Month 1

Introduction

Week 1: Introduction & Review

Introduction to CV

→ CV is a field of study focused on the problem of helping computers to see.

→ Multidisciplinary field broadly classified under AI and machine learning.

→ The goal of computer vision is to understand the content of digital images. Typically this involves a developing methods that attempt to reproduce the capability of human vision.

★ Computer vision and image processing

Image processing is process of creating new image from an existing image, typically simplifying or enhancing the content in some way. It is the type of digital signal processing and doesn't pay need to the info. of the image or the understanding the content of an image.

* <u>Tasks in Computer vision</u>

→ Object classification
→ Object identification
→ Object verification
→ Object detection
→ Object verification
→ Object landmark detection.
→ Object segmentation
→ Object recognition.

<u>Multilayer Perceptron</u>

we know that a perceptron is basically
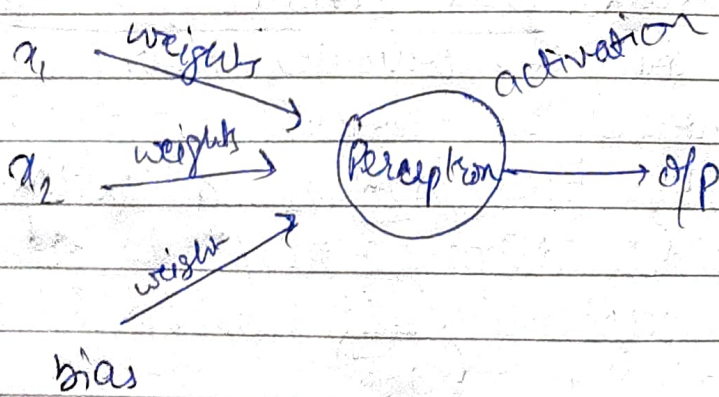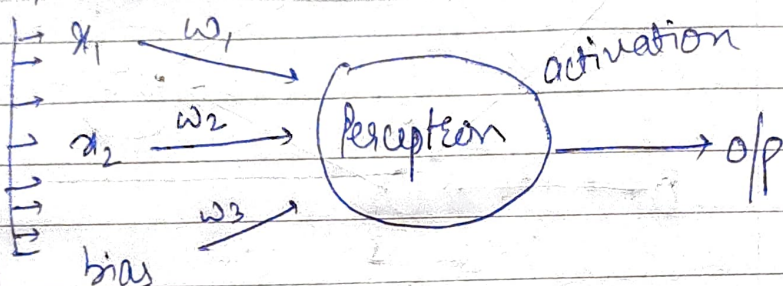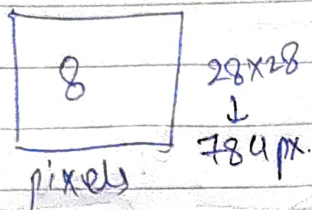

Inputs with weights added on goes in perceptron with bias ~~Goes the~~

Output of a perceptron goes a through the perceptron to create an o/p.

So why not stop only here? Why is this perceptron not good enough?

Suppose, we have a picture of number 6.

8 28×28
↓
784 px.

pixels

x_1 w_1

x_2 w_2 Perceptron activation → o/p

w_3

bias

Now, putting all the 784 values to give out ten probabilities by a single perceptron is quite a trivial job, as a perceptron is quite good for a linearly seperable problems.

The probability of recognizing the number is anything but linear.

