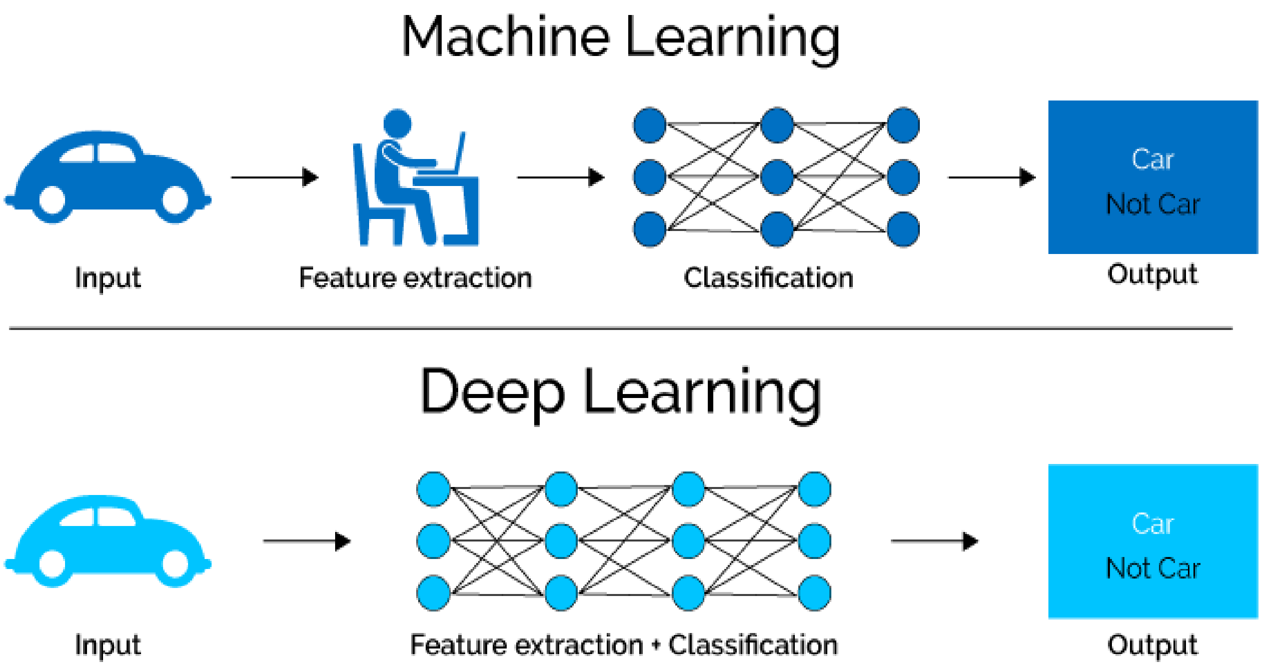
Deep Learning:

Table of Contents

Deep Learning:

Introduction:

Deep learning is a part of broader family of machine learning based on a set of techniques that allows a system to automatically discover the representations needed for feature detection or classification from raw data. It replaces the process of using domain knowledge of the data to create features that make machine learning algorithms work. It takes metadata as an input and process the data through a number of layers of the non-linear transformation of the input data to compute the output. In Deep Learning Neural Network, each hidden layer is responsible for training the unique set of features based on the output of the previous layer. As the number of hidden layers increases, the complexity and abstraction of data also increase.



Deep learning algorithms have shown superior learning and classification performance in areas such as transfer learning, handwritten character recognition, speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning is about automatically learning multiple levels of representations of the underlying distribution of the data to be modeled.

