Deep Learning

jen (jen@redshiftzero.com)

Southside Hackerspace: Chicago

April 20, 2014

Algorithms 0000000 Theano 00

What is Deep Learning?

- A relatively new area of machine learning, introduced to bring ML closer to its original goal of Artificial Intelligence (AI)
- An algorithm is "deep" if it has more than one stage of non-linear feature transformation. Deep neural networks have more than one hidden layer.
- **The advantage:** deep models can learn far more complex functions than a shallow one can

Algorithms 0000000 Theano 00

What is Deep Learning?

- A relatively new area of machine learning, introduced to bring ML closer to its original goal of Artificial Intelligence (AI)
- An algorithm is "deep" if it has more than one stage of non-linear feature transformation. Deep neural networks have more than one hidden layer.
- **The advantage:** deep models can learn far more complex functions than a shallow one can

Algorithms 0000000 Theano 00

What is Deep Learning?

- A relatively new area of machine learning, introduced to bring ML closer to its original goal of Artificial Intelligence (AI)
- An algorithm is "deep" if it has more than one stage of non-linear feature transformation. Deep neural networks have more than one hidden layer.
- **The advantage:** deep models can learn far more complex functions than a shallow one can

Difficulty associated with training deep architectures

- Availability of data: For many problems there may be insufficient training examples to fit a very complex model
- Computational speed: It's only recently that algorithmic advances and increase in computing speed (or using GPUs, etc.) have allowed large models to be fit

Difficulty associated with training deep architectures

- Availability of data: For many problems there may be insufficient training examples to fit a very complex model
- Computational speed: It's only recently that algorithmic advances and increase in computing speed (or using GPUs, etc.) have allowed large models to be fit

Algorithms

Theano 00

Human Intelligence

- Deep Learning is one approach to understanding human cognition. How do brains learn? What algorithms do we use?
- The "one learning hypothesis": brains learn using only one type of learning algorithm
 - Some supporting evidence: "neural re-wiring experiments" rerouting visual inputs to the auditory cortex or the somatosensory cortex the brain learns how to see in a different area of the brain
- One of the big motivations for this sort of work

Algorithms 0000000 Theano 00

Human Intelligence

- Deep Learning is one approach to understanding human cognition. How do brains learn? What algorithms do we use?
- The "one learning hypothesis": brains learn using only one type of learning algorithm
 - Some supporting evidence: "neural re-wiring experiments" rerouting visual inputs to the auditory cortex or the somatosensory cortex the brain learns how to see in a different area of the brain
- One of the big motivations for this sort of work

Algorithms 0000000

Human Intelligence

- Deep Learning is one approach to understanding human cognition. How do brains learn? What algorithms do we use?
- The "one learning hypothesis": brains learn using only one type of learning algorithm
 - Some supporting evidence: "neural re-wiring experiments" rerouting visual inputs to the auditory cortex or the somatosensory cortex the brain learns how to see in a different area of the brain
- One of the big motivations for this sort of work

Algorithms

Theand 00

Example



Feature visualization from Zeiler and Fergus 2013

Algorithms •000000 Theano 00

Neural Network Recap



The output of each stage is f(Wx) where W is a set of weights and x is the input layer. Each circle is representative of a neuron "active" if output is near 1, "inactive" is output is near 0.

Theanc 00

Activation function f(z)



f(z) is the activation function, often chosen to be the sigmoid function: $f(z) = \frac{1}{1 + \exp(-z)}$. Another common choice for f(z) is tanh(z)

Algorithms

Theano 00

Autoencoder neural networks



It tries to learn a function $h_{W,b}(x) \approx x$. Allows the auto encoder network to pick out distinguishing features of the data.

Algorithms 0000000 Theano 00

Autoencoder Example



Training an autoencoder with 100 hidden units on 10x10 pixel images. Each neuron picks out an edge feature

Greedy layer-wise training

- Typically, the layers of the deep network are trained one at a time, and typically only the last layer is supervised (the weights from training the layers individually are used to initialize the weights in the final deep network).
- The advantage of this method is the fact that by using unsupervised data, we can typically use a much larger dataset

Greedy layer-wise training

- Typically, the layers of the deep network are trained one at a time, and typically only the last layer is supervised (the weights from training the layers individually are used to initialize the weights in the final deep network).
- The advantage of this method is the fact that by using unsupervised data, we can typically use a much larger dataset

Convolutional Neural Networks (CNNs)

- Inspired from biology, each network is sensitive to a small sub-region of the input space, and tiled to cover the visual field.
- Sparsely connected:



• Widely used for vision-based problems (including the Kaggle Galaxy Zoo winning solution)

Other common algorithms used in deep learning

- Restricted Boltzmann Machines (RBMs): Very similar to neural networks but with different functions between layers
- Deep Belief Networks (DBNs): Uses RBMs at each layer, stacking them to create a more complex model

Other common algorithms used in deep learning

- Restricted Boltzmann Machines (RBMs): Very similar to neural networks but with different functions between layers
- Deep Belief Networks (DBNs): Uses RBMs at each layer, stacking them to create a more complex model

N	lo	ti	va	ti	on
	0	0	0		

- Theano is a python library to make writing deep learning models "easy" by focusing on making them fast
- It also allows you to transparently run your models on GPUs if you like, and to boost speed, Theano will compile parts of your code directly into CPU or GPU instructions
- In some cases, Theano will recognize numerically unstable expressions and fix them
- Used by software packages such as pylearn2 and theano-nets, and based on numpy, scipy
- Go here if you want to learn how to use it: deeplearning.net/software/theano/tutorial

- Theano is a python library to make writing deep learning models "easy" by focusing on making them fast
- It also allows you to transparently run your models on GPUs if you like, and to boost speed, Theano will compile parts of your code directly into CPU or GPU instructions
- In some cases, Theano will recognize numerically unstable expressions and fix them
- Used by software packages such as pylearn2 and theano-nets, and based on numpy, scipy
- Go here if you want to learn how to use it: deeplearning.net/software/theano/tutorial

N	lo	ti	va	ti	on
	0	0	0		

- Theano is a python library to make writing deep learning models "easy" by focusing on making them fast
- It also allows you to transparently run your models on GPUs if you like, and to boost speed, Theano will compile parts of your code directly into CPU or GPU instructions
- In some cases, Theano will recognize numerically unstable expressions and fix them
- Used by software packages such as pylearn2 and theano-nets, and based on numpy, scipy
- Go here if you want to learn how to use it: deeplearning.net/software/theano/tutorial

N	lo	ti	va	ti	on
	0	0	0		

- Theano is a python library to make writing deep learning models "easy" by focusing on making them fast
- It also allows you to transparently run your models on GPUs if you like, and to boost speed, Theano will compile parts of your code directly into CPU or GPU instructions
- In some cases, Theano will recognize numerically unstable expressions and fix them
- Used by software packages such as pylearn2 and theano-nets, and based on numpy, scipy
- Go here if you want to learn how to use it: deeplearning.net/software/theano/tutorial

N	lo	ti	va	ti	on
	0	0	0		

- Theano is a python library to make writing deep learning models "easy" by focusing on making them fast
- It also allows you to transparently run your models on GPUs if you like, and to boost speed, Theano will compile parts of your code directly into CPU or GPU instructions
- In some cases, Theano will recognize numerically unstable expressions and fix them
- Used by software packages such as pylearn2 and theano-nets, and based on numpy, scipy
- Go here if you want to learn how to use it: deeplearning.net/software/theano/tutorial

Algorithms

Theano ○●

More Resources

deeplearning.stanford.edu

deeplearning.net