Final Maths

ANFIS - Adaptive Network-Based Fuzzy Inference System

Q: Rule 1: IF x in small (A_1) AND y in small (B_2) THEN F_1 = small.

Rule 2: IF x in large (A_2) AND y in large (B_2) THEN F_2 = large.

A_1: $\mu_{A_1}(x) = \dfrac{1}{1+\left|\frac{x-1}{2}\right|^{v}}$

A_2: $\mu_{A_2}(x) = \dfrac{1}{1+\left|\frac{x-9}{2}\right|^{v}}$

B_1: $\mu_{B_1}(y) = \dfrac{1}{1+\left|\frac{y-2}{2}\right|^{v}}$

B_2: $\mu_{B_2}(y) = \dfrac{1}{1+\left|\frac{y-14}{2}\right|^{v}}$

↳ input node

$F_1 = 0.1x + 0.1y + 0.1$

$F_2 = 10x + 10y + 10$

Given ~~that~~ the trained fuzzy ~~model~~ system above and input values of $x = 3$ and $y = 4$, find output of the sugeno fuzzy system.

Soln: Given,

Input Values : $x = 3$
$y = 4$

$$\therefore A_1 : \mu_{A_1}(x) = \frac{1}{1 + \left|\frac{3-1}{2}\right|^2} = \frac{1}{1+1} = 0.5$$

and $$A_2 : \mu_{A_2}(x) = \frac{1}{1 + \left|\frac{3-9}{2}\right|^2} = \frac{1}{10} = 0.1$$

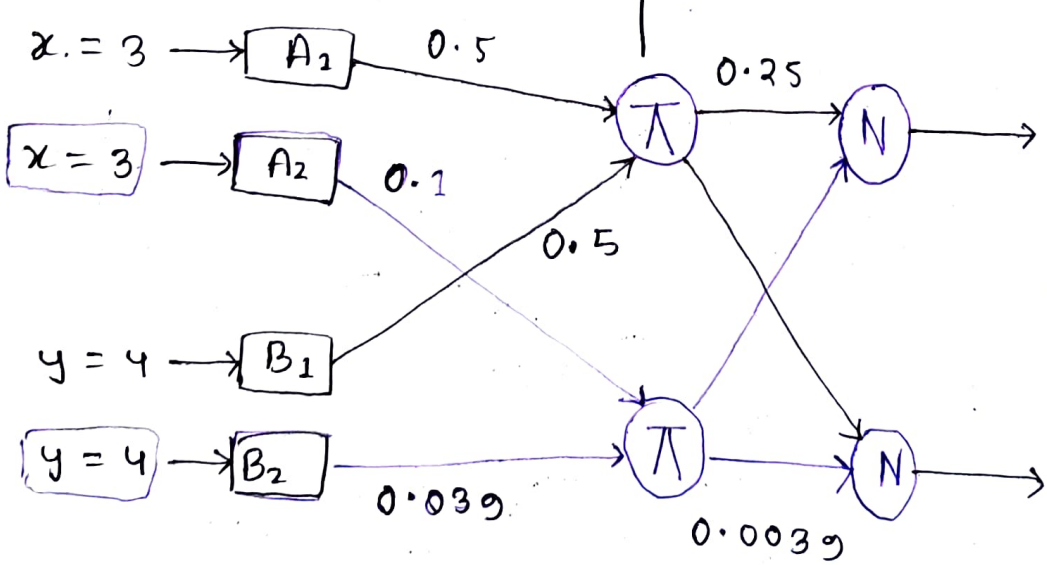
Again,

$$B_1 : \mu_{B_1}(y) = \frac{1}{1 + \left|\frac{4-2}{2}\right|^2} = \frac{1}{1} = 1$$

$$B_2 : \mu_{B_2}(y) = \frac{1}{1 + \left|\frac{4-14}{2}\right|^2} = \frac{1}{26} = 0.0039$$

Now,

(note: It worked as multiplication).