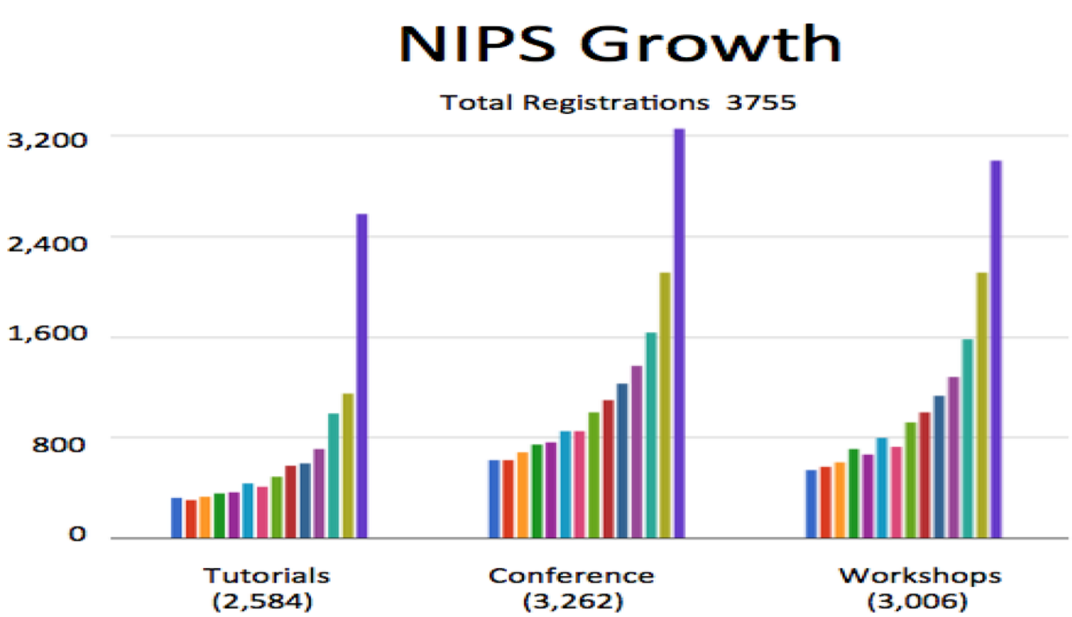
Deep Learning:

The Deep Learning Revolution:

The Big Bang of Artificial Intelligence

Neural Information Processing Systems:

The interest in deep learning increases a lot during the recent years. A good surrogate for interest in deep learning is attendance at the Annual Conference on Neural Information Processing Systems (NIPS). Interest in the conference has surged in the last 5 years:



Deep Learning:

NIPS is the main conference for deep learning research and have historically been where a lot of the new methodological research gets published. The chart only goes to 2015 and in the year 2016 NIPS had over 6,000 attendees.

Background:

Modern interest in deep learning began in 2006, with papers explaining how to train a type of neural network known as a deep belief network (DBN).Hinton declared that he knew how the brain works, and introduced the idea of unsupervised pre-training and deep belief nets. The idea was to train a simple 2-layer unsupervised model like a restricted Boltzmann machine, freeze all the parameters, stick on a new layer on top and train just the parameters for the new layer. You would keep adding and training layers in this greedy fashion until you had a deep network, and then use the result of this process to initialize the parameters of a traditional neural network. Using this strategy, people were able to train networks that were deeper than previous attempts, prompting a rebranding of ‘neural networks’ to ‘deep learning.

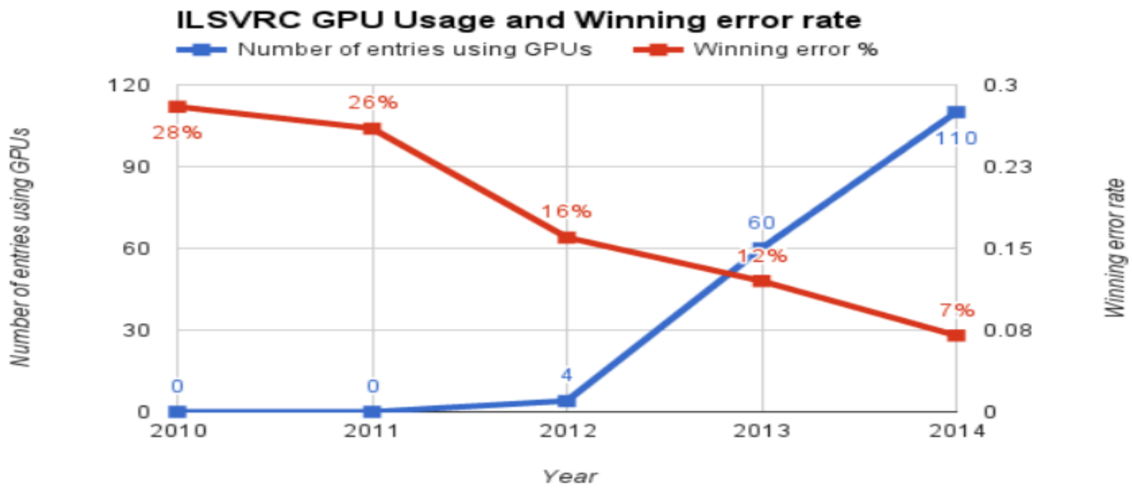
The Breakthrough (2012)

