

DIGIT RECOGNITION USING DEEP LEARNING

Just like we use our human brains to identify patterns and classify various types of information, deep learning algorithms can be taught to accomplish the same tasks for machines.

Example:

Automatic car driving system is a good example of deep learning. Deep Learning has many applications like smart vision, speech recognition and natural language processing

Deep learning exists from 1970 but it became popular in the last 5 years due to the following reasons.

- More availability of data and less cost of storage devices
- More computational power availability and more optimized algorithm which takes less power and utilizes less resources.

Because of this deep learning has created a boom in computer vision.

Convolutional Neural Networks are a part of Deep learning which are mostly used for computer vision problem. They are called Convolutional Neural networks because most of the layers in the network involves Convolution operation. The whole CNN started with a Perceptron model.

Perceptron model is inspired from human neuron. The idea of perceptron is that it outputs 1 if the equated sum of inputs and weights are greater than or equal to threshold. It outputs 0 if the equated sum is less than threshold. Perceptron model has some limitations. Example: XOR function is not linearly separable so we cannot use perceptron to predict truth table of XOR. The output of perceptron is always binary. So, researchers have taught that network of perceptron model solves these issues and they came up with an architecture called multi-layer perceptron architecture called neural network. If the neural network contains more of convolution operation, then it is called Convolutional Neural Networks.

CNN contains many layers. Some of the important layers are listed below.

- Convolution operation extracts the features of the image like vertical edges, horizontal edges etc. Pooling Operation reduces the size of the input while preserving the important features of the image.
- Pooling operation does down sampling. There are two types of pooling.
 - Maxpooling : Takes the max of given kernel size
 - AvgPooling: Takes the average of given kernel size.
- Fullyconnected layer: this layer is called fully connected since each input is connected to every neuron in hidden layer.

How to approach an Image classification problem?

- First, we need a dataset to train our model. We can prepare our dataset, or we have some opensource datasets. Some of the opensource datasets are ImageNet, MNIST, VisualQA etc
- Next we need to have a loss function and optimizer. Loss function calculates the loss of our model and optimizer optimizes this loss in such a way that our model learns its weights faster. Some loss functions: MSE, Catogorical loss entropySome Optimizors: SGD, Adam
- Then we need a visual algorithm like LENET, Googlenet etc for classifying the imagesAnd
- we needed test dataset to test our model.

- Then we need to choose any of the popular frameworks in Python for deep learning like caffe, caffe2, pytorch, tensorflow and keras etc to implement it.

Keras is very user friendly and good for deep learning beginners So, we will start our hello world (digit recognition using keras) program of deep learning using Keras.