$$W_1 = \frac{0.25}{0.25 + 0.0039}$$

$$\therefore W_1 = \underline{0.985}$$

Average

$$W_2 = \frac{0.0039}{0.25 + 0.0039}$$

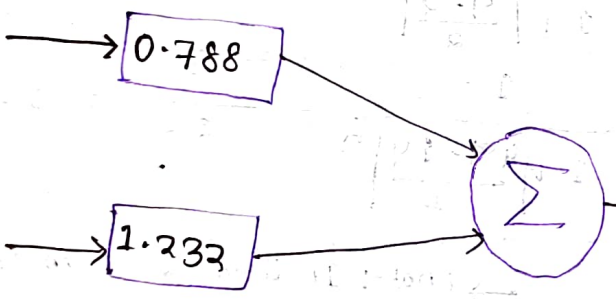
$$= \underline{0.0154}$$

Given, Functions: $f_1 = 0.1x + 0.1y + 0.1$

$f_2 = 10x + 10y + 10$

N $\xrightarrow{W_1 = 0.985}$ $\boxed{W_1 f_1 = (0.985) \times (0.1 \times 3 + 0.1 \times 4 + 0.1) = 0.788}$

N $\xrightarrow{W_2 = 0.0154}$ $\boxed{W_2 f_2 = (0.0154) \times (10 \times 3 + 10 \times 4 + 10) = 1.232}$



$0.788 + 1.232 = 2.02$

$O = 2.02$

Ans:

Fuzzy C Means

Q: Input Number of objects = 6, Number of clusters = 2

X	Y	C1	C2
1	6	0.8	0.2
2	5	0.9	0.1
3	8	0.7	0.3
4	4	0.3	0.7
5	7	0.5	0.5
6	9	0.2	0.8

Soln:

Step-1: Initialize the membership matrix.

Step-2: Find the constraint using the equation.

Formula

$$c_j = \left[\frac{\sum_i [u_j(x_i)]^m x_i}{\sum_i [u_j(x_i)]^m}, \frac{\sum_i [u_j(y_i)]^m y_i}{\sum_i [u_j(y_i)]^m} \right]$$

distance from cluster. power

~~Here, distance~~

Here, m is a fuzzification parameter. The range of m is always [1.25, 2]

$$= \left[\frac{u_1(x_1)^m x_1 + u_1(x_2)^m x_2 + u_1(x_3)^m x_3 + u_1(x_4)^m x_4 + u_1(x_5)^m x_5 + u_1(x_6)^m x_6}{u_1(x_1)^m + u_1(x_2)^m + u_1(x_3)^m + u_1(x_4)^m + u_1(x_5)^m + u_1(x_6)^m}, \right.$$

$$\left. \frac{u_1(y_1)^m y_1 + u_1(y_2)^m y_2 + u_1(y_3)^m y_3 + u_1(y_4)^m y_4 + u_1(y_5)^m y_5 + u_1(y_6)^m y_6}{u_1(y_1)^m + u_1(y_2)^m + u_1(y_3)^m + u_1(y_4)^m + u_1(y_5)^m + u_1(y_6)^m} \right]$$

$$c_1 = \left[\frac{1*0.8^2 + 2*0.9^2 + 3*0.7^2 + 4*0.3^2 + 5*0.5^2 + 6*0.2^2}{0.8^2 + 0.9^2 + 0.7^2 + 0.3^2 + 0.5^2 + 0.2^2}, \right.$$

$$\left. \frac{6*0.8^2 + 5*0.9^2 + 8*0.7^2 + 4*0.3^2 + 7*0.5^2 + 9*0.2^2}{0.8^2 + 0.9^2 + 0.7^2 + 0.3^2 + 0.5^2 + 0.2^2} \right]$$

$$c_1 = \frac{5.58}{2.32}, \frac{14.28}{2.32}$$

$$\therefore c_1 = (2.4, 6.1)$$

Again, $c_2 = \left[\frac{1*0.2^2 + 2*0.1^2 + 3*0.3^2 + 4*0.7 \ldots *0.5^2}{0.2^2 + 0.1^2 + 0.3^2 + 0.7^2 +} \right.$

$$\frac{+ 6*0.8^2}{0.5^2 + 0.8^2}, \frac{6*0.2^2 + 5*0.1^2 + 8*0.3^2 +}{0.2^2 + 0.1^2 + 0.3^2 + 0.7^2 +}$$

$$\left. \frac{4*0.7 + 7*0.5^2 + 9*0.8^2}{0.5^2 + 0.8^2} \right]$$

$$c_2 = \frac{7.38}{1.52}, \frac{10.48}{1.52}$$

$$\therefore c_2 = (4.8, 6.8)$$

Step-3: Find Distance.

$$D_i = \sqrt{(x_2 - x_1)^2 + (y_2 + y_1)^2}$$

Centroid 1:

$(\underset{x_1}{1}, \underset{y_1}{6})(\underset{x_2}{2.4}, \underset{y_2}{6.1}) = \sqrt{(1.4)^2 + (0.1)^2} = \sqrt{1.96 + 0.01} = \sqrt{1.97} = 1.40$

$(2,5)(2.4, 6.1) = \sqrt{(2-2.4)^2 + (5-6.1)^2} = \sqrt{0.16 + 1.21} = \sqrt{1.37} = 1.17$

$(3,8)(2.4, 6.1) = \sqrt{0.36 + 3.61} = \sqrt{3.97} = 1.99$

$(4,4)(2.4, 6.1) = \sqrt{2.56 + 4.41} = \sqrt{6.97} = 2.64$

$(5,7)(2.4, 6.1) = \sqrt{6.76 + 0.81} = \sqrt{7.57} = 2.75$

$(6,9)(2.4, 6.1) = \sqrt{12.96 + 8.41} = \sqrt{21.37} = 4.62$

Centroid 2:

$(1,6)(4.8, 6.8) = \sqrt{14.44 + 0.64} = \sqrt{15.08} = 3.88$

$(2,5)(4.8, 6.8) = \sqrt{7.84 + 3.24} = \sqrt{11.08} = 3.32$

$(3,8)(4.8, 6.8) = \sqrt{3.24 + 1.44} = \sqrt{4.68} = 2.16$

$(4,4)(4.8, 6.8) = \sqrt{0.64 + 7.84} = \sqrt{8.48} = 2.91$

$(5,7)(4.8, 6.8) = \sqrt{0.04 + 0.04} = \sqrt{0.08} = 0.28$

$(6,9)(4.8, 6.8) = \sqrt{1.44 + 4.84} = \sqrt{6.28} = 2.50$

cluster 1		cluster 2		
Datapoint	Distance	Datapoint	Distance	Distance
(1,6)	1.40	(1,6)	~~(1,6)~~	3.88
(2,5)	1.17	(2,5)	~~(2,5)~~	3.32
(3,8)	2.09	(3,8)	~~(3,8)~~	2.16
(4,4)	2.64	(4,4)	~~(4,4)~~	2.91
(5,7)	2.75	(5,7)	~~(5,7)~~	0.28
(6,9)	4.62	(6,9)	~~(6,9)~~	2.50

Step-4: Update the membership value.

$$u_j(x_i) = \frac{\left[\frac{1}{d_{ji}}\right]^{1/m-1}}{\sum_{k=1}^{c}\left[\frac{1}{d_{ki}}\right]^{1/m-1}}$$

here, $m=2$, i - first data point, j - first cluster
c=cluster

Cluster 1:
Now,

$$u_{11} = \frac{\left(\frac{1}{d_{11}}\right)^{\frac{1}{2-1}}}{\sum_{k=1}^{c}\left(\frac{1}{d_{ki}}\right)^{\frac{1}{2-1}}}$$

Here,
$c=2$
(cause there are 2 clusters.)

$$u_{11} = \frac{\left(\frac{1}{d_{11}}\right)^1}{\left(\frac{1}{d_{11}}\right)^1 + \left(\frac{1}{d_{21}}\right)^1}$$

$$= \frac{\frac{1}{1.4}}{\left(\frac{1}{1.4}\right)^1 + \left(\frac{1}{3.88}\right)^1} = \frac{0.71}{0.71+0.25}$$

$$= 0.7$$

$$u_{12} = \frac{\frac{1}{d_{12}}}{\left(\frac{1}{d_{12}}\right)+\left(\frac{1}{d_{22}}\right)} = \frac{\frac{1}{1.17}}{\frac{1}{1.17}+\frac{1}{3.32}} = \frac{0.56}{0.86} = 0.6$$

$$u_{13} = \frac{\frac{1}{d_{13}}}{(1/d_{13})+(1/d_{23})} = \frac{1/1.99}{(1/1.99)+(1/2.16)} = \frac{0.50}{0.96}$$
$$= 0.5$$

$$u_{14} = \frac{(1/d_{14})}{(1/d_{14})+(1/d_{24})} = \frac{1/2.64}{(1/2.64)+(1/2.91)} = \frac{0.37}{0.37+0.34}$$
$$= 0.5$$

$$u_{15} = \frac{(1/d_{15})}{(1/d_{15})+(1/d_{25})} = \frac{1/2.75}{(1/2.75)+(1/0.28)} = \frac{0.36}{3.03} = 0.1$$

$$u_{16} = \frac{(1/d_{16})}{(1/d_{16})+(1/d_{26})} = \frac{(1/4.62)}{(1/4.62)+(1/2.50)} = \frac{0.21}{0.61} = 0.3$$

