Introduction to Deep Learning

AI+X
INSTITUTE FOR ARTIFICIAL INTELLIGENCE

What is Deep Learning?

- It is a branch of machine learning
- Machine learning is the field of study where algorithms build predictive or descriptive models from data
- Deep learning models are a form of neural networks

What is a Neural Network?

- Very loosely inspired by the human brain
- You have neurons which have an activation function
- You have connections between neurons (like synapses)
- They can learn a model given a labeled or unlabeled data set

Neural Network

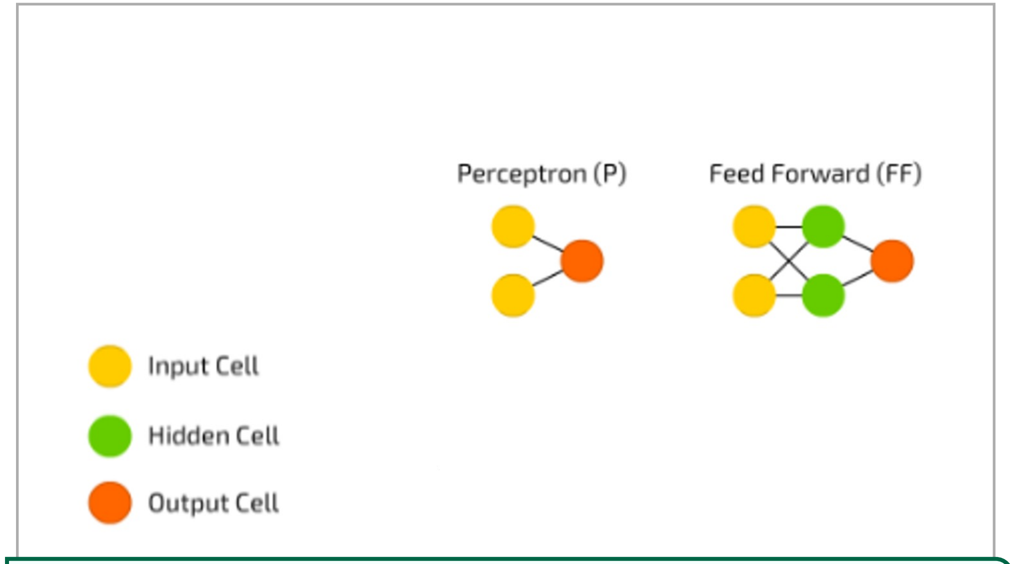

Figure 1. Perceptron (P) and Feed Forward (FF) network.

- The Perceptron (P) can only work well for linearly separable data
- Feed Forward (FF) backpropagation neural networks are universal approximators.
- However, we do not know the transfer function, configuration OR if we can learn the function for a given problem.

Are neural networks new?

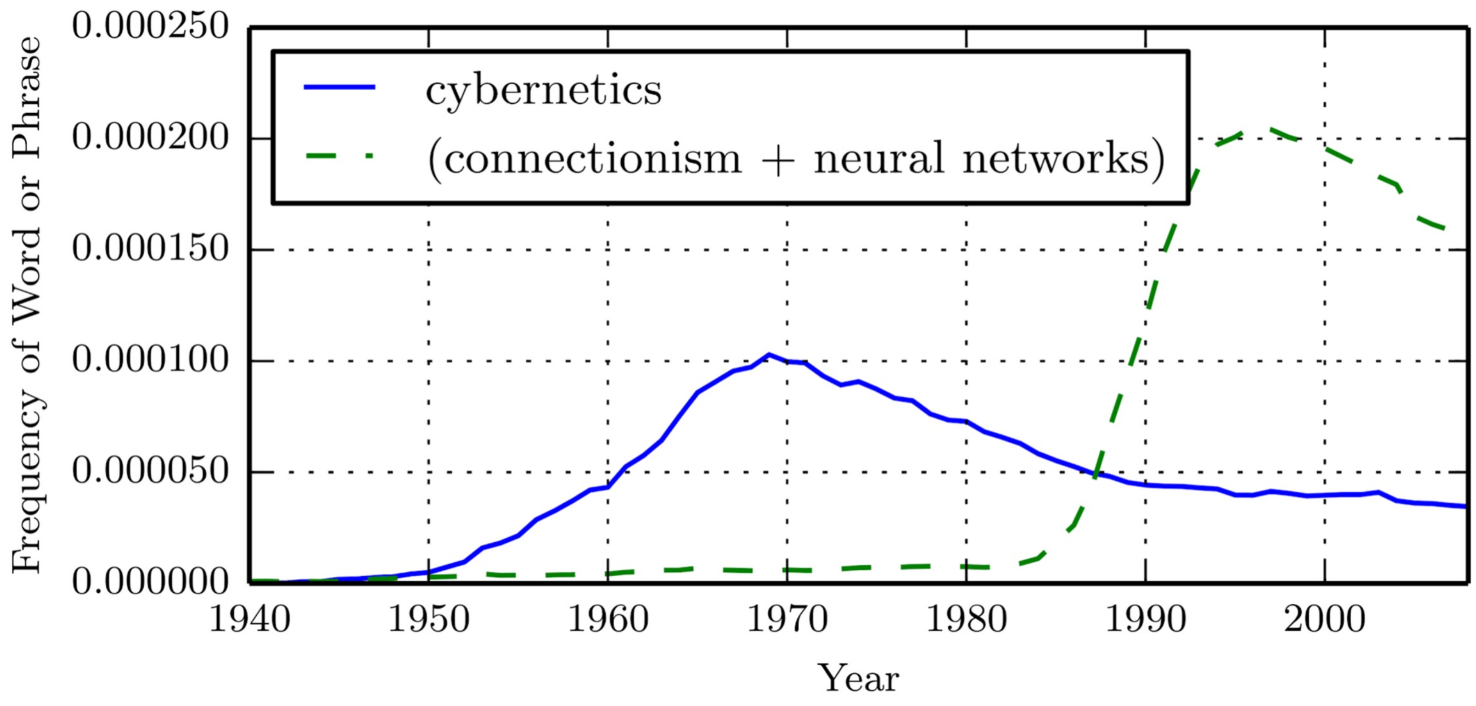

Figure 2. Two of the three historical waves of artificial neural nets research.

Why neural networks over other methods?

- There is a theorem that says a neural network with a layer that is neither input nor output can arbitrarily closely approximate any function
- However, how to actually find the right architecture and parameters is unknown
- So, for a while other algorithms were used for most machine learning tasks

How deep networks got popular?

