RAISONI GROUP — a vision beyond — D2

Assignment NO-3

Title :-

convolutional Neural Network (CNN).

objective :-

- To implement different deep learning models.
- To illustrate the concepts of AI/ML.

Problem statement :-

- using MNIST Fashion dataset and create a classifier to classify fashion clothing into categories.

S/W & H/W Requirements :-

- 64-bit open source operating system or its derivative.
- programming language :- Python.

Theory :-

convolutional Neural Network :-

- convolutional neural network are a specialized type of artificial neural networks that use a mathematical operation called convolution. in place of general matrix multiplication in at least one of their layer

They are specially designed to process pixel data and are used in image recognition and processing.

- In MNIST-like classification problem multilayer perceptron (MLP) are mostly used and provide very good result, furthermore, the training time remains reasonable. However, for larger datasets with larger images the no. of parameter for the model increases exponentially, therefore, training becomes longer and performance poorer.

* The convolution algorithm :-

The convolution is a kind of product operation of a filter - also called a kernel - with a matrix of images to extract from it some pre-determined characteristics.

Literally - speaking we use a convolution filter to "filter" the image to and display only what really matter to us.

$$\begin{bmatrix} 2 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

* Deep Learning approach for convolution:-

In this classification problem, we have two categories namely dog & cat. classifying an image, in one of these categories, depends on singular characteristics such as the shape of the skull, shape of the ears, shapes of the eyes etc.

only these characteristics matter to perform this classification task, the other information is not important to us.

The ideal would thus be starting from an image, to be able to extract the main characteristics interesting for the classification problem by using appropriate filters. In a deep Learning context, it will be the model to determine these filters by training on the dataset.

RAISONI GROUP
a vision beyond

* Padding and edge effect :-

- lets take again the example of the convolution animated above and look at the dimensions of the output image, also called feature map. The input image is a matrix 6x6 px sized, the filter is a 3x3 matrix we can see that the feature map is 4x4 px sized. By the way generally - speaking, the size of the feature map is,

$$m = n - p + 1$$

If we want to output a feature map with same size as the input image, we have to add zeros around the input image before the convolution, the input image is 'padded with zeros' hence the name ~~is~~ of this operation padding.

* Activation function :-

during training, we know that the coefficients of filters are updated these can be

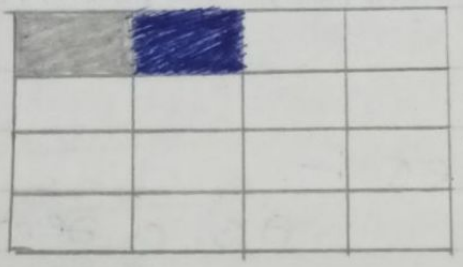
Rohit Patil

negative as we have seen above with the sobel filter, it then comes that the coefficient of the feature map can hold large negative value during training. since weknow these values represents pixel levels and therefore positive, we can apply a function to replace the negative values with zeros and keep the positive values as they are. This is an activation function called relu.

* convolution layer :-

As a reminder convolution is to apply a filter or kernal to an input image, we then get a feature map that highlights characteristics or features of the input image: outline, spots, shapes etc. Each filter has a simple and precise task to achieve. so, to solve our classification problem we will have to use several filters ; and by combining the features highlighted by those filters. such as shape of ears, eyes and contours, our model

0	1	0	1	1	0
1	1	1	0	0	1
0	1	0	1	1	0
0	0	1	1	0	1
0	1	0	1	1	0
1	0	1	1	1	0



Respective fields of neurons.

RAISONI GROUP
a vision beyond

will be able to get trained to distinguish a dog from a cat.

* Advanced approach of convolution:-

- In the previous description it is said that convolution is to multiply a sliding matrix with and input matrix image. Although this explanation is slightly different but not much more complex, lets ~~let~~ take the example of our first convolution with a 3x3 kernal and a 6x6 Px matrix. we use the formula above to predict the size of the feature map, 4x4 Px.

- The value of the pixel at position top left of the feature map directly depends both on the pixel values in the input image and the values of the convolution kernal. The value of this pixel is according to the convolution algorithm.

$$\sum_{i=1}^{9} x_i w_i$$

RAISONI GROUP
a vision beyond

where the wi are the coefficients of the convolution kernal and the xi, the coefficients of the matrix in the ~~gge~~ pencil box. Lets remind the convolution kernel,

2	2	0
0	2	0
0	1	1

Convolution kernel

conclusion :-

The purpose of this article was to introduce convolutional neural networks with their major interest. In general, these networks provide excellent results for classification and recognition tasks. If the problem to solve is to look for a pattern in a sequence then convolution networks will be good candidates.