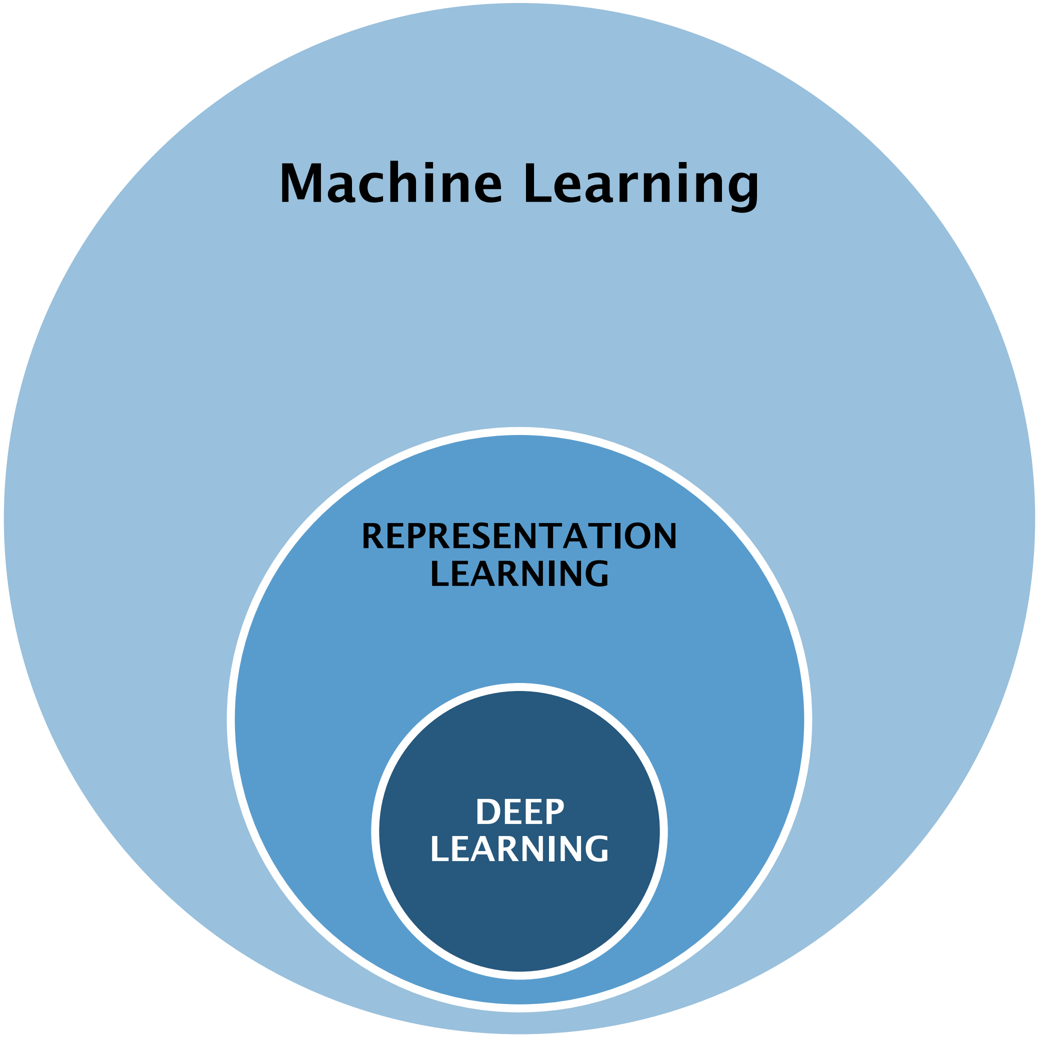
Deep Learning

Overview

- Today we'll focus on a specific kind of machine learning called deep learning


Machine Learning
REPRESENTATION LEARNING
DEEP LEARNING

[LBH15] Yann LeCun, Yoshua Bengio, Geoffrey Hinton (2015). Deep learning. *Nature*, *521*(7553), 436–444.