Figure 3. ImageNet challenge logo.

Performance on Object Classification

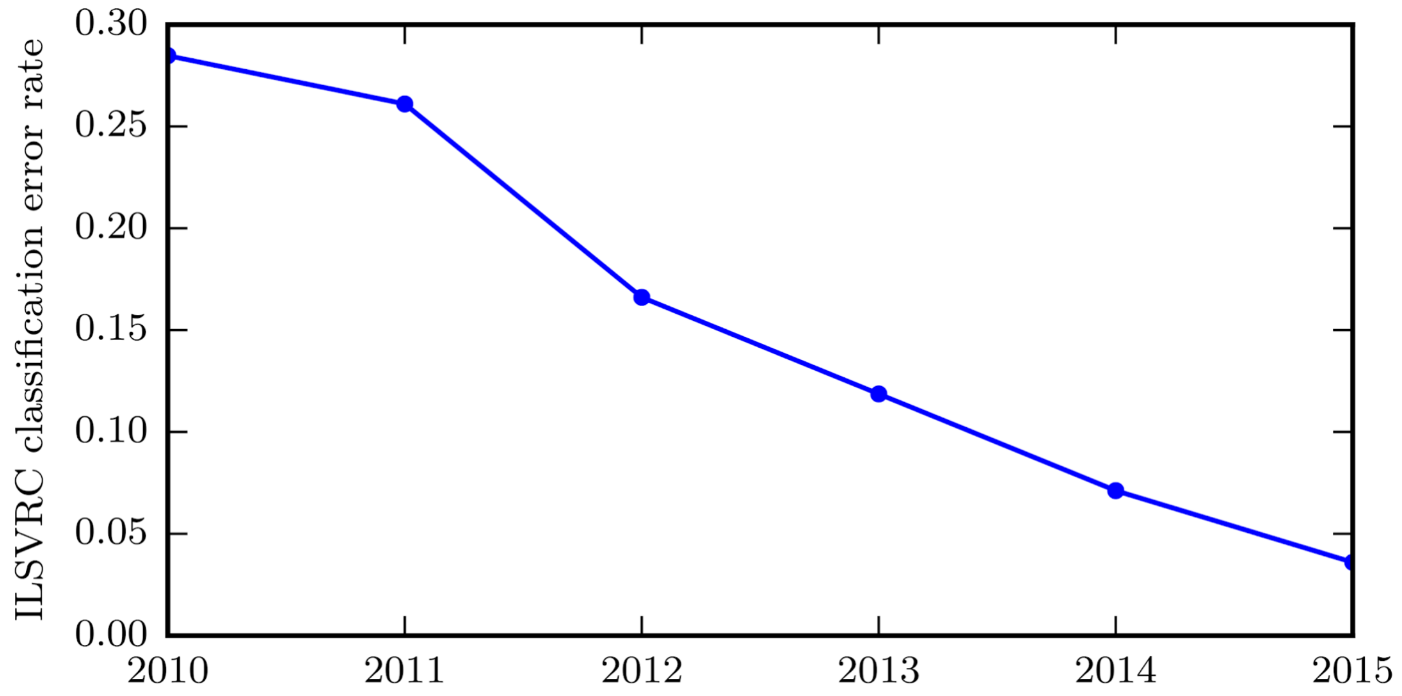

Figure 4. Decreasing error rate over time.

Knowledge Check 1

Can you guess the performance of the second best participant?

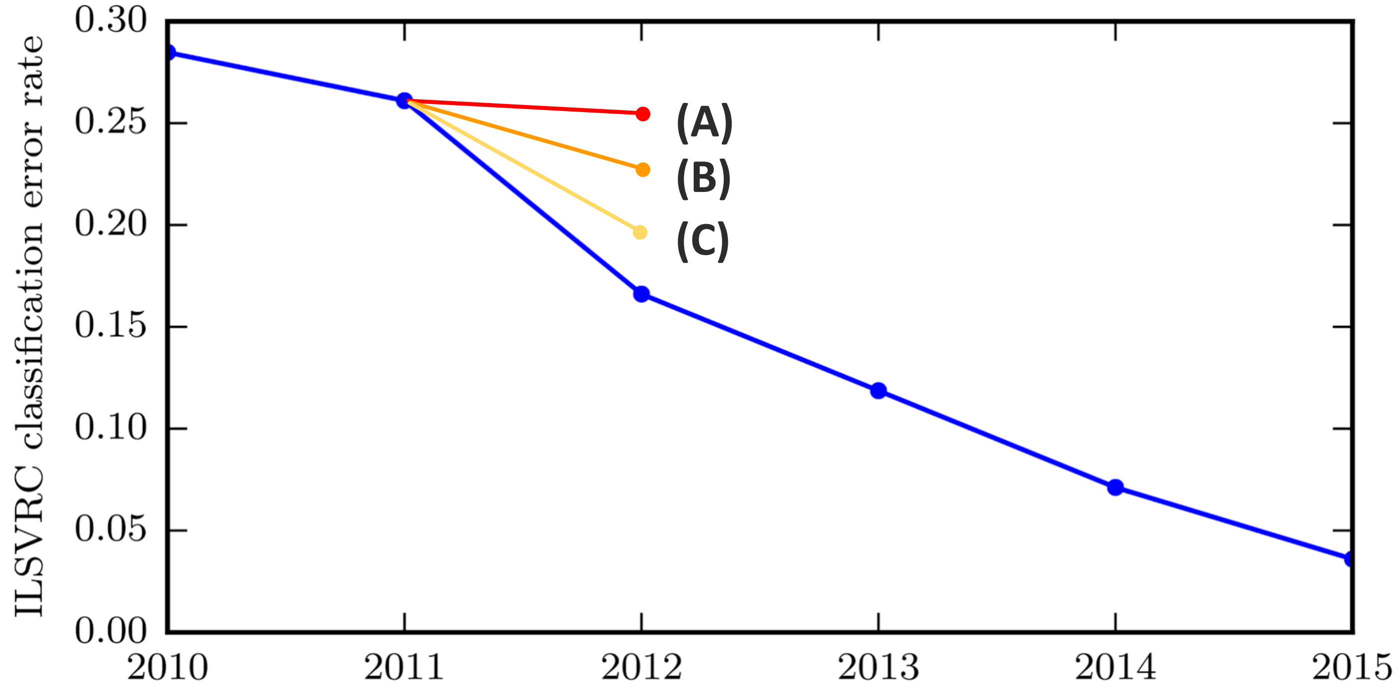

Figure 5. Decreasing error rate over time.

