The Five Big Ideas in AI

David S. Touretzky
Carnegie Mellon University

What is Artificial Intelligence?

Artificial Intelligence is a
branch of computer science

…

What is Artificial Intelligence?

Artificial Intelligence is a
branch of computer science

concerned with techniques that
allow computers to do things

…

What is Artificial Intelligence?

Artificial Intelligence is a
branch of computer science

concerned with techniques that
allow computers to do things

that are generally considered
to **require intelligence**.

We Use AI-Powered Technologies Every Day

- Speech recognition
  - Home assistants: "Alexa, please dim the lights."
  - Automated video subtitles; automated transcription services (otter.ai, sonix.ai)
- Computer vision
  - Use your face to unlock your iPhone; play with Snapchat and Tiktok filters
  - Self-driving cars: Tesla autopilot; automated lane departure warnings; emergency braking
- Language understanding
  - Google search: "What does an alligator weigh?" "What's the second largest city in Honduras?"
  - Google Translate: automatic translation among any of 100+ languages; restaurant menus
- Recommender systems
  - Netflix→movies; Amazon→products; Facebook→news stories; Google→ads
- Robotics
  - Roombas clean up; Kiva robots automate warehouses; drones follow you

AI Poses Challenges for Society

- Many types of work will become more highly automated.
  - Some people will **lose their jobs** -- including white collar jobs.
  - People will find that their **jobs have changed**; new skills are needed.
  - New **jobs will be created** around AI, robotics, and machine learning.

- Automated decision making systems can be opaque and biased.

- AI is powering the surveillance state:
  - Ubiquitous face recognition: loss of privacy
  - Automated monitoring of phone calls and social media

- Deep fakes: seeing is no longer believing; hard to know what to trust

- Super-intelligent large language models might wipe out humanity

Five Big Ideas in AI

1. **Perception:** Computers perceive the world using sensors.

2. **Representation and reasoning:** Agents maintain representations of the world and use them for reasoning.

3. **Learning:** Computers can learn from data.

4. **Natural interaction:** Intelligent agents require many kinds of information to interact naturally with humans.

5. **Societal impact:** AI can impact society in both positive and negative ways.
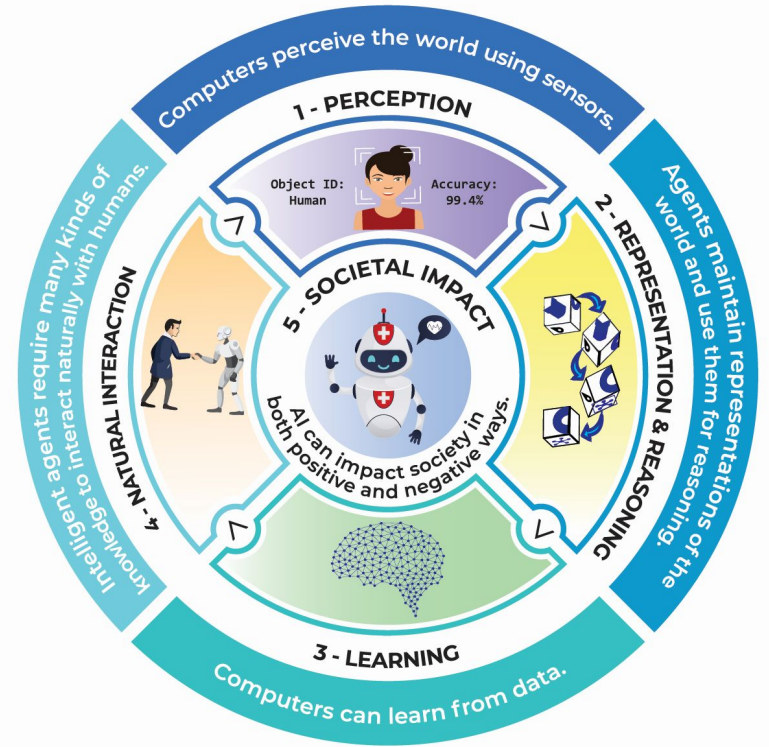
Five Big Ideas in Artificial Intelligence v.2

1. Perception

Computers perceive the world using sensors. Perception is the process of extracting meaning from sensory signals. Making computers "see" and "hear" well enough for practical use is one of the most significant achievements of AI to date.

