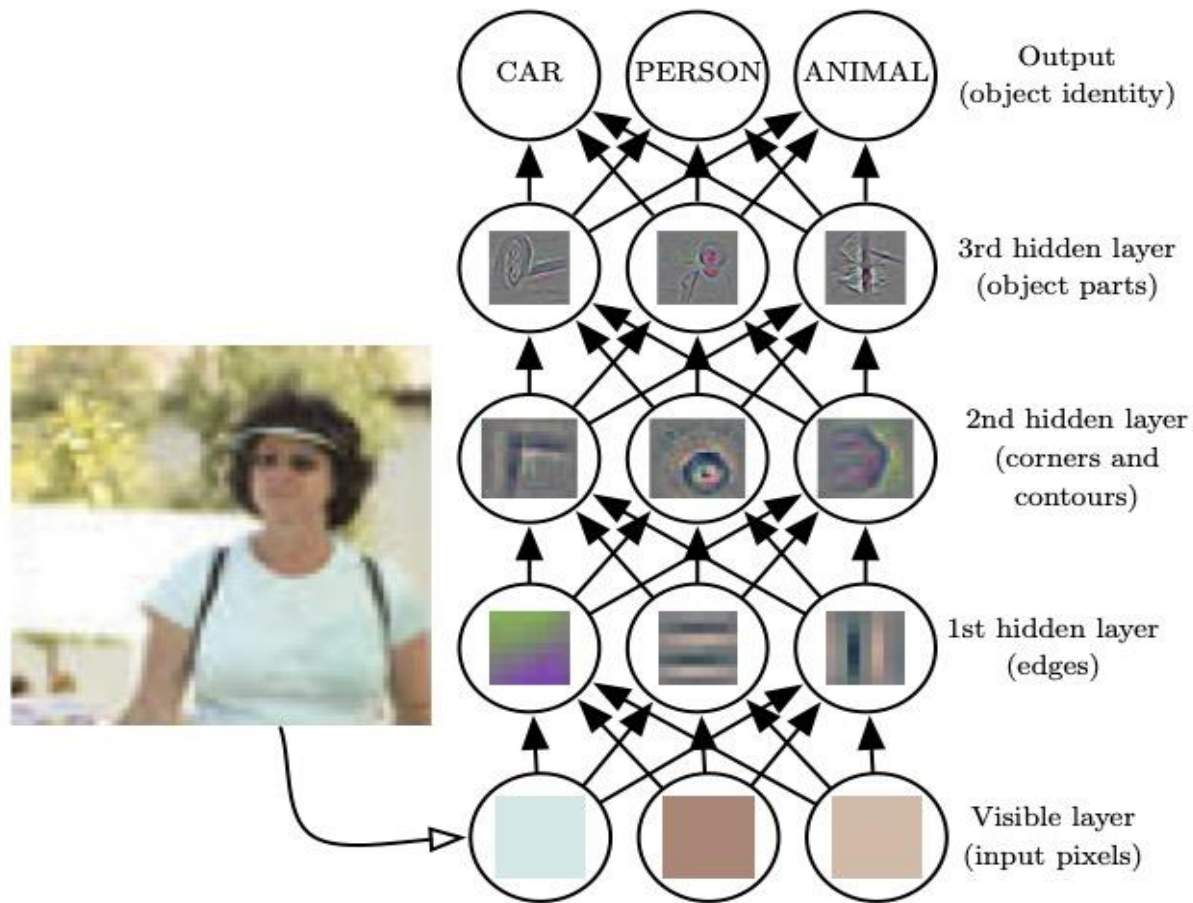# Depth: Repeated Composition



Figure 1.2

(Goodfellow 2016)

Zeiler and Fergus (2014)