A

B

C

Performance on Object Classification

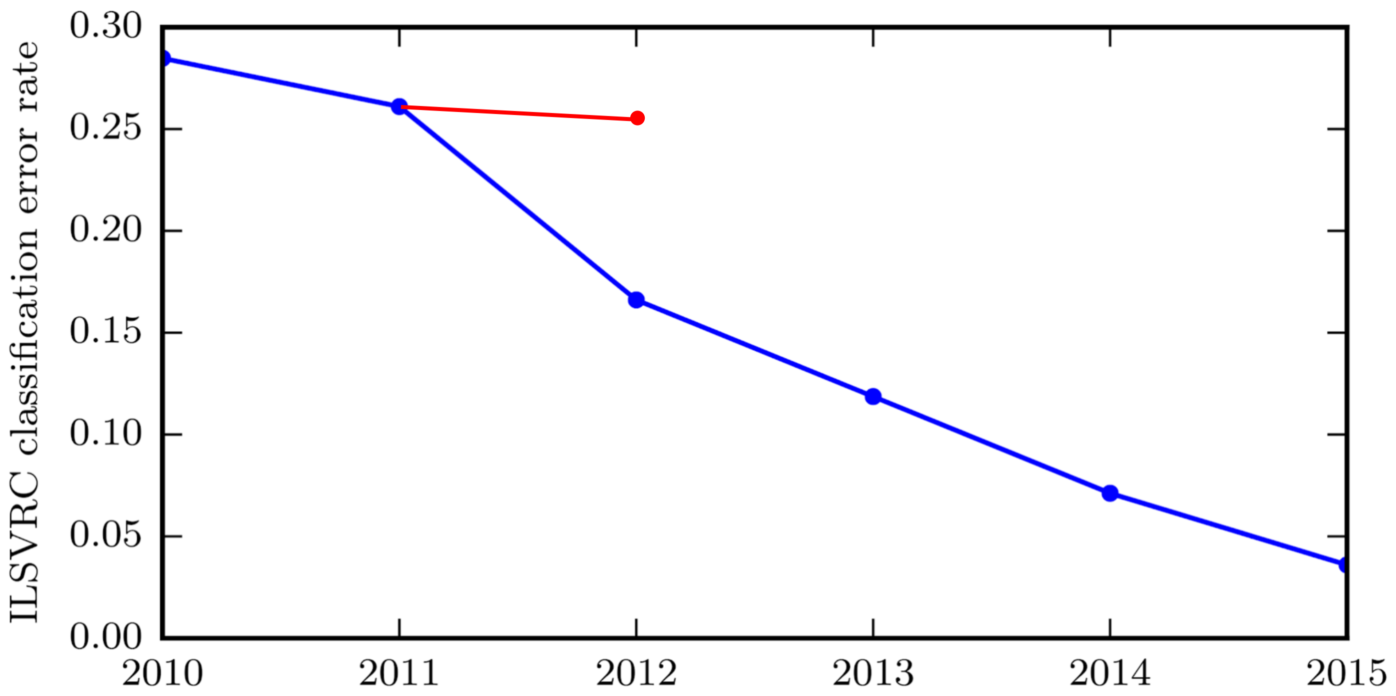

Figure 6. Decreasing error rate over time.

Performance on Object Classification

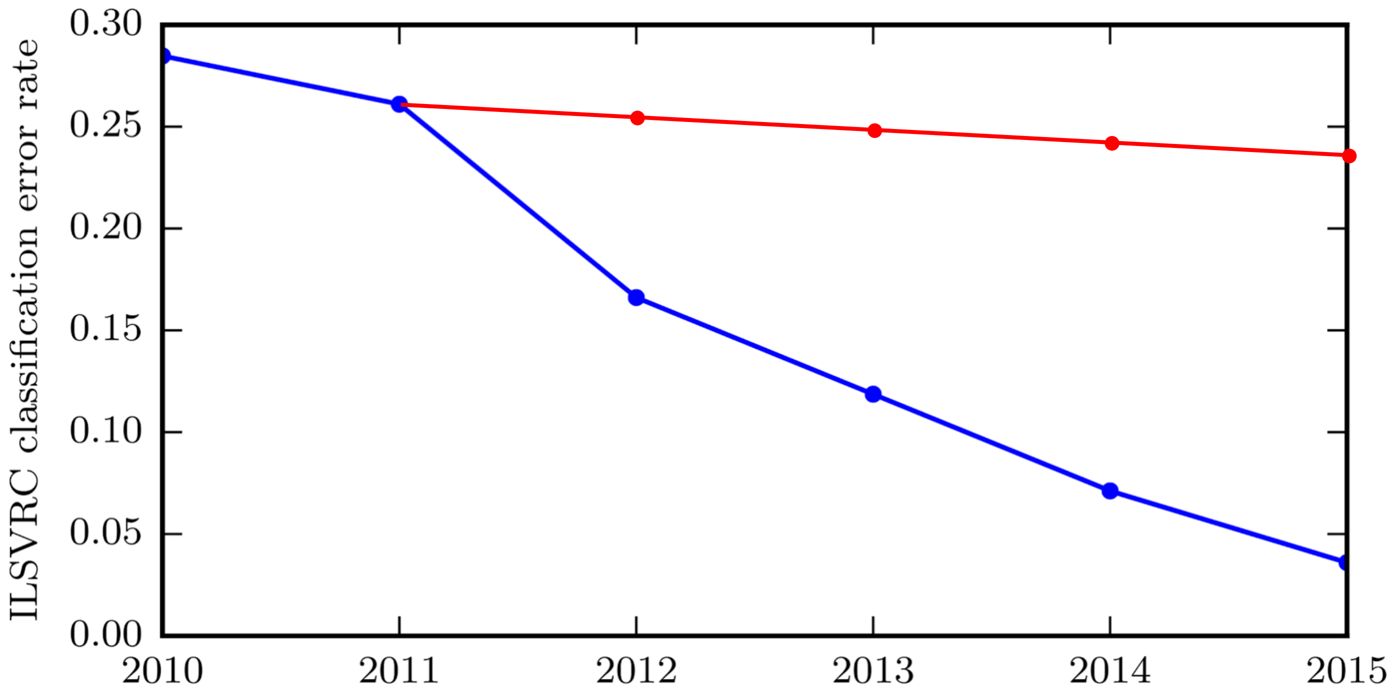

Figure 7. Decreasing error rate over time.