REVIEW

doi:10.1038/nature14539

Deep learning

Yann LeCun[1,2], Yoshua Bengio[3] & Geoffrey Hinton[4,5]

Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech.

Machine-learning technology powers many aspects of modern society: from web searches to content filtering on social networks to recommendations on e-commerce websites, and it is increasingly present in consumer products such as cameras and smartphones. Machine-learning systems are used to identify objects in images, transcribe speech into text, match news items, posts or products with users' interests, and select relevant results of search. Increasingly, these applications make use of a class of techniques called deep learning.

Conventional machine-learning techniques were limited in their ability to process natural data in their raw form. For decades, constructing a pattern-recognition or machine-learning system required careful engineering and considerable domain expertise to design a feature extractor that transformed the raw data (such as the pixel values of an image) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input.

Representation learning is a set of methods that allows a machine to be fed with raw data and to automatically discover the representations needed for detection or classification. Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. With the composition of enough such transformations, very complex functions can be learned. For classification tasks, higher layers of representation amplify aspects of the input that are important for discrimination and suppress irrelevant variations. An image, for example, comes in the form of an array of pixel values, and the learned features in the first layer of representation typically represent the presence or absence of edges at particular orientations and locations in the image. The second layer typically detects motifs by spotting particular arrangements of edges, regardless of small variations in the edge positions. The third layer may assemble motifs into larger combinations that correspond to parts of familiar objects, and subsequent layers would detect objects as combinations of these parts. The key aspect of deep learning is that these layers of features are not designed by human engineers: they are learned from data using a general-purpose learning procedure.

Deep learning is making major advances in solving problems that have resisted the best attempts of the artificial intelligence community for many years. It has turned out to be very good at discovering intricate structures in high-dimensional data and is therefore applicable to many domains of science, business and government. In addition to beating records in image recognition[1–4] and speech recognition[5–7], it has beaten other machine-learning techniques at predicting the activity of potential drug molecules[8], analysing particle accelerator data[9,10], reconstructing brain circuits[11], and predicting the effects of mutations in non-coding DNA on gene expression and disease[12,13]. Perhaps more surprisingly, deep learning has produced extremely promising results for various tasks in natural language understanding[14], particularly topic classification, sentiment analysis, question answering[15] and language translation[16,17].

We think that deep learning will have many more successes in the near future because it requires very little engineering by hand, so it can easily take advantage of increases in the amount of available computation and data. New learning algorithms and architectures that are currently being developed for deep neural networks will only accelerate this progress.

Supervised learning

The most common form of machine learning, deep or not, is supervised learning. Imagine that we want to build a system that can classify images as containing, say, a house, a car, a person or a pet. We first collect a large data set of images of houses, cars, people and pets, each labelled with its category. During training, the machine is shown an image and produces an output in the form of a vector of scores, one for each category. We want the desired category to have the highest score of all categories, but this is unlikely to happen before training. We compute an objective function that measures the error (or distance) between the output scores and the desired pattern of scores. The machine then modifies its internal adjustable parameters to reduce this error. These adjustable parameters, often called weights, are real numbers that can be seen as 'knobs' that define the input–output function of the machine. In a typical deep-learning system, there may be hundreds of millions of these adjustable weights, and hundreds of millions of labelled examples with which to train the machine.

To properly adjust the weight vector, the learning algorithm computes a gradient vector that, for each weight, indicates by what amount the error would increase or decrease if the weight were increased by a tiny amount. The weight vector is then adjusted in the opposite direction to the gradient vector.

The objective function, averaged over all the training examples, can

[1]Facebook AI Research, 770 Broadway, New York, New York 10003 USA. [2]New York University, 715 Broadway, New York, New York 10003, USA. [3]Department of Computer Science and Operations Research Université de Montréal, Pavillon André-Aisenstadt, PO Box 6128 Centre-Ville STN Montréal, Quebec H3C 3J7, Canada. [4]Google, 1600 Amphitheatre Parkway, Mountain View, California 94043, USA. [5]Department of Computer Science, University of Toronto, 6 King's College Road, Toronto, Ontario M5S 3G4, Canada.