In 2010, a large database known as ImageNet containing millions of labeled images was created and published by Fei-Fei Li’s group at Stanford. This database was coupled with the annual LSVRC, where contestants would build computer vision models, submit their predictions, and receive a score based on how accurate they were. In the first two years of the contest, the top models had error rates of 28% and 26%. In 2012, Alex Krizhevsky, Ilya Sutskever, and Geoff Hinton entered a submission that would halve the existing error rate to 16%. The model combined several critical components that would go one to become mainstays in deep learning models.

Deep learning and the rise of the GPU:

Probably the most important piece was the use of graphics processing units (GPUs) to train the model. GPUs are essentially parallel floating-point calculators with 100s-1000s of cores. The speedup offered by GPUs meant they could train larger models, which led to lower error rates. They also introduced a method to reduce overfitting known as dropout and used the rectified linear activation unit (ReLU), both of which are now bread and butter components in modern deep learning.

Deep Learning:



Deep learning architecture:

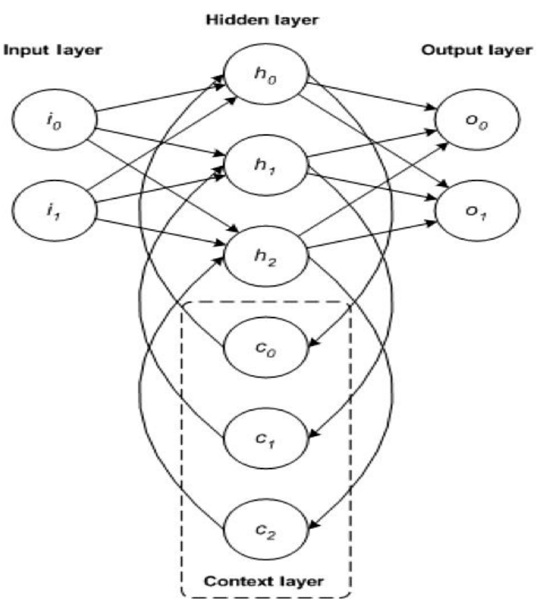
The number of architectures and algorithms that are used in deep learning are many. But here we will discuss five major architectures.

Architectures	Applications
RNN	Speech recognition, handwriting recognition
LSTM/GRU networks	Natural language text compression, handwriting recognition, speech recognition, gesture recognition, image captioning
CNN	Image recognition, video analysis, natural language processing
DBN	Image recognition, information retrieval, natural language understanding, failure prediction
DSN	Information retrieval, continuous speech recognition

Deep Learning:

Recurrent Neural Networks:

The RNN is one of the foundational network architectures from which other deep learning architectures are built. The primary difference between a typical multilayer network and a recurrent network is that rather than completely feed-forward connections, a recurrent network might have connections that feed back into prior layers or into the same layer.