5. Societal Impact

AI can impact society in both positive and negative ways. AI technologies are changing the ways we work, travel, communicate, and care for each other. But we must be mindful of the harms that can potentially occur. For example, biases in the data used to train an AI system could lead to some people being less well served than others. Thus, it is important to discuss the impacts that AI is having on our society and develop criteria for the ethical design and deployment of AI-based systems.

2. Representation & Reasoning

Agents maintain representations of the world and use them for reasoning. Representation is one of the fundamental problems of intelligence, both natural and artificial. Computers construct representations using data structures, and these representations support reasoning algorithms that derive new information from what is already known. While AI agents can reason about very complex problems, they do not think the way a human does.

4. Natural Interaction

Intelligent agents require many kinds of knowledge to collaborate and interact naturally with humans. Ideally, agents will converse with us using natural language, draw upon cultural knowledge to infer intentions from observed behavior, and respond appropriately to body language, facial expressions, and emotions. Advances in deep neural networks such as large language models and convolutional neural networks are making this possible.
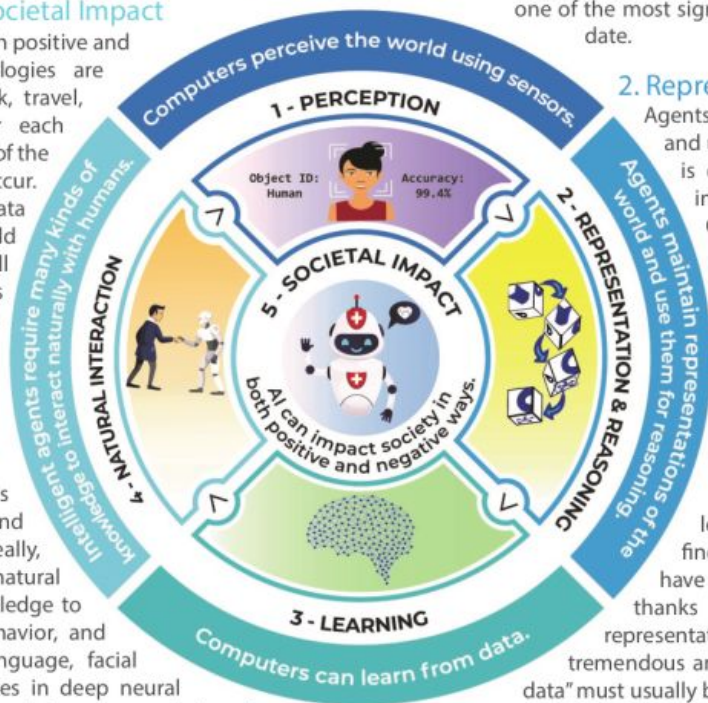
3. Learning

Computers can learn from data. Machine learning is a kind of statistical inference that finds patterns in data. Many areas of AI have progressed significantly in recent years thanks to learning algorithms that create new representations. For the approach to succeed, tremendous amounts of data are required. This "training data" must usually be supplied by people, but is sometimes acquired by the machine itself.

Computers perceive the world using sensors.

1 - PERCEPTION

Object ID: Human   Accuracy: 99.4%

Agents maintain representations of the world and use them for reasoning of the

2 - REPRESENTATION & REASONING

5 - SOCIETAL IMPACT

AI can impact society in both positive and negative ways.

Intelligent agents require many kinds of knowledge to interact naturally with humans.

4 - NATURAL INTERACTION

3 - LEARNING

Computers can learn from data.

AI4K12.org

Big Idea #1: Perception

Computers perceive the world using sensors.

Perception is the extraction of *meaning* from sensory signals using knowledge.

- Human senses vs. computer sensors

- Types of perception: vision, speech recognition etc.