Representation Learning

“a set of methods that allows a machine to be fed with raw data and to automatically discover the representations needed for detection or classification.” [LBH15]

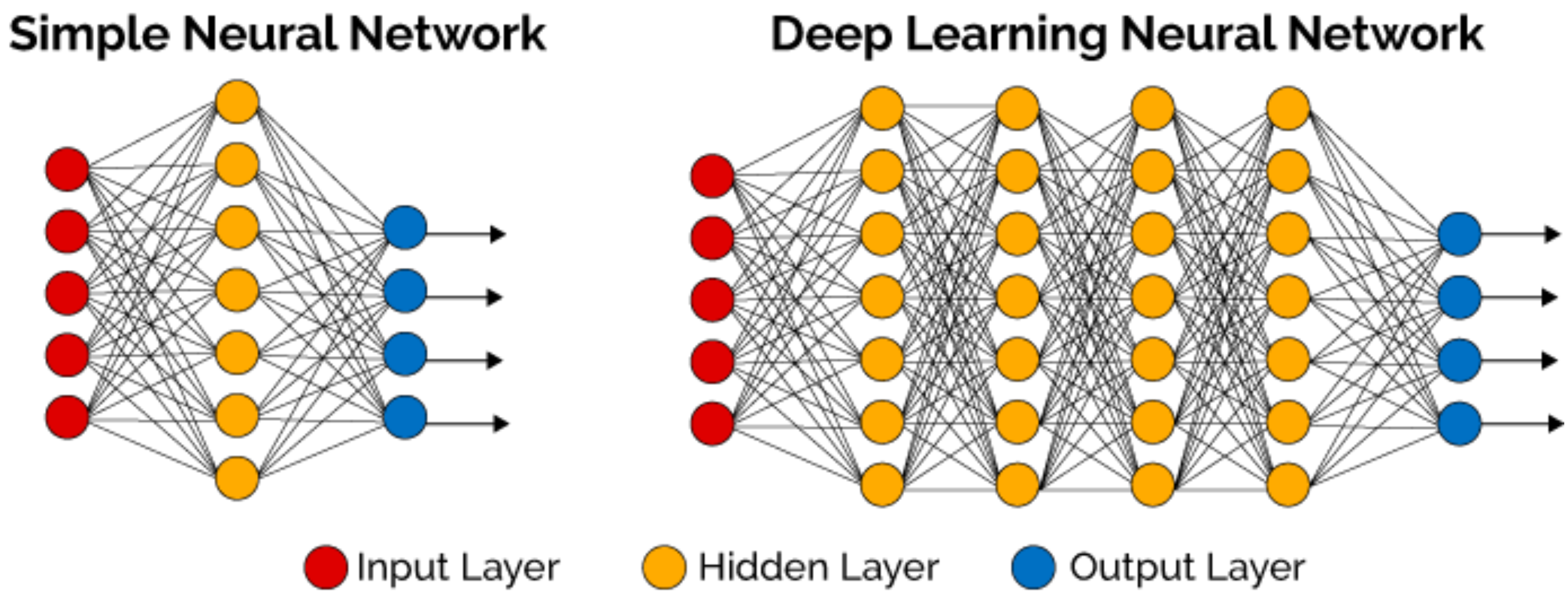
[LBH15] Yann LeCun, Yoshua Bengio, Geoffrey Hinton (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Deep Learning

"… are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. With the composition of enough such transformations, very complex functions can be learned. For classification tasks, higher layers of representation amplify aspects of the input that are important for discrimination and suppress irrelevant variations." [LBH15]

[LBH15] Yann LeCun, Yoshua Bengio, Geoffrey Hinton (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Deep Learning



https://hackernoon.com/log-analytics-with-deep-learning-and-machine-learning-20a1891ff70e