Neural Network

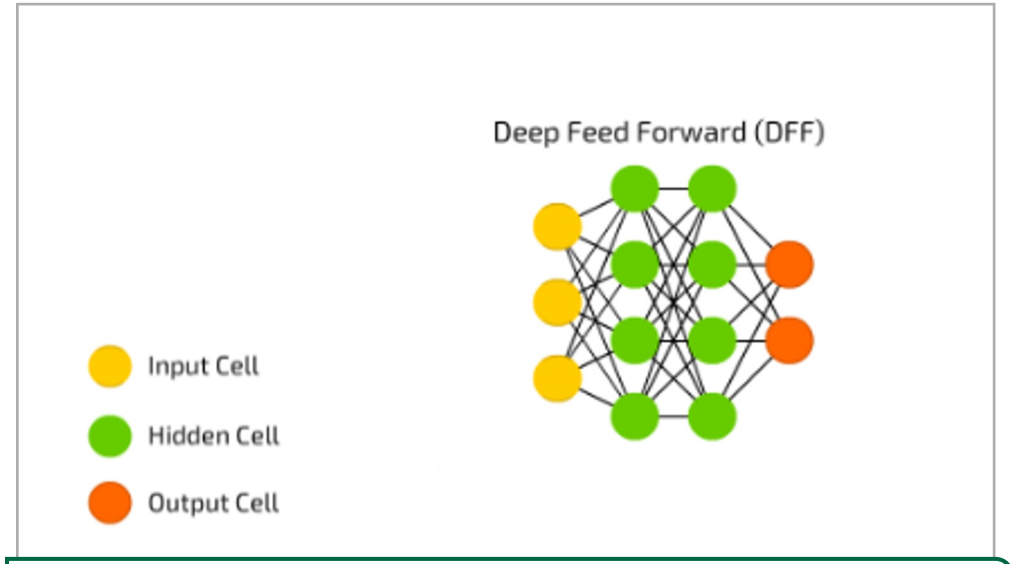

Figure 8. Deep Feed Forward (DFF) network.

- A Deep network has more than 1 hidden layer. A bit more brain like and often several more hidden layers.

Neural Network

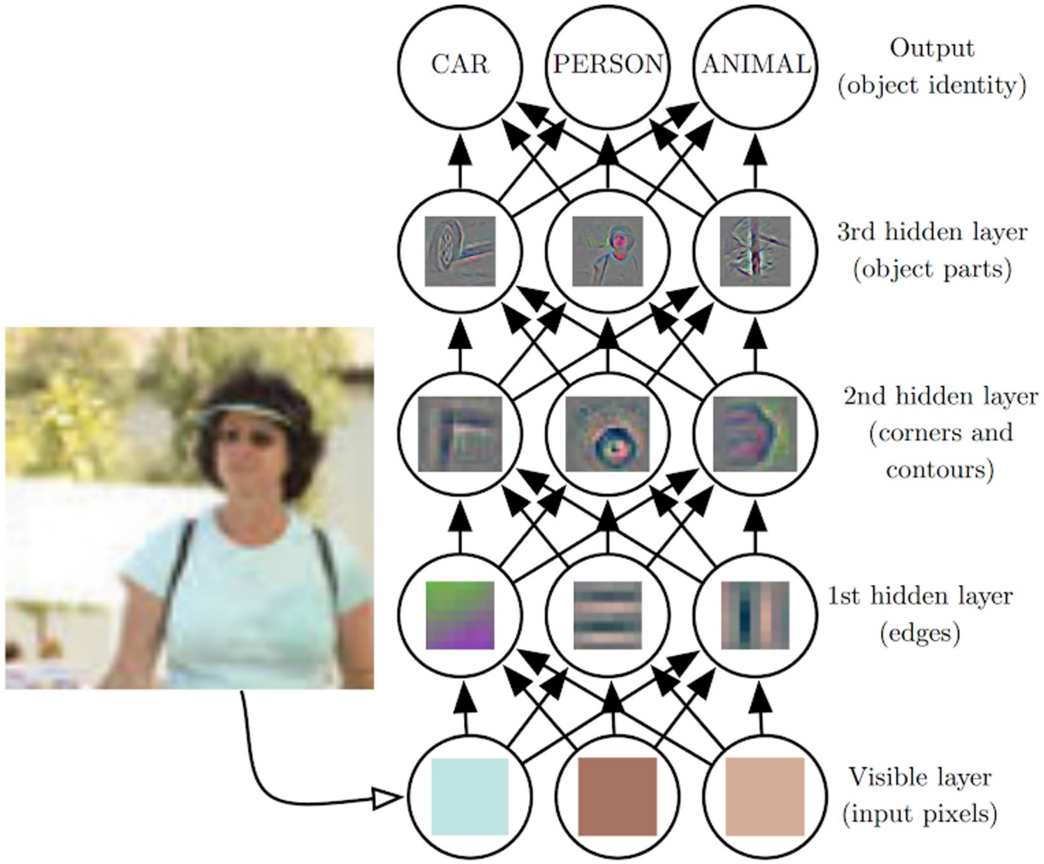

Figure 9. Illustration of a deep learning model.

Knowledge Check 2

For which of the following datasets can we use Linear Discriminant Analysis to distinguish between blue circles and green triangles?

