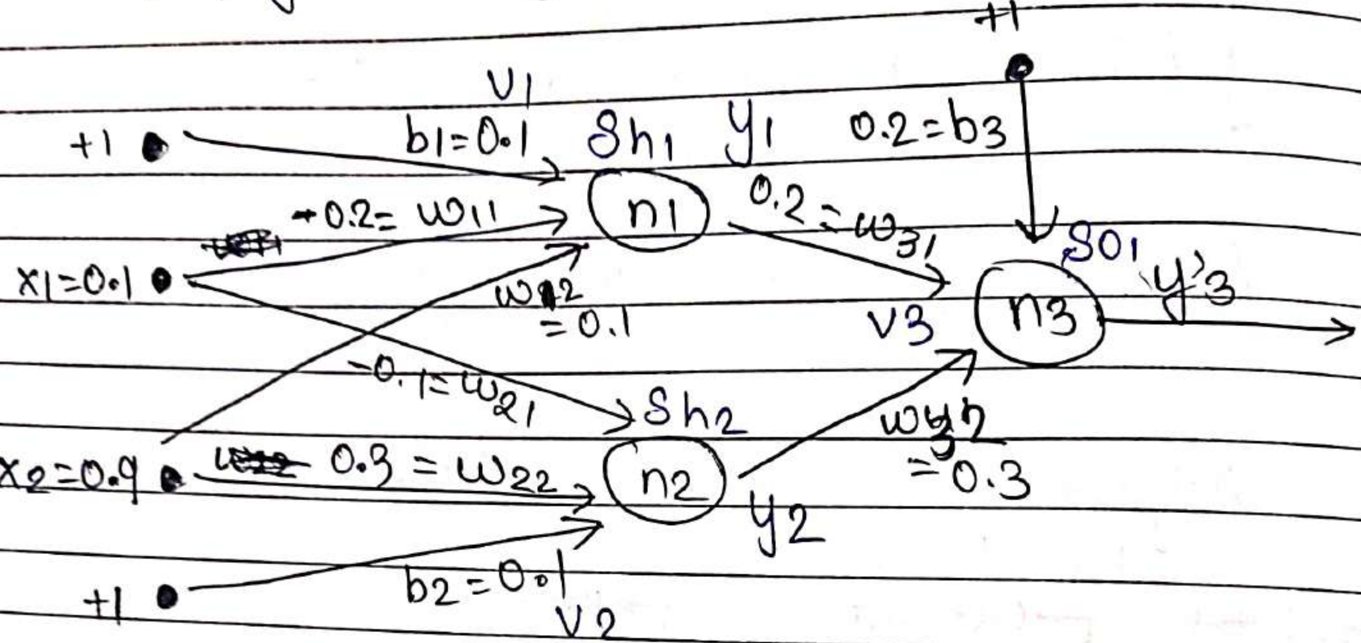
Example.

→ A complete forward & backward sweep of the feedforward network (2-2-1 architecture) using backpropagation algorithm.



Assume target o/p = 0.9 , $\eta = 0.25$ & $\alpha = 0.0001$

$d_3 = 0.9$

Sigmoid func^n $= g(z) = \dfrac{1}{1+e^{-z}}$

→ Forward Pass $\quad g'(z) = g(z)(1-g(z))$

$v_1 = b_1 + w_{11} \cdot x_1 + w_{12} \cdot x_2$

$= 0.1 + (-0.2) \cdot (0.1) + (0.1)(0.9)$

$= 0.17$

$y_1 = \dfrac{1}{1+e^{-0.17}} = g(0.17)$

$= 0.542$

$v_2 = b_2 + x_1 \cdot w_{21} + x_2 \cdot w_{22}$

$= 0.1 + (0.1)(-0.1) + (0.9)(0.3)$

$= 0.36.$

$$y_2 = \frac{1}{1+e^{-(-0.36)}} = 0.58$$

$$v_3 = b_3 + y_1 \cdot (w_{31}) + y_2 \cdot (w_{32})$$
$$= 0.2 + (0.542)(0.2) + (0.58)(0.3)$$
$$= 0.48$$

$$y_3 = \frac{1}{1+e^{-0.48}} = 0.61\cancel{8}$$

$\therefore$ Error, $e = d_3 - y'_3$
$$= 0.9 - 0.61 = 0.29$$

→ Backward Pass.