Implementations of Deep Learning

- Multiple levels of representation learning can be implemented in various ways including:
  - Deep neural networks
    - Deep convolution neural networks (ConvNets)
    - Recurrent neural networks (RNNs)
  - Deep believe networks

Deep Learning Platforms

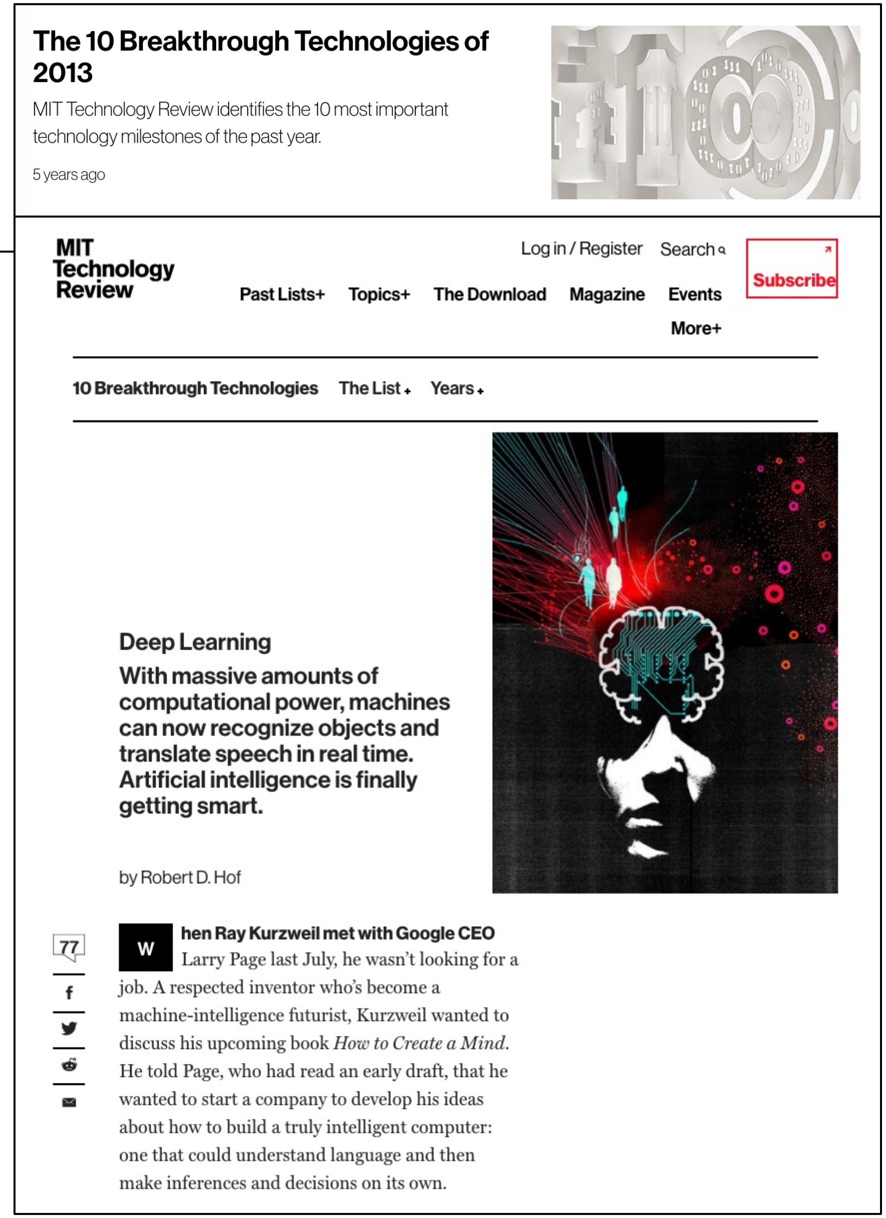
- **Example: TensorFlow**
  - TensorFlow is an open-source machine learning library developed by the Google Brain Team
  - For CSCI 6100G course projects that involve deep learning I recommend you use TensorFlow
  - *Tutorial:* https://www.tensorflow.org/tutorials/wide_and_deep

TensorFlow

The Impact of Deep Learning

- Deep learning was identified as one of the Top 10 Breakthrough Technologies by the MIT Tech Review, 2013
- Since than its application has proliferated many areas of IT technology

https://www.technologyreview.com/s/513696/deep-learning/



The 10 Breakthrough Technologies of 2013

MIT Technology Review identifies the 10 most important technology milestones of the past year.