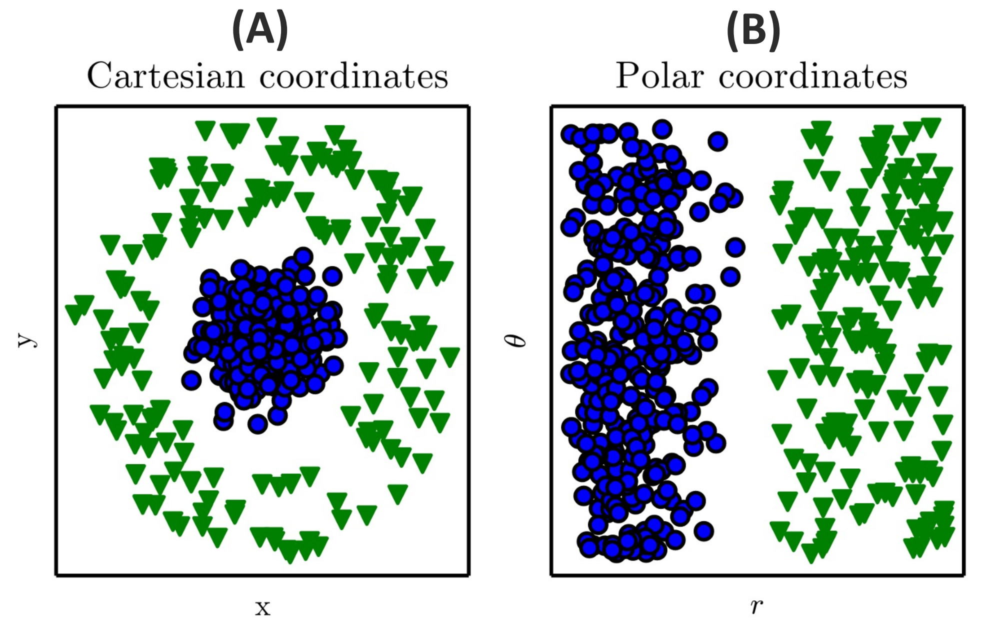

Figure 10. Cartesian Coordinates & Polar Coordinates

A Cartesian Coordinates

B Polar Coordinates

Why Deep Neural Networks?

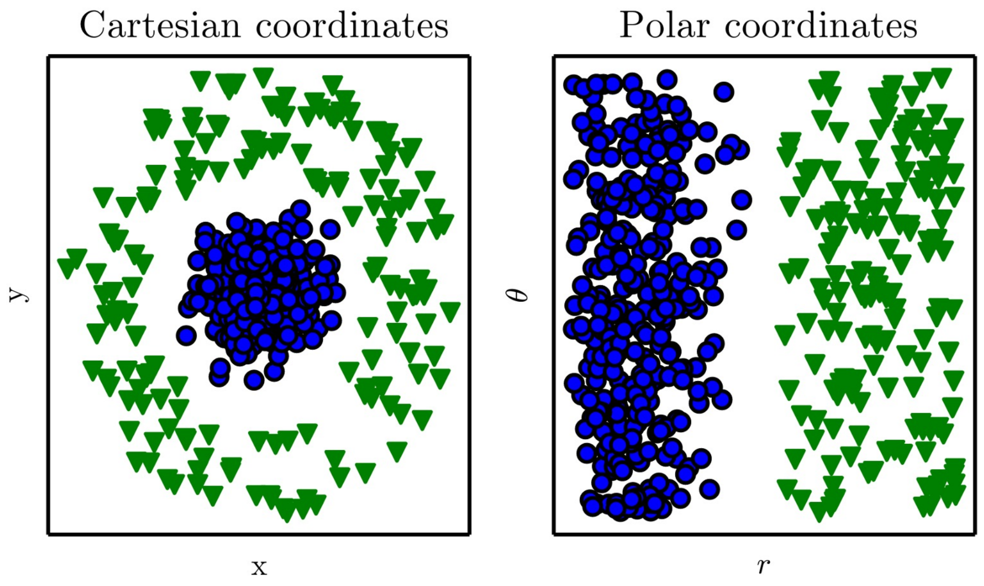

Figure 11. Different representations for the same data. Only one of them is linearly separable.

- No need of feature engineering

Why Deep Neural Networks?

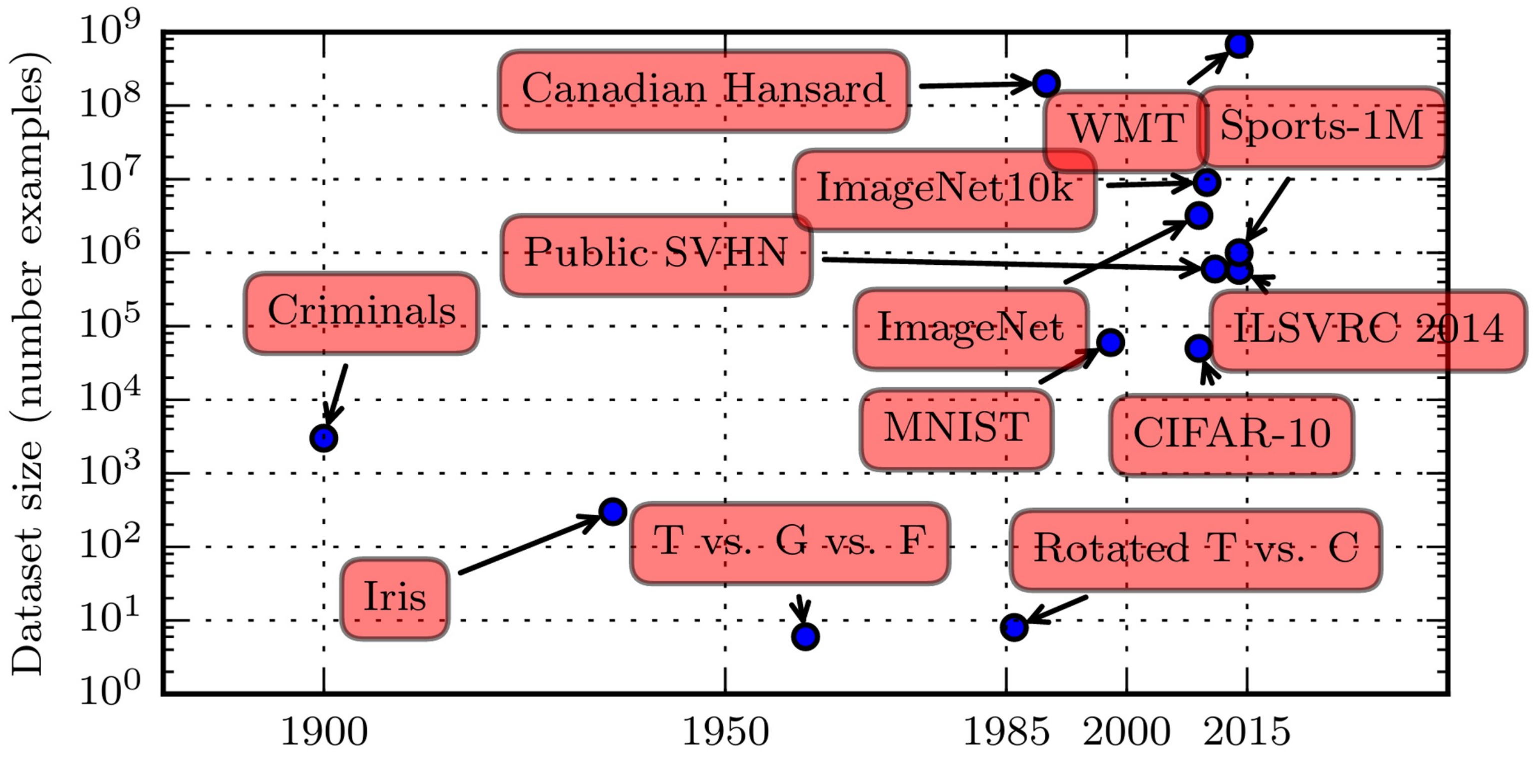

Figure 12. Increasing dataset size over time.