- How perception works: algorithms
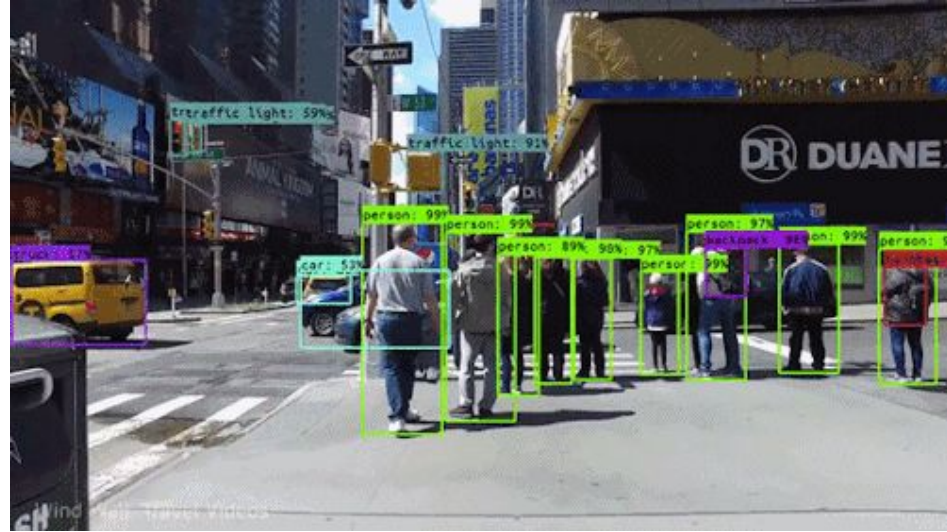
Example Guidelines
- Identify sensors on computers, robots, and intelligent appliances.
- Explain how sensor limitations affect computer perception.
- Explain that perception systems may draw on multiple algorithms as well as multiple sensors.
- Build an application using multiple sensors and types of perception (possibly with Scratch plugins, or Calypso).

Are Supermarket Doors Intelligent?

This is what you get when your automatic doors have sensing but not perception.

There is a deer in Aldi....

Big Idea #2: Representation and Reasoning

Agents maintain representations of the world, and use them for reasoning.

- Types of representations

- Families of algorithms and the work they do

- Representations support reasoning: algorithms operate on representations

Example Guidelines
- Create/design a representation of an (animal) classification system using a tree structure.
- Draw a search tree for tic-tac-toe
- Describe how AI representations support reasoning to answer questions
- Describe the differences between types of search algorithms

Big Idea #3: Learning

Computers can learn from data.

- Nature of learning

- Fundamentals of neural networks

- Data sets
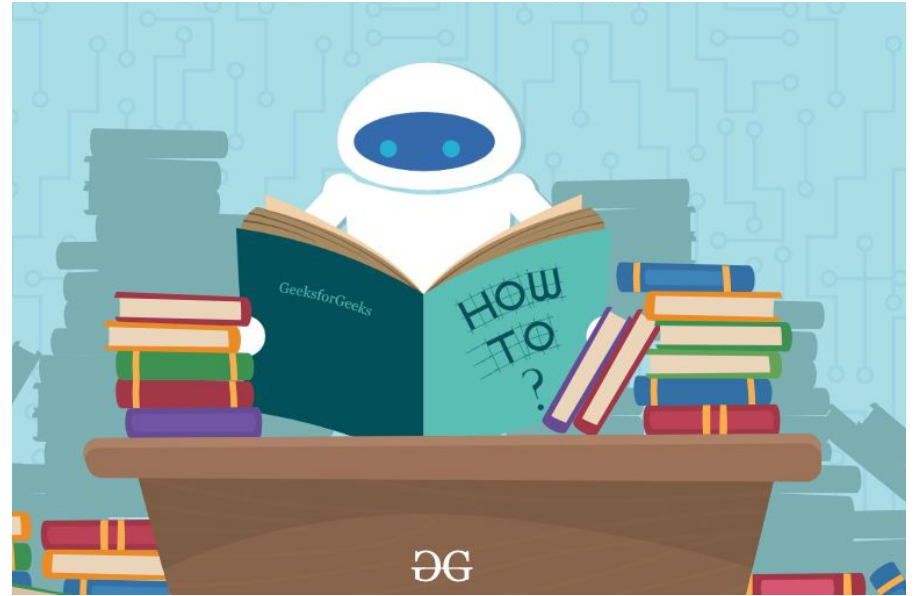
Example Guidelines
- Modify an interactive machine learning project by training its model..
- Describe how algorithms and machine learning can exhibit biases.
- Identify bias in a training data set and extend the training set to address the bias
- Train a neural net (1-3 layers) using *TensorFlow Playground*
- Trace and experiment with a simple ML algorithm