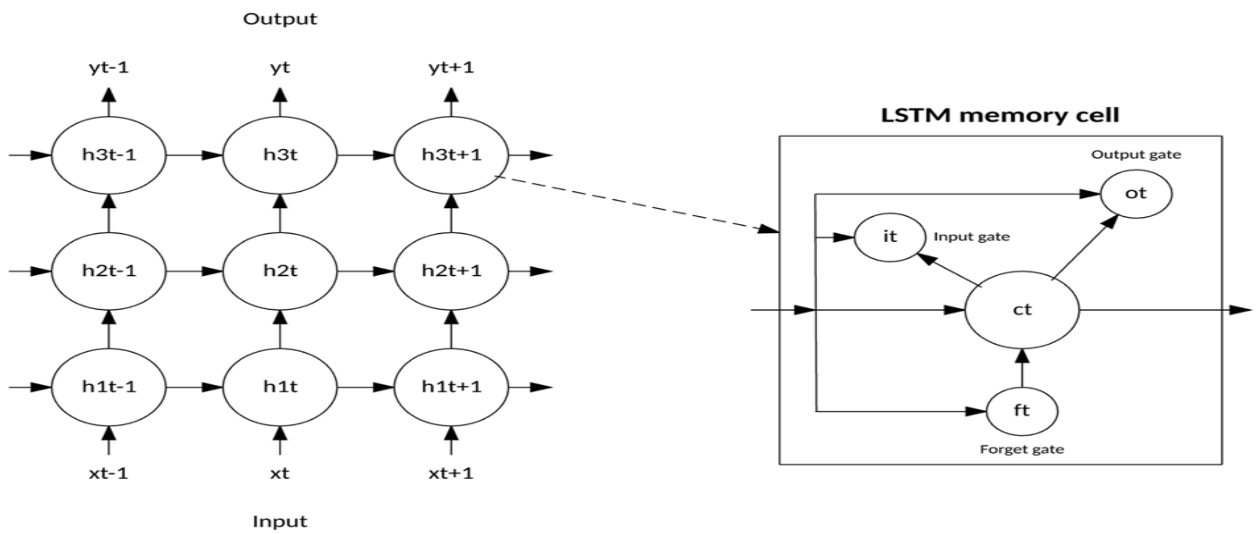
This feedback allows RNNs to maintain memory of past inputs and model problems in time. RNNs consist of a rich set of architectures. The key differentiator is feedback within the network, which could manifest itself from a hidden layer, the output layer, or some combination of both.

LSTM:

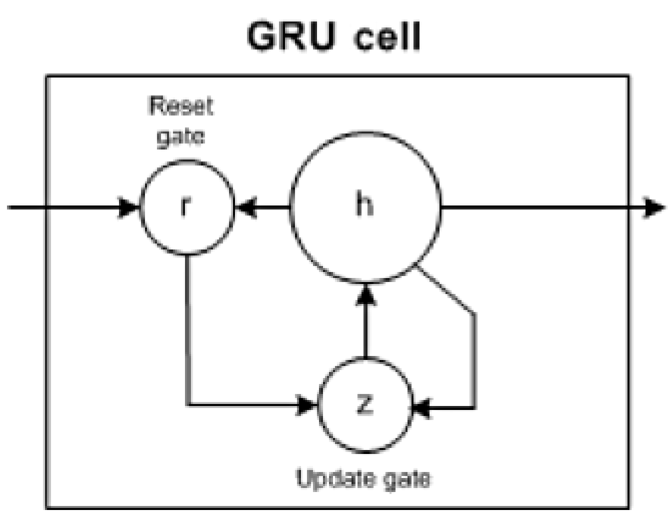
Long short-term memory (LSTM) block or network is a simple recurrent neural network which can be used as a building component or block (of hidden layers) for an eventually bigger recurrent neural network. The LSTM block is itself a recurrent network because it contains recurrent connections similar to connections in a conventional recurrent neural network.

Deep Learning:

The expression long short-term refers to the fact that LSTM is a model for the short-term memory which can last for a long period of time. An LSTM is well-suited to classify process and predict time series given time lags of unknown size and duration between important events.



A simplification of the LSTM was introduced called the **gated recurrent unit** (GRU).This model has two gates, getting rid of the output gate present in the LSTM model. For many applications, the GRU has performance similar to the LSTM, but being simpler means fewer weights and faster execution.