Why Deep Neural Networks?

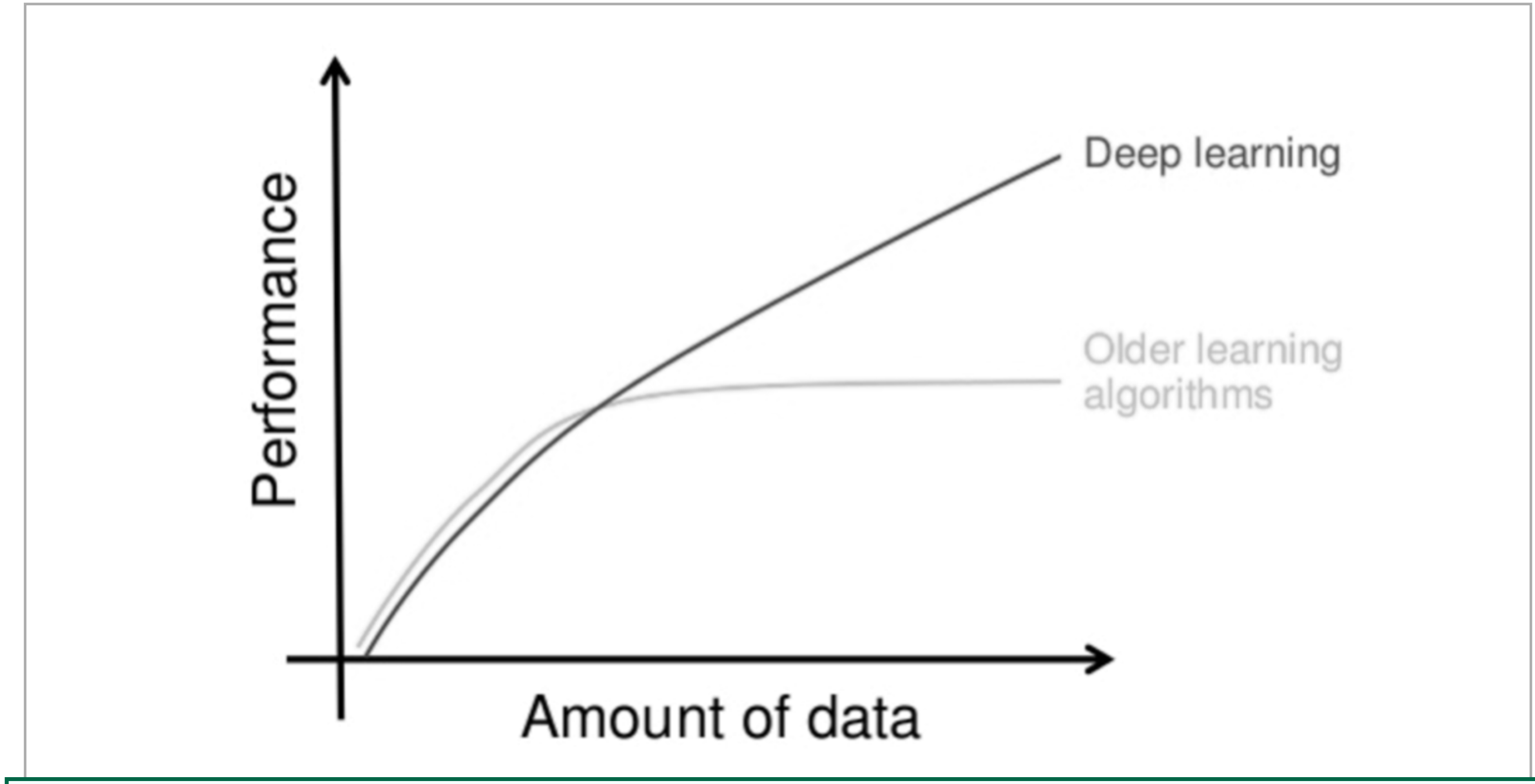

Figure 13. Deep learning approaches can handle larger amounts of data.

Why Deep Neural Networks?

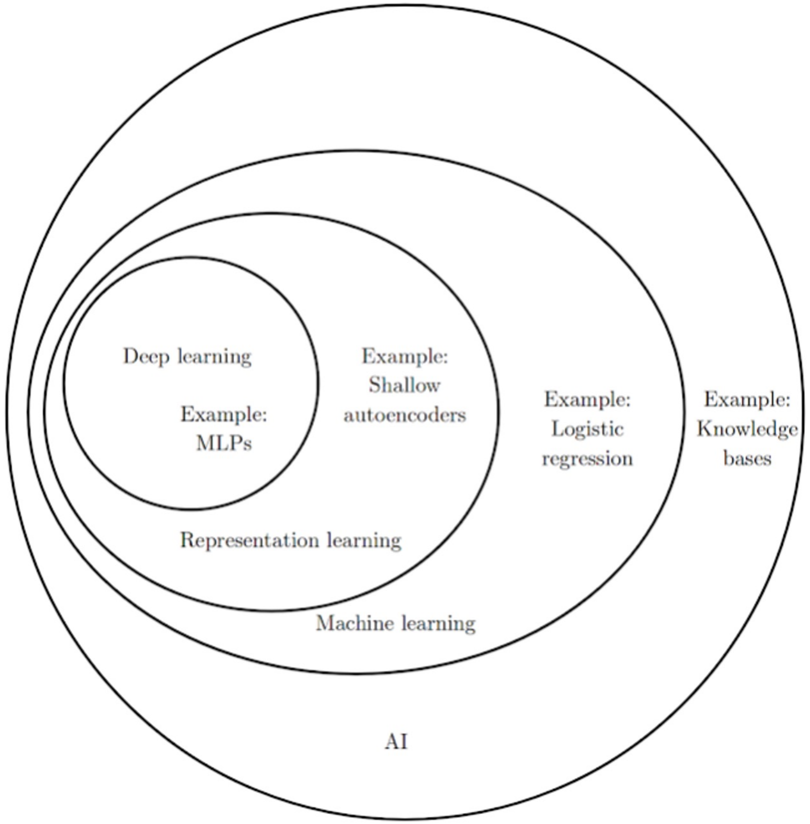

Figure 14. Deep learning is a kind of representation learning.

