Deepan Sanghavi

Deep Learning

- Yann LeCun, Yoshua Bengio & Geoffrey Hinton

Abstract

This paper concentrates on explaining the subject Deep Learning. It allows computational model to learn through multiple processing layers and representing data with multiple level of abstraction. This concept has applications in diverse field. 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. Initially paper starts with explaining the root of Deep Learning i.e. Machine Learning and further explaining the sub category which is Supervised Learning. Moreover, this paper explains other deep learning models such as Recurrent neural network, Backpropagation and Convolutional neural network.

Introduction:

Paper starts with explaining about what is machine learning and deep learning. Further, it concentrates on how deep learning can be useful and areas where it is used most. Deep learning specifically consist of nonlinear models. Paper concentrates on supervised learning and explained it with an image dataset example. Paper mentions how the scoring of the class is done and have mentioned about the cost function which tries to reduce the error. This cost functions adjust weights and parameters in such a way that you get lowest error rate based in training data. Basically, it gives definition about various deep learning algorithm terms such as gradient vector, stochastic gradient descent. Later they, compare linear vs nonlinear classifiers.

Backpropagation:

Aim of the research is to transfer hand-engineered features into trainable multilayer networks. The backpropagation procedure to compute the gradient of an objective function with respect to the weights of a multilayer stack of modules is nothing more than a practical application of the chain rule for derivatives. To go from one layer to the next, a set of units compute a weighted sum of their inputs from the previous layer and pass the result through a non-linear function. This section of paper explains how backpropagation developed through time.

Convolutional Neural Network:

one deep, feedforward network that was much easier to train and generalized much better than networks with full connectivity between adjacent layers. This was the convolutional neural network and it is widely accepted. ConvNets are designed to process data that come in the form of multiple arrays, for example a color image composed of three 2D arrays containing pixel intensities in the three color channels. This explains how CNN works and it series of stages explain each one of them. It has two type of layers: Convolutional Layer and Pooling layer. Units in a convolutional layer are organized in feature maps, within which each unit is connected to local patches in the feature maps of the previous layer

through a set of weights called a filter bank. The role of the pooling layer is to merge semantically similar features into one. Further, they have explained the application of CNN in Image classification problem.

Distributed representations and language processing:

Over here, it explains difference between classical and deep learning algorithms. The hidden layers of a multilayer neural network learn to represent the network's inputs in a way that makes it easy to predict the target outputs.

RNN's:

RNN's are used for sequential inputs such as speech and language. RNNs process an input sequence one element at a time, maintaining in their hidden units a 'state vector' that implicitly contains information about the history of all the past elements of the sequence. They mentioned the problem with RNN's which is know as vanishing gradient problem. RNNs are very powerful dynamic systems, but training them has proved to be problematic because the backpropagated gradients either grow or shrink at each time step, so over many time steps they typically explode or vanish.

Strengths:

- Paper is very well organized
- Used several examples to explain the concepts
- Good overview of each algorithm without going very deep into the concept
- Perfect paper to get an overview of deep learning

Drawbacks:

- Should have implemented a system for each of the methods and how results are different
- Explanation could have been of more detail for some section such as convolutional neural networks