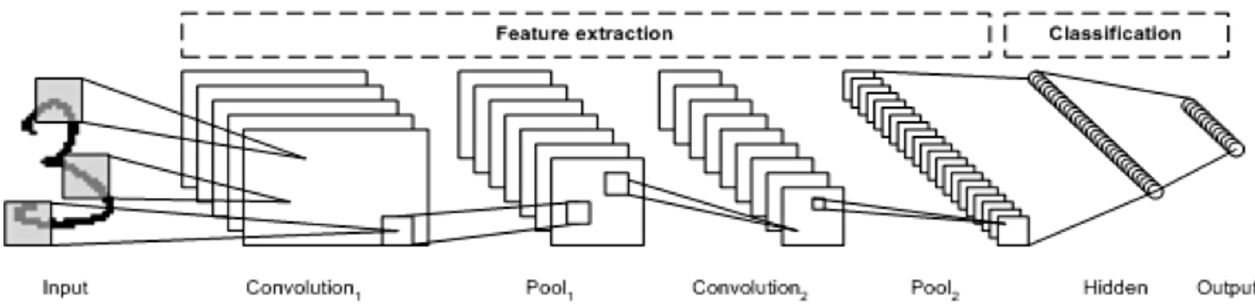
Deep Learning:

The GRU includes two gates: an update gate and a reset gate. The update gate indicates how much of the previous cell contents to maintain. The reset gate defines how to incorporate the new input with the previous cell contents. A GRU can model a standard RNN simply by setting the reset gate to 1 and the update gate to 0.

The GRU is simpler than the LSTM, can be trained more quickly, and can be more efficient in its execution. However, the LSTM can be more expressive and with more data, can lead to better results.

Convolutional neural networks:

A CNN is a multilayer neural network that was biologically inspired by the animal visual cortex. The architecture is particularly useful in image-processing applications. The first CNN was created by Yann Lacuna; at the time, the architecture focused on handwritten character recognition, such as postal code interpretation.



Apart from image processing, the CNN has also successfully applied to video recognition and natural language processing.

Deep Belief Networks:

In order to overcome the limitation of earlier neural networks, Professor Geoffrey Hinton introduces Deep Belief Networks. Deep Belief Networks (DBN) consists of two different types of neural networks – Belief Networks and Restricted Boltzmann Machines.

In contrast to perceptron and backpropagation neural networks, DBN is unsupervised learning algorithm. A deep belief network (DBN) is a class of deep

Deep Learning:

neural network, composed of multiple layers of variables that are not directly observed but are rather inferred (through a mathematical model) from other variables that are observed (directly measured),with connections between the layers but not between units within each layer.

Belief nets:

A belief net is a directed acyclic graph composed of stochastic variables.
We get to observe some of the variables and we would like to solve two problems.

- ***The inference problem:***

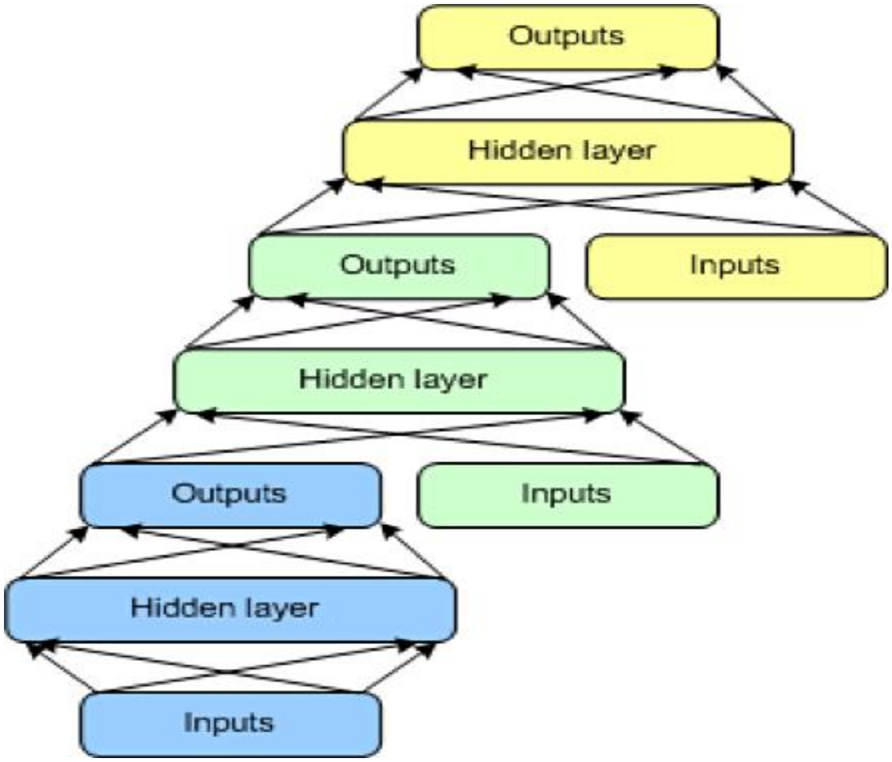
  Infer the states of the unobserved variables.
- ***The learning problem:***