Why Deep Neural Networks?

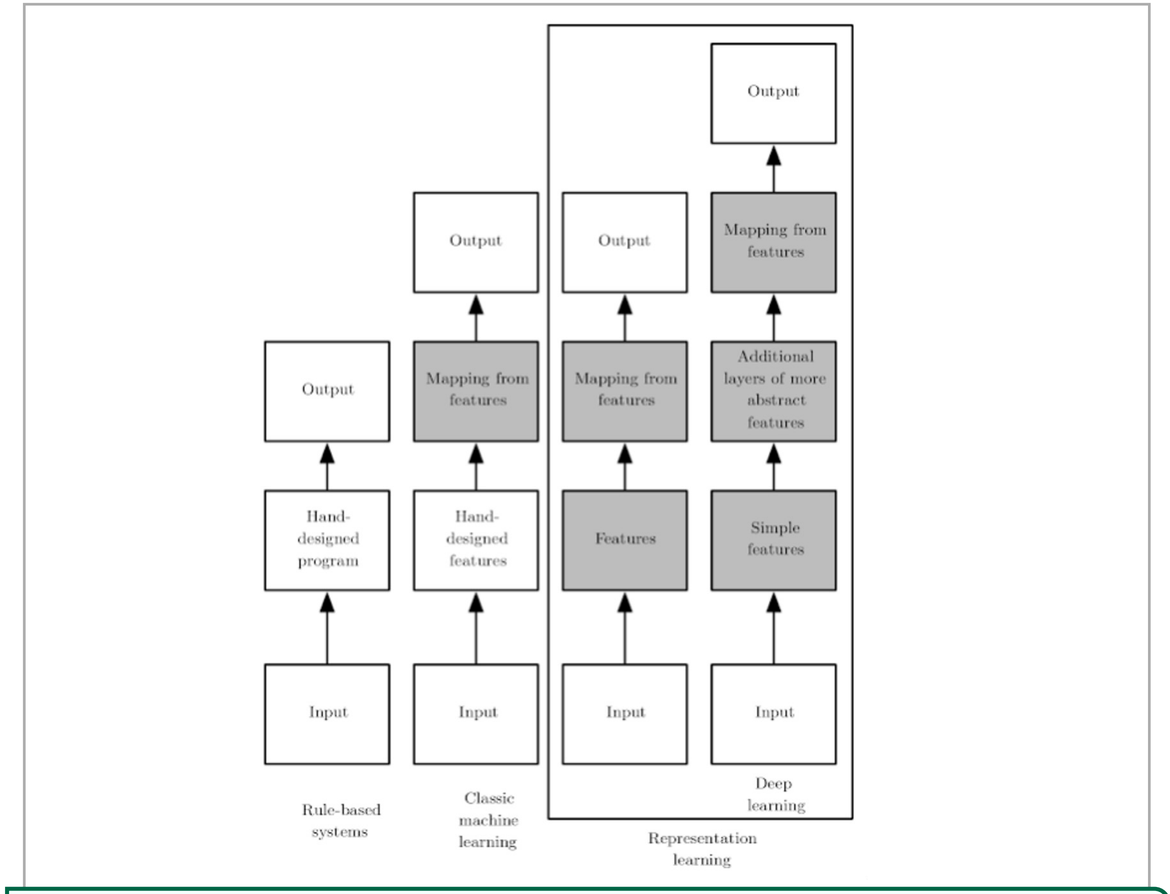

Figure 15. Flowcharts showing how the different parts of an AI system relate to each other within different AI disciplines..

Why Deep Neural Networks?

- No need of feature engineering
- Deep networks are way more accurate at challenging tasks like:
    - Image recognition
    - Speech recognition
    - Speech translation

Why Deep Neural Networks?

- “The complexity of the computer function grows exponentially with depth” [1]
  - More layers >>> bigger layers
- Thanks to parallel processing capability of GPU
  - Generative Pretrained Transformer 3 (GPT-3) [2]
  - 96 layers, 175B parameters
  - 355 years to train on a Tesla V100
  - ~$4,600,000 to train in the cheapest GPU cloud service
  - Supercomputer with more than 285,000 CPU cores & 10,000 GPUs

Why Deep Neural Networks?