AI and Machine Learning
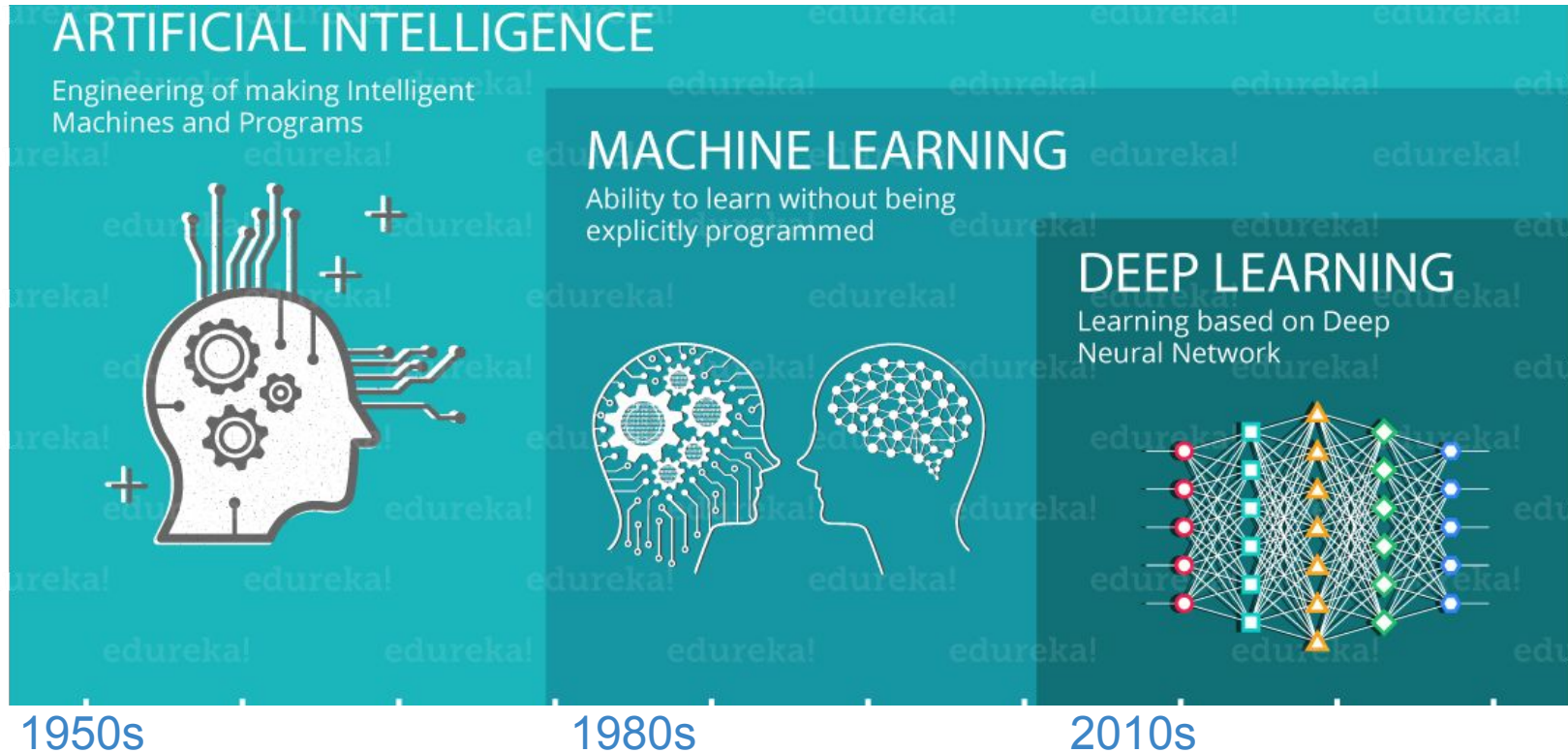
Big Idea #4: Natural Interaction

Intelligent agents require many kinds of knowledge to interact naturally with humans.