5 years ago

MIT Technology Review

Log in / Register Search Subscribe

Past Lists+ Topics+ The Download Magazine Events More+

10 Breakthrough Technologies The List+ Years+

Deep Learning

With massive amounts of computational power, machines can now recognize objects and translate speech in real time. Artificial intelligence is finally getting smart.

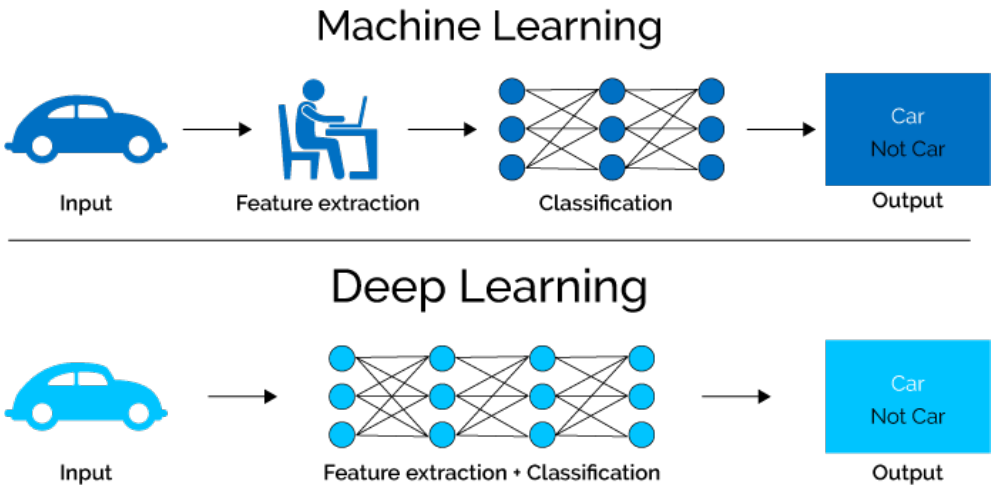
by Robert D. Hof

77

When Ray Kurzweil met with Google CEO Larry Page last July, he wasn't looking for a job. A respected inventor who's become a machine-intelligence futurist, Kurzweil wanted to discuss his upcoming book *How to Create a Mind*. He told Page, who had read an early draft, that he wanted to start a company to develop his ideas about how to build a truly intelligent computer: one that could understand language and then make inferences and decisions on its own.

The Benefit of Deep Learning

- Not constrained by traditional machine learning's limitations with respect to processing raw data
  - Processing raw data requires expertise and domain knowledge



https://hackernoon.com/log-analytics-with-deep-learning-and-machine-learning-20a1891ff70e

Criticisms of Deep Learning

- *"deep learning and AI in general ignore too much of the brain's biology in favor of brute-force computing."* [MIT-DL]
- *"Google's attitude is: lots of data makes up for everything"* – viewpoint of Jeff Hawkins, founder of Palm Computing, on Google's approach to deep learning [MIT-DL]
- Concerns about bias and comprehension of deep learning algorithms also increase as the algorithms get more complex and the data gets bigger

[MIT-DL] https://www.technologyreview.com/s/513696/deep-learning/

Deep Learning

Summary

- Today we introduced a type of representation learning known as deep learning
- We learned that deep learning can be implemented with neural networks (discussed the TensorFlow platform)
- We also discussed the benefits, potential and criticisms of deep learning