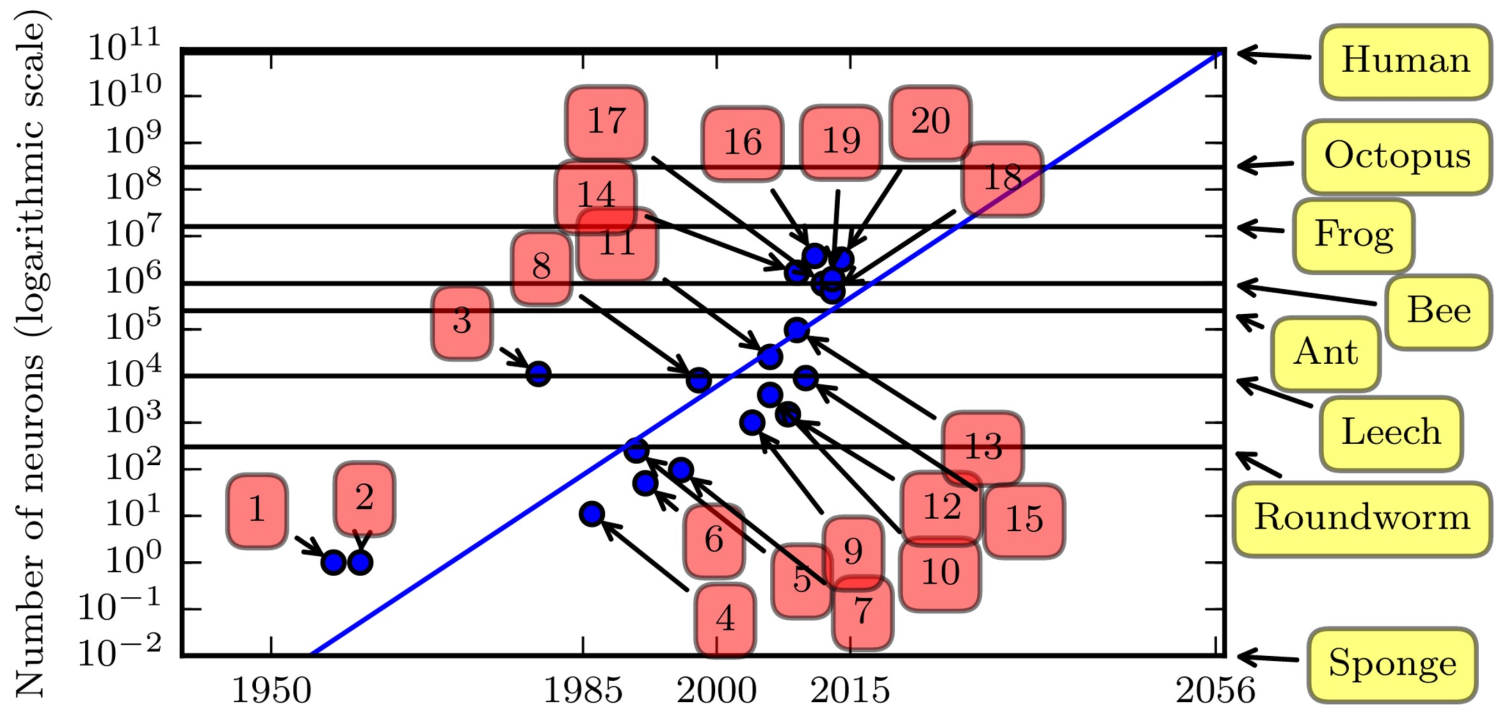

Figure 16. Increasing neural network size over time.

Why Deep Neural Networks?

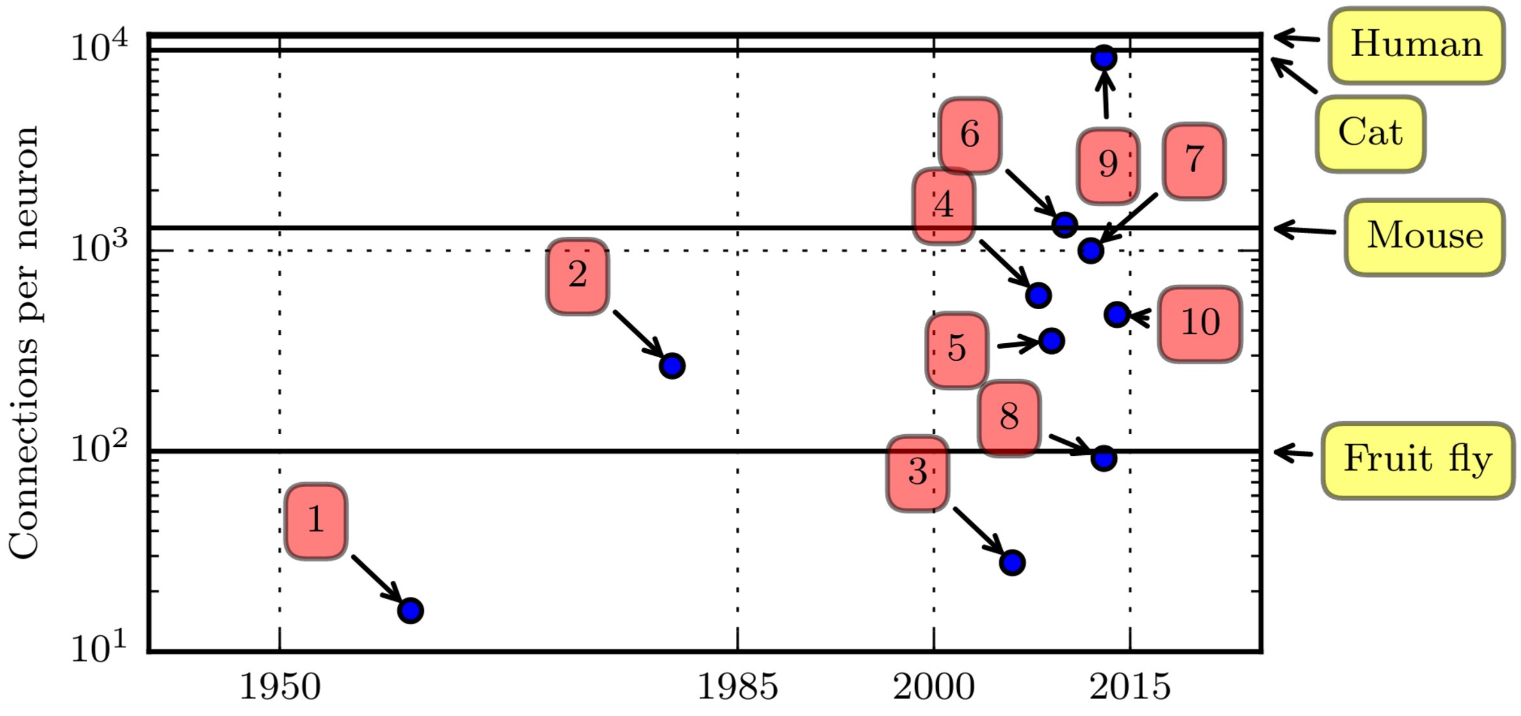

Figure 17. Number of connections per neuron over time.

You have reached the end of the lecture.

Image/Figure References

Figure 1. Perceptron (P) and Feed Forward (FF) network. Needs Source
Figure 2. Two of the three historical waves of artificial neural nets research. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 3. ImageNet Challege Logo. Retrieved from: www.image-net.org.
Figure 4. Decreasing error rate over time. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 5. Decreasing error rate over time. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 6. Decreasing error rate over time. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 7. Decreasing error rate over time. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 8. Deep Feed Forward (DFF) network. Need Source.
Figure 9. Illustration of a deep learning model. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 10. Cartesian Coordinates & Polar Coordinates. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 11. Cartesian Coordinates & Polar Coordinates. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 12. Increasing dataset size over time. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 13. Deep learning approaches can handle larger amounts data. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 14. Deep learning is a kind of representation learning. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 15. Flowcharts showing how the different parts of an AI system relate to each other within different AI disciplines. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 16. Increasing neural network size over time. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Figure 17. Number of connections per neuron over time. Retrieved from: Goodfellow, Bengio, and Courville, Deep Learning, MIT Press, 2016.
Other images were purchased from Getty Images and with permission to use.