- Natural language understanding

- Common sense reasoning

- Affective computing & interaction (e.g. with robots, or speech agents)

- Consciousness and philosophy of mind

Example Guidelines
- Recognize and label facial expressions into appropriate emotions (happiness, sadness, anger) and explain why they are labeled the way they are
- Experiment with software that recognizes emotions in facial expressions
- Construct a simple chatbot
- Describe some tasks where AI outperforms humans, and tasks where it does not
- Explain and give examples of how language can be ambiguous
- Reason about the nature of intelligence, and identify approaches to determining whether an agent is or is not intelligent.

Big Idea #5: Societal Impact (1 of 3)

*"Artificial Intelligence can impact society in both
positive and negative ways."*

- **Ethics of AI making decisions about people**
  - Fairness and bias
  - Transparency and explainability
  - Accountability

Example Guidelines
- Critically explore the positive and negative impacts of an AI system.

- Describe ways that AI systems can be designed for inclusivity.
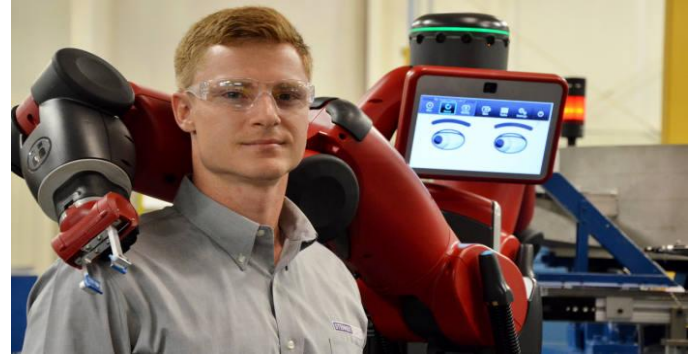

Machine Bias: ProPublica.org

Big Idea #5: Societal Impact (2 of 3)

"Artificial Intelligence can impact society in both positive and negative ways."

- **Economic impacts of AI**
  - Increased productivity
  - New types of services
  - Reduction in of some types of jobs
  - New career opportunities

Example Guidelines
- Design and explain how an AI system can be used to address a social issue.

- Understand tradeoffs in the design of AI systems and how decisions can have unintended consequences in the function of a system.
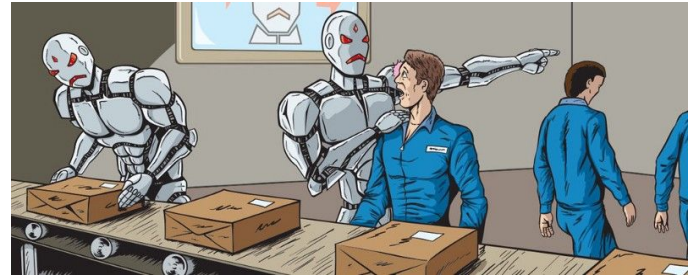
"Artificial Intelligence can impact society in both positive and negative ways."

- **AI & Culture**
  - Living with intelligent assistants and robot companions.
  - Would you let your child travel unaccompanied in a self-driving car?
  - New YouTube genre: self-driving car mishaps.

Example Guidelines
- Critically explore the positive and negative impacts of an AI system.

- Describe the debate about whether people should be polite to agents and robots.

Teachable Machine

18

What is a neural network?

- A large, complex mathematical function that maps inputs to outputs.
  - Because the function is so complex, the network can do sophisticated things.