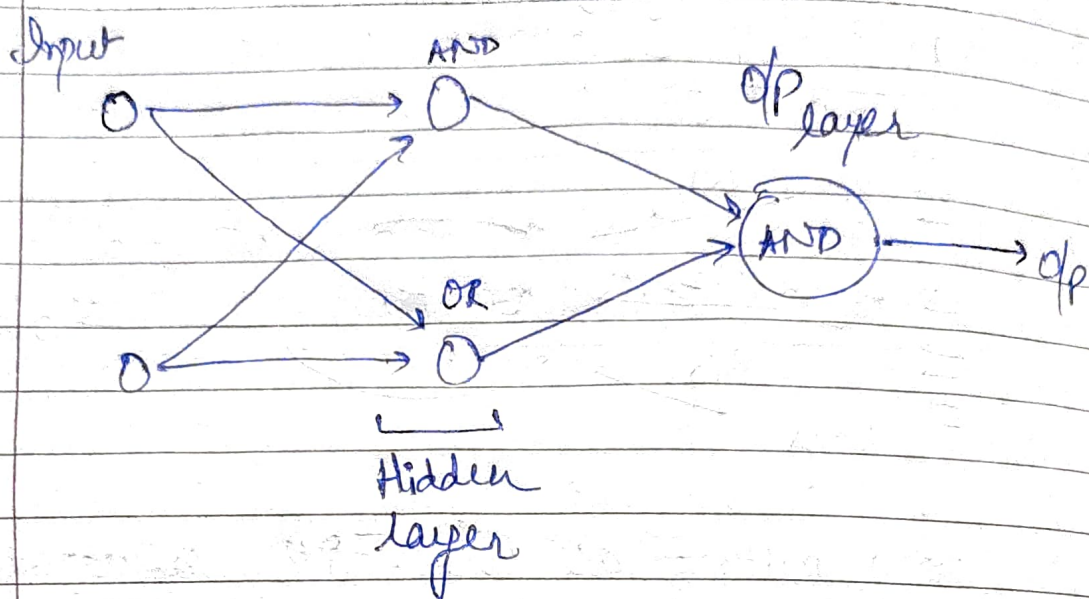For example: A XOR problem is not at all linear seperable problem.

A XOR B

A/B	T	F
T	F	T
F	T	F

⇓

This can't be divided by a single line

So, what we can do is, add 3 perceptrons with AND and OR to create a single unit

Hence, more complicated problems need multi-layered perceptrons to solve them.

Input

AND

O/P layer

AND → o/p

OR

Hidden layer

Convolutional Neural Network

Convolutional neural networks (CNNs) are type of Artificial Neural Network (ANN) which is broadly used for image analysis, but can be also used for different type data analysis.

CNNs can be seen as to have some special kind of ability to learn or recognize patterns in a data and makes sense from them.

So, what is the difference b/w a CNN and simple MLP?

CNNs can and has other layers in it. There is no problem until the dimensions are taken care of.

CNN's apply the "convolutional function" to the input in the convolutional layers.

Convolutional layers have filters which helps the detect the patterns in an image quite easily.

Pattern can be edges, corners, color changes, etc.

So, the CNN uses layers of these convolutional layers to detect objects, etc. Deeper the layer, more sophisticated is the recognition.

when initiating convolutional layer, it is necessary to instantiate the number of filters the layer would have.

A filter can be seen a matrix of set dimensions with some arbitary values.

eg:

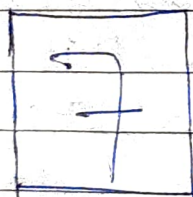| 0.25	0.17	0.9
0.89	0.36	0.63
0.7	0.24	0.82

3×3

The filter slides off over the pixels of the image.

The dot product of the pixels and the filter is stored and then slide to next 3×3 pxs and this goes on.

Hence, the dot products of the filter and pixels are the output of that convolutional layer and ip to another.

suppose the handwritten digit 7 from the MNIST dataset.

4 filters of 3×3 each

filter 1

-1	-1	-1
1	1	1
0	0	0

filter 2

-1	1	0
-1	1	0
-1	1	0

filter 3

0	0	0
1	1	1
-1	-1	-1

filter 4

0	1	-1
0	1	-1
0	1	-1

Now, in the filters, we can say that the rows with

−1 $\Rightarrow$ detects Black

0 $\Rightarrow$ detects Gray

1 $\Rightarrow$ detects white

These filters are very basic and used for edge detection. More sophisticat-ed filters are found later in the network.