  Adjust the interactions between variables to make the network more likely to generate the observed data.

Deep Stacking Networks:

The final architecture is the DSN, also called a deep convex network. A DSN is different from traditional deep learning frameworks in that although it consists of a deep network, it's actually a deep set of individual networks, each with its own hidden layers. This architecture is a response to one of the problems with deep learning: the complexity of training. Each layer in a deep learning architecture exponentially increases the complexity of training, so the DSN views training not as a single problem but as a set of individual training problems.

Deep Learning:



The DSN consists of a set of modules, each of which is a subnetwork in the overall hierarchy of the DSN. In one instance of this architecture, three modules are created for the DSN. Each module consists of an input layer, a single hidden layer, and an output layer. Modules are stacked one on top of another, where the inputs of a module consist of the prior layer outputs and the original input vector. This layering allows the overall network to learn more complex classification than would be possible given a single module.

The DSN permits training of individual modules in isolation, making it efficient given the ability to train in parallel. Supervised training is implemented as back-propagation for each module rather than back-propagation over the entire network. For many problems, DSNs can perform better than typical DBNs, making them popular and efficient network architecture.

Deep Learning:

Deep Learning Frameworks

Tensorflow

One of the most popular Deep Learning libraries out there, Tensorflow, was developed by the Google Brain team and open-sourced in 2015. Positioned as a ‘second-generation machine learning system’, Tensorflow is a Python-based library capable of running on multiple CPUs and GPUs. It is available on all platforms, desktop, and mobile. It also has support for other languages such as C++ and R and can be used directly to create deep learning models, or by using wrapper libraries (for e.g. Keras) on top of it.

Keras

Although TensorFlow is a very good deep learning library, creating models using only Tensorflow can be a challenge, as it is a pretty low-level library and can be quite complex to use for a beginner. To tackle this challenge, Keras was built as a simplified interface for building efficient neural networks in just a few lines of code and it can be configured to work on top of TensorFlow. Written in Python, Keras is very lightweight, easy to use, and pretty straightforward to learn. Because of these reasons, Tensorflow has incorporated Keras as part of its core API.

Caffe

Built with expression, speed, and modularity in mind, Caffe is one of the first deep learning libraries developed mainly by Berkeley Vision and Learning Center (BVLC). It is a C++ library which also has a Python interface and finds its primary application in modeling Convolutional Neural Networks. One of the major benefits of using this library is that you can get a number of pre-trained networks directly from the Caffe Model Zoo, available for immediate use.

Torch

Torch is a Lua-based deep learning framework and has been used and developed by big players such as Facebook, Twitter and Google. It makes use of the C/C++ libraries as well as CUDA for GPU processing. Torch was built with an aim to achieve maximum flexibility and make the process of building your models extremely simple. More recently, the Python implementation of Torch, called PyTorch, has found popularity and is gaining rapid adoption.