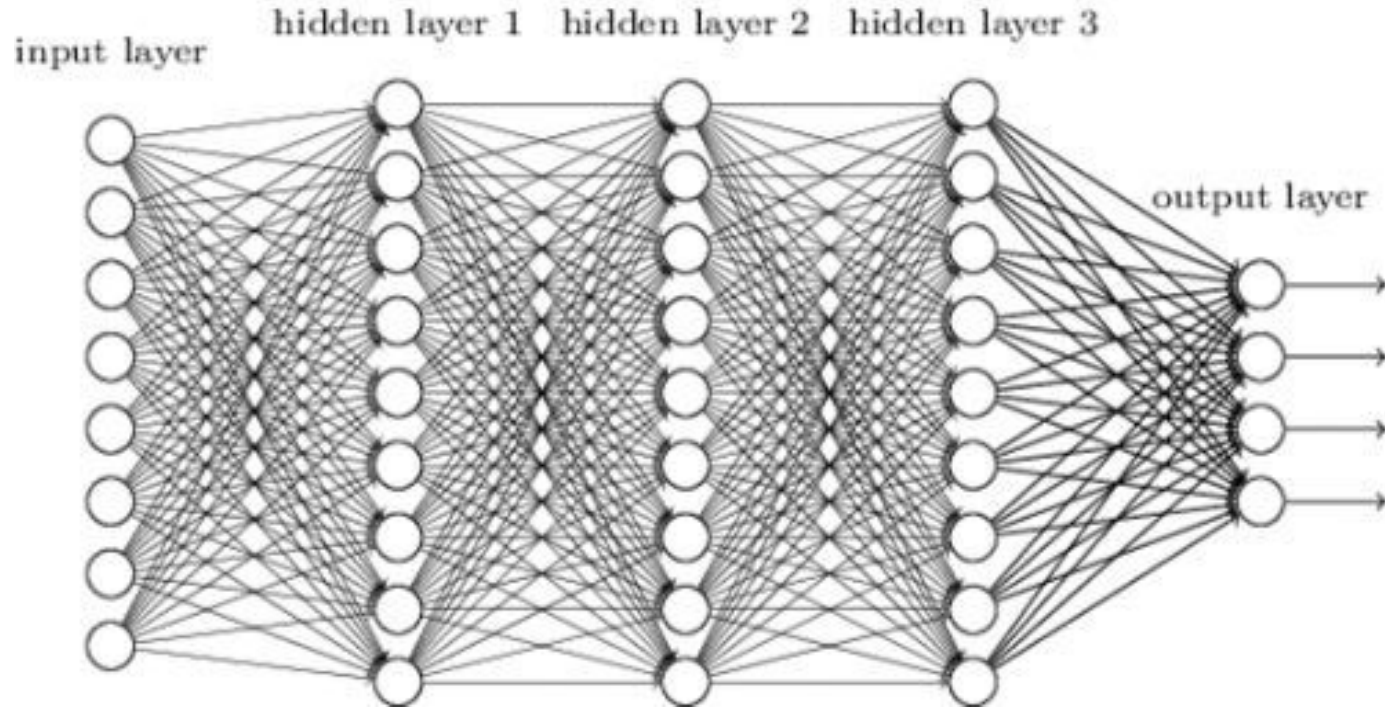- Each layer is composed of many simple functions, called "units" or "neurons".
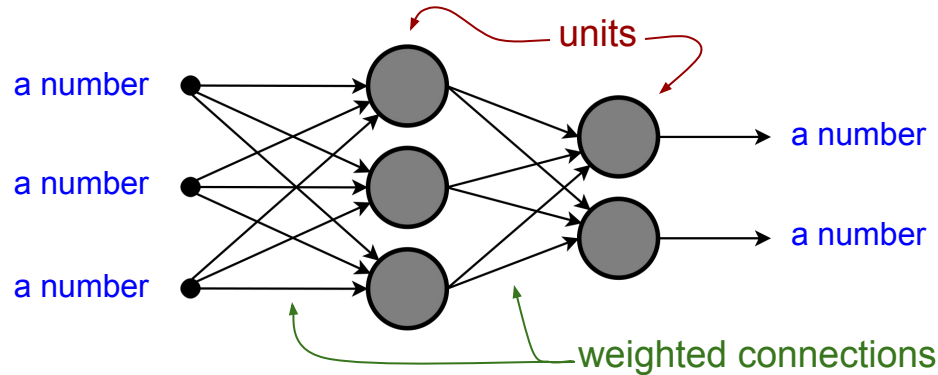
- Each unit takes multiple numbers as input and produces a number as output.