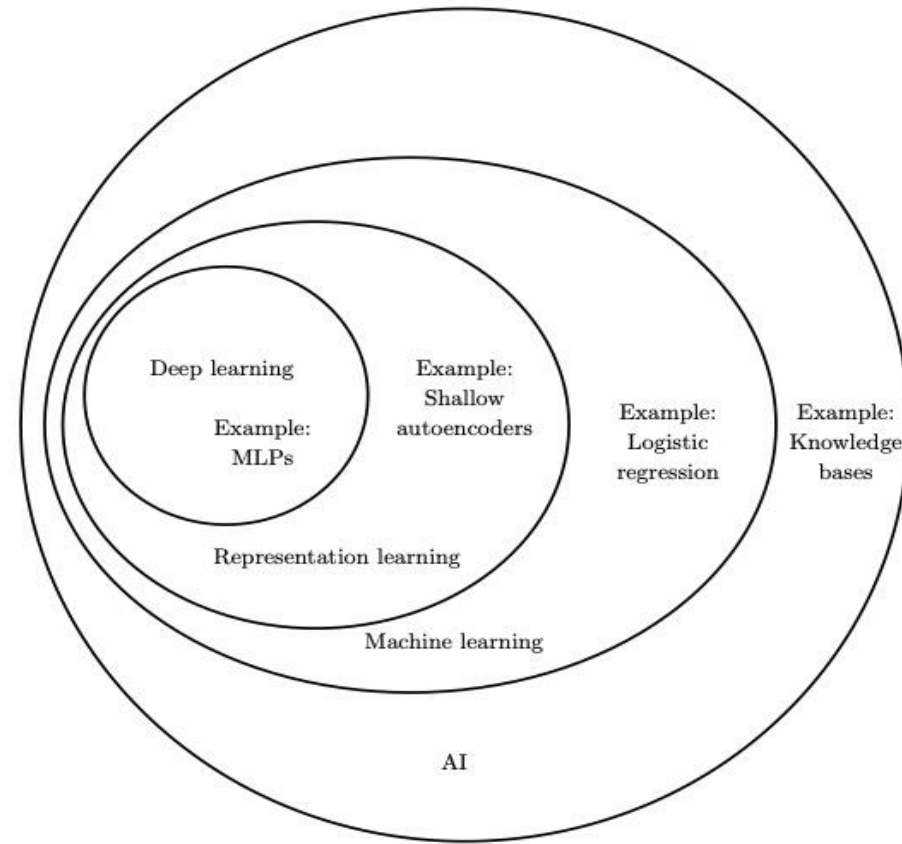Deep learning is an approach to AI
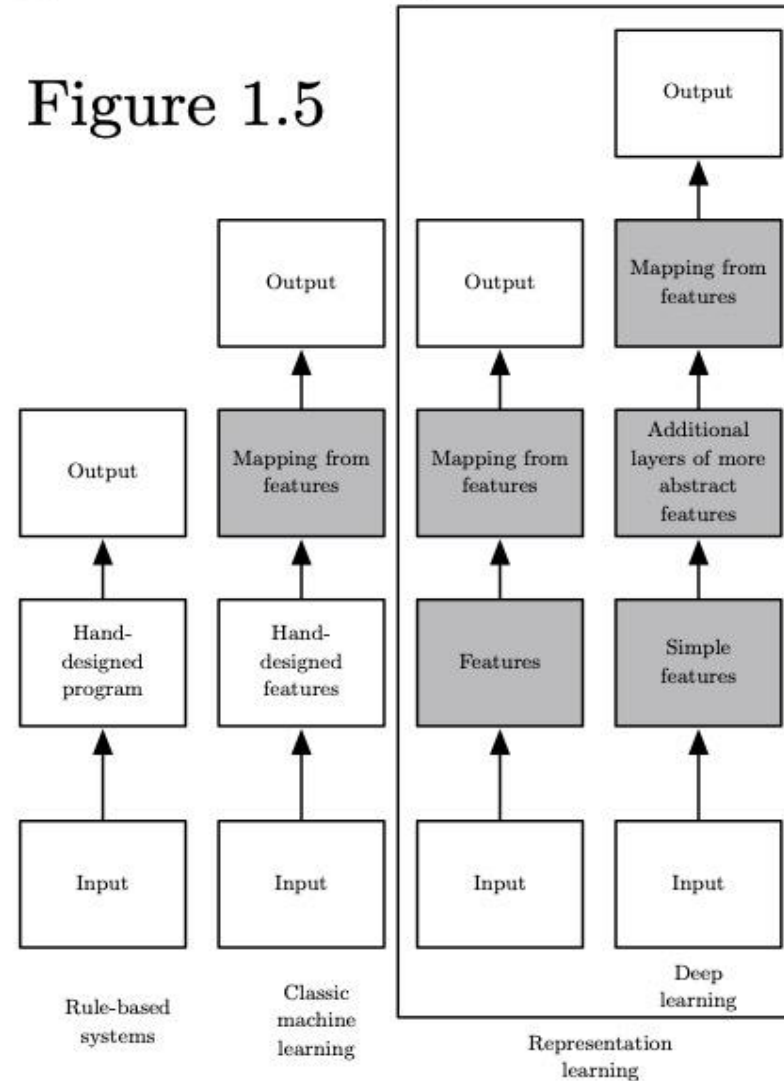


Machine Learning and AI

Figure 1.4

(Goodfellow 2016)

# Learning Multiple Components

## Parts of an AI system

Figure 1.5



(Goodfellow 2016)

# Conclusion

- Machine learning is the only viable approach to build AI systems that can operate in complicated real-world environments.

- Deep learning is a particular kind of machine learning
  - achieves great power and flexibility by representing the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts

- Deep Learning makes it possible for more abstract representations computed in terms of less abstract ones.