# **Optimization Algorithms**

## **A quick history of DL to set the context**

1. The following illustration shows the progress of Deep Learning over the last 3 decades
2. Some of the salient points in the DL-timeline are as follows
   1. **1989-1991**
      1. Universal Approximation Theorem: we will be able to approximate any kind of function with our Neural Network
      2. Backpropagation: Derivative calculation happens backwards from the output layer to the input, ie back propagation. It is nothing but Gradient Descent(**1847**) applied with the chain rule
   2. **1993-1994**
      1. A lot of work was done on Recurrent Neural Networks
   3. **1998**
      1. LSTMs (Long Short-Term Memory) were proposed
      2. Work done on Convolutional Neural Networks
   4. **2006**
      1. Revival of DL by Hinton et. al. with the proposal of Unsupervised Pre-training
      2. People’s interest in DL started increasing.
   5. **2019**
      1. Better Learning Algorithms, Initializations, Activation and Regularization
      2. More Data, compute and democratization.