Neural Networks Are Organized In Layers

Input Layer

Hidden Layer(s)

Output Layer
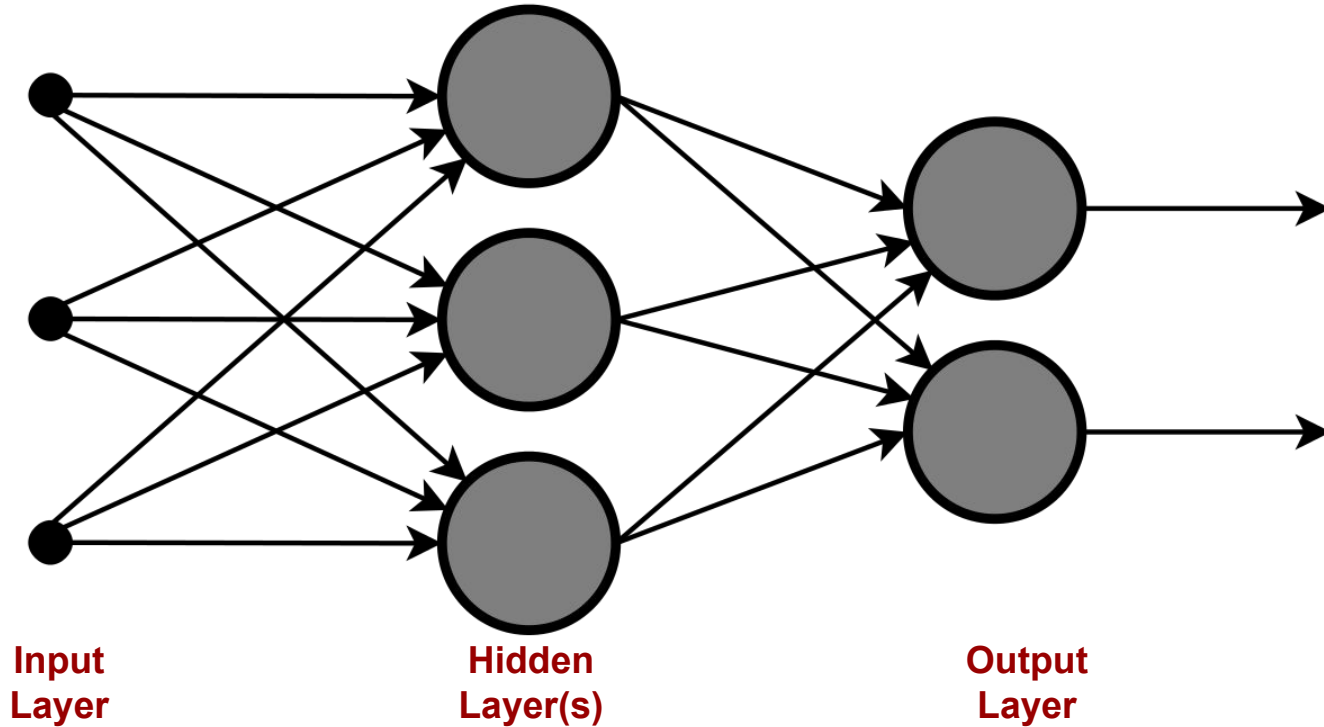
Deep Neural Networks Have Many Layers