Future of Deep Learning:

Unsupervised learning had a catalytic effect in reviving interest in deep learning, but has since been overshadowed by the successes of purely supervised learning. Although we have not focused on it in this Review, we expect unsupervised learning to become far more important in the longer term. Human and animal learning is largely unsupervised: we discover the structure of the world by observing it, not by being told the name of every object.

Human vision is an active process that sequentially samples the optic array in an intelligent, task-specific way using a small, high-resolution fovea with a large, low-resolution surround. We expect much of the future progress in vision to come from systems that are trained end-to-end and combine ConvNets with RNNs that use reinforcement learning to decide where to look. Systems combining deep learning and reinforcement learning are in their infancy, but they already outperform passive vision systems at classification tasks and produce impressive results in learning to play many different video games.

Natural language understanding is another area in which deep learning is poised to make a large impact over the next few years. We expect systems that use RNNs to understand sentences or whole documents will become much better when they learn strategies for selectively attending to one part at a time.

Ultimately, major progress in artificial intelligence will come about through systems that combine representation learning with complex reasoning. Although deep learning and simple reasoning have been used for speech and handwriting recognition for a long time, new paradigms are needed to replace rule-based manipulation of symbolic expressions by operations on large vectors.

How to stay current:

This field moves fast, and papers get old on a timescale of months, so keeping your finger on the pulse can be tough. I've developed a few coping mechanisms to try and help drink from the deep learning firehose:

- Read arXiv, and specifically the machine learning subsection. Close to 100% of deep learning papers get posted in some form on arXiv, so it will always be the definitive source for the cutting edge. The problem is that the volume of new work coming out is too large for any one person to consume. The main way to cope is a combination of social media like twitter and reddit and blogs from individuals and various research groups.
- arxiv-sanity is a machine learning (though not deep learning!) powered tool from Andrej Karpathy to help you shift through the vast swaths of research that is produced daily. Makes it easy to find papers relevant to your interests and even comes with a Netflix style recommendation engine.
- The keras blog has a lot of great tutorials on how to implement state of the art models in keras.
- Twitter is probably the best way to stay up to date, because everything happens in real time.

  - ✓ Yann Lecun - Head of Facebook's AI research lab.
  - ✓ Andrej Karpathy - Research scientist at OpenAI, also has a great blog.
  - ✓ Ian Goodfellow - Researcher at OpenAI, creator of generative adversarial networks.

Deep Learning:

- ✓ Francois Chollet - Creator of keras, researcher at Google.
- ✓ Russ Salakhutdinov - Director of AI Research at Apple and professor at CMU.
- ✓ Fei-Fei Li - Imagenet creator, Stanford professor and director of the Stanford AI lab, Chief Scientist at Google Cloud.
- ✓ Andrew Ng - Cofounder of coursera, former Standford professor, Chief scientist at Baidu,
- ✓ Sander Dieleman - Creator of the lasagne deep learning framework and research scientist at Deep Mind.
- ✓ Hugo Larochelle - Google Brain researcher.
- ✓ Rachel Thomas - Cofounder of fast.ai, a fantastic deep learning MOOC.

References

Deep Learning:

1. Dettmers, Tim. *Deep Learning in a Nutshell: History and Training.* 12 15, 2015. https://devblogs.nvidia.com/deep-learning-nutshell-history-training/.

2. Yann LeCun, Yoshua Bengio& Geoffrey Hinton. "Deep learning." 27 May 2015.

3. Hinton, G. E., Osindero, S. and Teh, Y. (2006) A fast learning algorithm for deep belief nets.

4. CY Lee, S Xie, P Gallagher, Z Zhang, Z Tu - Artificial Intelligence and Statistics, 2015 Deeply-supervised nets

5. Andrew, L.Beam. Deep Learning- 02 20 , 2017. History and Background https://beamandrew.github.io/deeplearning/2017/02/23/deep_learning_101_part1.html

6. Amey Varangaonkar. Deep Learning Frameworks May 25, 2017 https://datahub.packtpub.com/deep-learning/top-10-deep-learning-frameworks/