Step 1: Calculate local gradients

$$\delta o_1 = g'(v_3) * (d_3 - y_3)$$
$$= g(v_3)(1 - g(v_3)) * (0.29)$$
$$= 0.61 (1 - 0.61)(0.29)$$
$$= 0.06$$

$$\delta h_1 = g'(v_1) * \delta o_1 * w_{31}$$
$$= 0.542 (1 - 0.542)(0.06)(0.2) = 0.003$$

$$\delta h_2 = g'(v_2) * \delta o_1 * w_{32}$$
$$= 0.58 (1 - 0.58) * 0.06 * 0.3$$
$$= 0.004$$

Step 2: Adjust the weights of the network using the learning rule:

$$w(n+1) = w(n) + \alpha w(n-1) + \eta \delta(n) * y$$

→ Assume

$$w_{31}(n+1) = w_{31}(n) + w_{31}(n-1) * \alpha + \eta * \delta_{O1} * y_1$$
$$= 0.2 + (0.2) * 0.0001 + 0.25 * 0.08 * 0.542$$
$$= 0.20815 \approx 0.2090$$

(take all value 4 after decimal)

$$w_{32}(n+1) = w_{32}(n) + w_{32}(n-1) * \alpha + \eta\, \delta_{O1} * y_2$$
$$= 0.3 + 0.3 * (0.0001) + (0.25) * (0.06)(0.58)$$
$$= 0.30$$

$$w_{11}(n+1) = w_{11}(n) + \alpha \cdot w_{11}(n) + \eta \delta_{h_1} \cdot x_1$$
$$= -0.2 + (0.0001)(-0.2) + (0.05)(0.0033)(0.1)$$
$$= \cancel{-0.819}$$
$$= -0.1999$$

$$w_{21}(n+1) = (-0.1) + 0.001 * (-0.1) + (0.25)(0.0049)(1)$$
$$= -0.0999$$

Similarly

$$w_{12}(n+1) = 0.1008$$
$$w_{22}(n+1) = 0.3011$$

for

$$b(n+1) = b(n) + \alpha * b(n-1) + \eta * \delta(n) * 1$$
$$b_3 \approx 0.2166$$
$$b_1 = 0.1008 \qquad b_2 = 0.1012$$

Next

We can do forward pass using new weights & bias.

~~Face~~

Face Detection

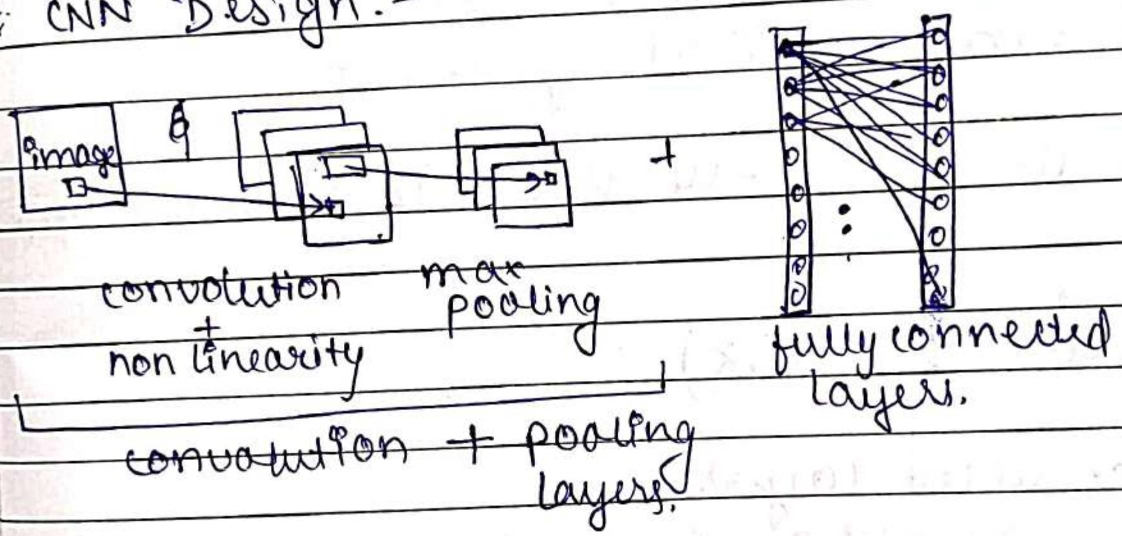
Methodology

i/p image → Pre-Processing → Feature extracⁿ with CNN → Softmax classifier → Result

3 steps mainly –

1) Resize image
~~Algorith~~ 2) Build a CNN model to extract feature
3) use Softmax classifier, for classificaⁿ.

CNN Design:-



Convolutional Layer-

- ~~to~~ The primary purpose of convolution layer is to extract features from the i/p data which is an image
- Convolution preserves the spatial relationship b/w pixels by learning image features.
- This produces a feature map or activation map in the o/p image & after that the feature maps are fed as i/p data to the next convolutional layer.

* Pooling Layer:

- Pooling layer reduces the dimensionality of each activation map but continues to have the most imp info.

- The i/p image are divided into a set of non-overlapping rectangles. Each region is down sampled by a non-linear operaⁿ such as average or maximum.

* RELU Layer — ~~Activation funcⁿ used to~~ make sure that model doesn't deviate much from original gradient

- It reconstitutes all -ve values in feature map by 0.

$$ReLU = \max(0, x)$$

* Fully connected layer:
- FCL is regarded as final pooling layer feeding the features to a classifier that uses softmax activation function for classificaⁿ.

- The goal of employing the FCL is to employ the features for classifying the i/p image into various classes based on the training dataset.