The Weights Are Learned by Training on a Dataset

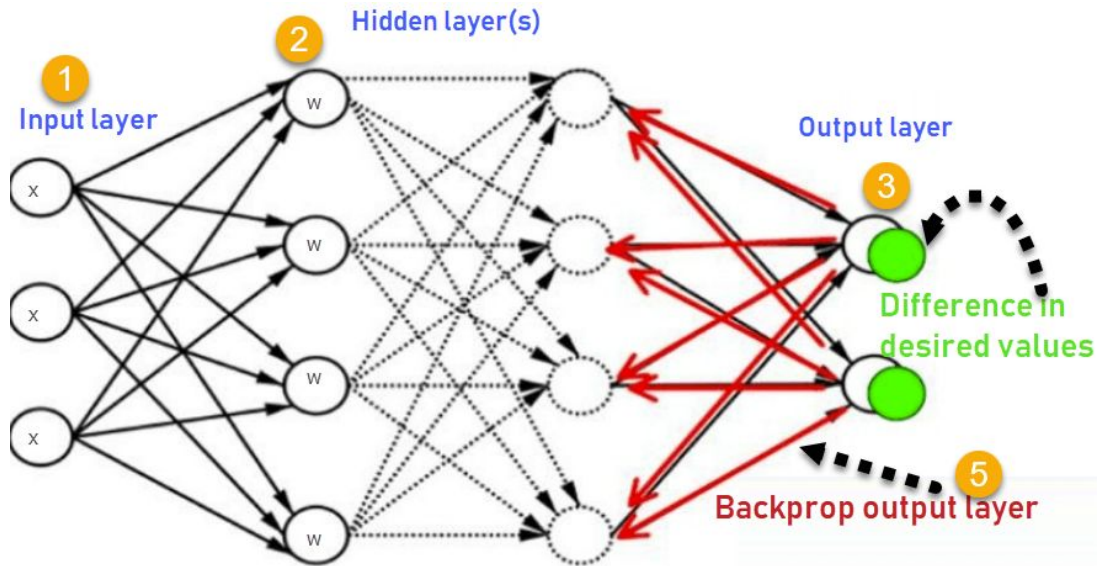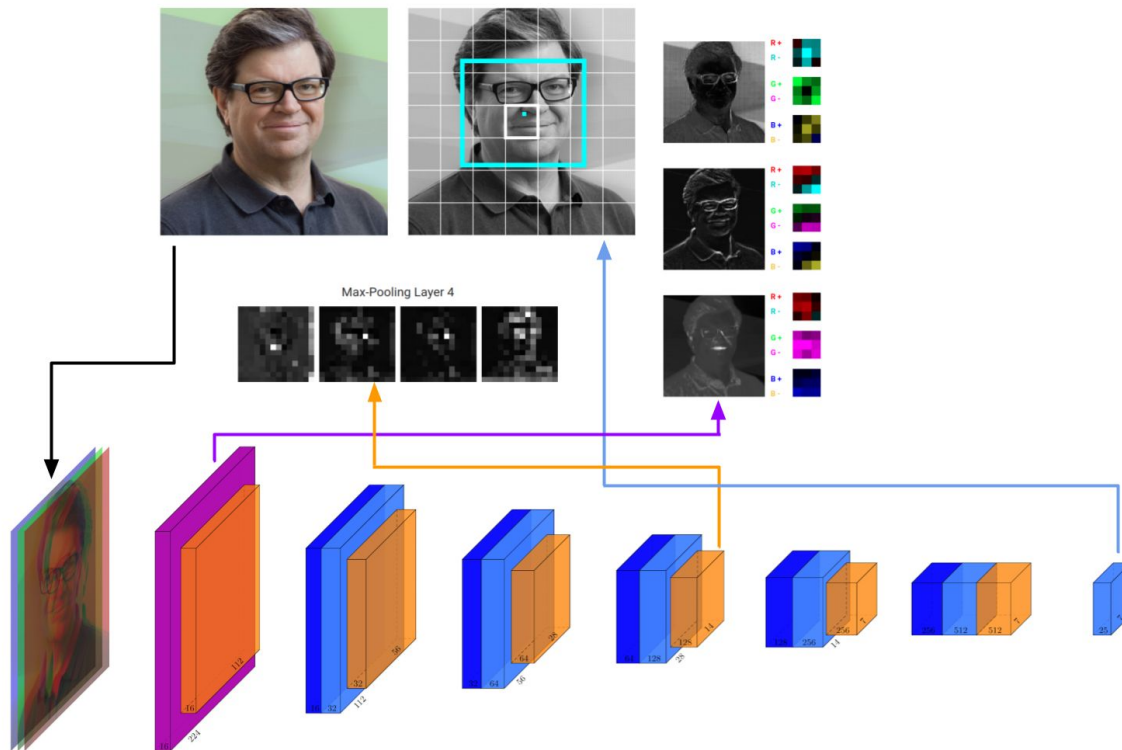"Backpropagation" is a learning algorithm for adjusting the weights.

Illustration from https://www.guru99.com/backpropogation-neural-network.html

Abstraction Pipeline In A Deep Neural Network

https://www.cs.cmu.edu/~dst/FaceDemo

Max-Pooling Layer 4

AI4K12.org Resource Directory

https://ai4k12.org/resources/list-of-resources/

Check out our resource directory for:

- Online demos
- Educational videos
- Curriculum resources
- Books
- and more!

Join Us in Growing the Community of
K-12 AI Educators and Resource Developers

Visit us:
https://AI4K12.org

Join the mailing list:
https://aaai.org/Organization/mailing-lists.php