Cluster 2:

$$u_{21} = \frac{(1/d_{21})}{(1/d_{12})+(1/d_{21})} = \frac{1/3.88}{(1/1.40)+(1/3.88)} = \frac{0.25}{0.96} = 0.3$$

$$u_{22} = \frac{(1/d_{22})}{(1/d_{12})+(1/d_{22})} = \frac{(1/3.32)}{(1/1.17)+(1/3.32)} = \frac{0.30}{0.86} = 0.4$$

$$u_{23} = \frac{(1/d_{23})}{(1/d_{13})+(1/d_{23})} = \frac{(1/2.16)}{(1/1.99)+(1/2.16)} = \frac{0.46}{0.96} = 0.5$$

$$u_{24} = \frac{(1/d_{24})}{(1/d_{14})+(1/d_{24})} = \frac{1/2.09}{(1/2.64)+(1/2.09)} = \frac{0.34}{0.71}$$

$$= \underline{0.5}$$

$$u_{25} = \frac{(1/d_{25})}{(1/d_{15})+(1/d_{25})} = \frac{(1/0.28)}{(1/2.75)+(1/0.28)} = \frac{3.57}{3.93}$$

$$= \underline{0.9}$$

$$u_{26} = \frac{(1/d_{26})}{(1/d_{16})+(1/d_{26})} = \frac{0.4}{0.21+0.4} = \frac{0.4}{0.61} = \underline{0.7}$$

Now, the new membership value is:

X	Y	C1	C2
1	6	0.7	0.3
2	5	0.6	0.4
3	8	0.5	0.5
4	4	0.5	0.5
5	7	0.1	0.9
6	9	0.3	0.7

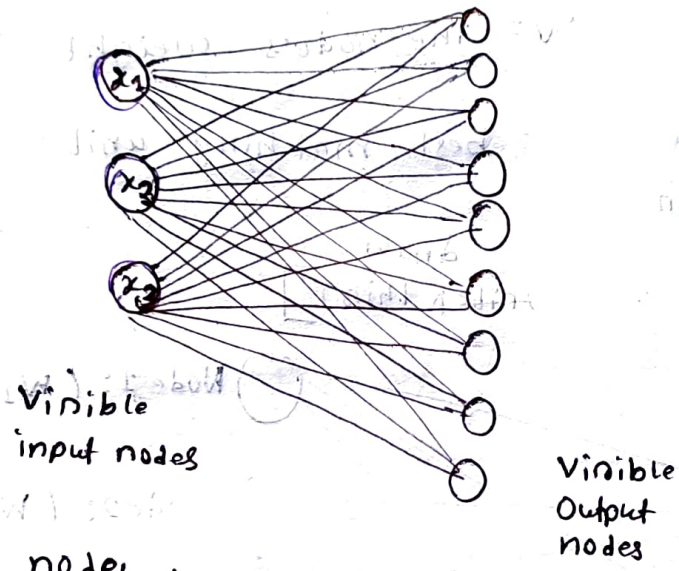
Am:

→ Exam এ যদি Obj 6 টা থাকে তাহলে 1/2 টা point এর updated value ফিক্সই হবে। সবগুলো করা লাগবে না i.e. $u_{11}, u_{12}, u_{21}, u_{22}$ ফিক্সই হবে।

Self Organizing Maps (SOM)

Q: Question can be like this:-

Consider the structure of Self Organizing which has 3 visible input nodes and 9 outputs that are connected ~~to input~~ directly to input as shown below fig.



Q: Calculate the best matching unit.

input nodes values are:

$x_1 = 0.7$

$x_2 = 0.6$

$x_3 = 0.9$

The values of weights:

$W_{1,1} = 0.31$	$W_{1,2} = 0.22$	$W_{1,3} = 0.10$
$W_{2,1} = 0.21$	$W_{2,2} = 0.34$	$W_{2,3} = 0.39$
$W_{3,1} = 0.39$	$W_{3,2} = 0.42$	$W_{3,3} = 0.45$
$W_{4,1} = 0.25$	$W_{4,2} = 0.32$	$W_{4,3} = 0.62$
$W_{5,1} = 0.24$	$W_{5,2} = 0.31$	$W_{5,3} = 0.16$
$W_{6,1} = 0.52$	$W_{6,2} = 0.33$	$W_{6,3} = 0.42$
$W_{7,1} = 0.31$	$W_{7,2} = 0.22$	$W_{7,3} = 0.10$
$W_{8,1} = 0.12$	$W_{8,2} = 0.41$	$W_{8,3} = 0.19$
$W_{9,1} = 0.34$	$W_{9,2} = 0.40$	$W_{9,3} = 0.51$