Given the following data, apply PCA and determine principal components (4,11), (8,4), (13,5), (7,14)

Soln

feature	sample 1	sample 2	sample 3	sample 4
x	4	8	13	7
y	11	4	5	14

step 1 No of features, $n = 2$

No of samples, $N = 4$

step 2 Find mean $\bar{x}$, $\bar{y}$

$$\bar{x} = \frac{(4+8+13+7)}{4} = 8$$

$$\bar{y} = \frac{(11+4+5+14)}{4} = 8.5$$

step 3 Compute covariance matrix

The ordered pairs are $(x,x), (x,y), (y,x), (y,y)$

i) Covariance of all ordered pairs

$$\operatorname{cov}(x_i, x_j) = \frac{1}{N-1} \sum_{k=1}^{N} (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) = \frac{1}{(N-1)} \sum_{k=1}^{N} (x_i - \bar{x})^2$$

(here $i=j$)

$$= \frac{1}{(4-1)} \left[(4-8)^2 + (8-8)^2 + (13-8)^2 + (7-8)^2 \right]$$

$$\boxed{\operatorname{cov}(x,x) = 14}$$

$$cov(x,y) = \frac{1}{N-1}\sum_{k=1}^{N}(x_i-\bar{x}_i)(y_i-\bar{y}_i)$$

$$cov(x,y) = \frac{1}{4-1}\left[(4-8)(11-8.5)+(8-8)(4-8.5)+(13-8)(5-8.5) + (7-8)(14-8.5)\right]$$

$$\boxed{cov(x,y) = -11}$$

$$\boxed{cov(y,x) = cov(x,y) = -11}$$

$$cov(y,y) = \frac{1}{4-1}\left[(11-8.5)^2+(4-8.5)^2+(5-8.5)^2+(14-8.5)^2\right]$$

$$\boxed{cov(y,y) = 23}$$

ii) covariance matrix of $n \times n$ i.e 2×2

$$S = \begin{bmatrix} cov(x,x) & cov(x,y) \\ cov(y,x) & cov(y,y) \end{bmatrix}$$

$$= \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

<u>step 4</u> Find Eigen value, Eigen vector then normalise

i) Eigen value

$\det(S-\lambda I)=0$ $\quad I \rightarrow$ identity matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(2)

$$\det\left(\begin{bmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{bmatrix}\right) = 0$$

$$(14-\lambda)(23-\lambda) - [(-11)(-11)] = 0$$

$$\lambda^2 - 37\lambda + 201 = 0 \qquad \Rightarrow \frac{-b \pm \sqrt{b^2-4ac}}{2a}$$

$$\underline{\lambda = 30.3849,\ 6.6151} \qquad a=1,\ b=-37,\ c=201$$

arrang $\lambda_1 > \lambda_2$

$\lambda_1 = 30.3849 \Rightarrow$ first principal (largest) component

$\lambda_2 = 6.6151 \Rightarrow$ second principal component

x —x —·(Extra to understand how dimension reduces)x —— x ——

ii) Eigen vector of λ_1

$$(S - \lambda_1 I)\,U_1 = 0 \qquad U_1 = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} 14-\lambda_1 & -11 \\ -11 & 23-\lambda_1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} (14-\lambda_1)u_1 - 11u_2 + \\ -11u_1 + (23-\lambda_1)u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(14-\lambda_1)u_1 - 11u_2 = 0.$$

$$-11u_1 + (23-\lambda_1)u_2 = 0.$$

$$\frac{u_1}{11} = \frac{u_2}{14-\lambda_1} = t$$

when $t=1$

Eigen vector U_1 of $\lambda_1 = \begin{bmatrix} 11 \\ 14-\lambda_1 \end{bmatrix}$

$$= \begin{bmatrix} 11 \\ 14-30.3849 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -16.3849 \end{bmatrix}$$

ii) Normalize U_1

$$e_1 = \begin{bmatrix} 11/\sqrt{(11^2+(-16.3849)^2)} \\ \dfrac{-16.3849}{\sqrt{11^2+(-16.3849)^2}} \end{bmatrix}$$

$$e_1 = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$$

(4)

Similarly, $e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$

Step 5 New data using PC_1

PC_1	P_{11}	P_{12}	P_{13}	P_{14}
	-4.3052	3.7361	5.6928	-5.1238

$$P_{11} = e_1^T \begin{bmatrix} x_1 - \bar{x}_i \\ y_i - \bar{y}_i \end{bmatrix}$$

$$= \begin{bmatrix} 0.5574 & -0.8303 \end{bmatrix} \begin{bmatrix} -4 \\ 2.5 \end{bmatrix}$$

$$= -4.3052$$

Similarly

$P_{12} = 3.7361$

$P_{13} = 5.6928$

$P_{14} = -5.1238$

dimension reduction from 2 to 1 using principal component 1

original data

x	4	8	13	7
y	11	4	5	14

new data (reduced dimension)

PC_1	P_{11}	P_{12}	P_{13}	P_{14}
	-4.3052	3.7361	5.6928	-5.1238

5