Sol^n: calculating the best matching unit:

We know,

The Euclidean ~~dian~~ distance:

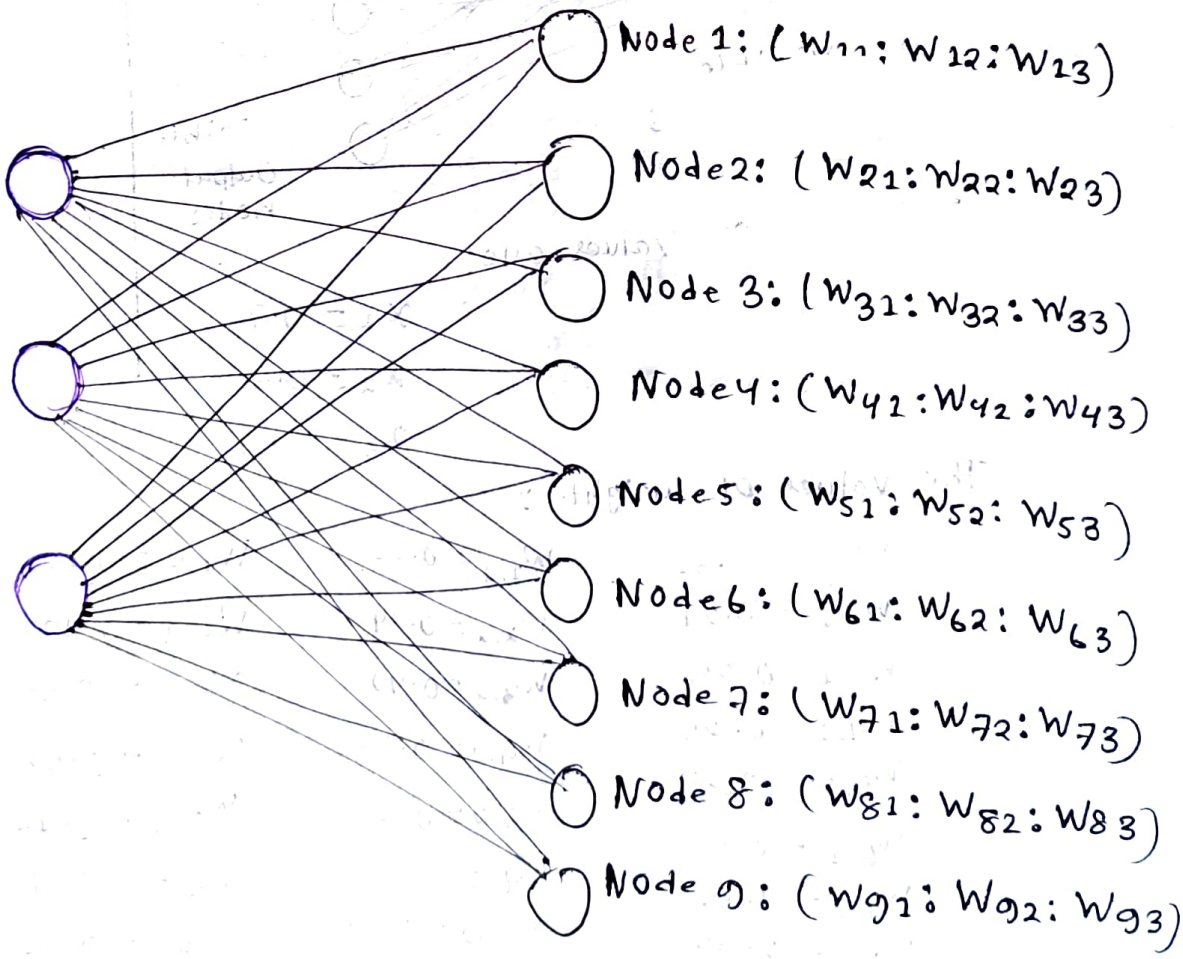
$$\text{Distance} = \sqrt{\sum_{i=0}^{i=n} (X_i - W_i)^2}$$

Here, X = The current input vector

W = The node's weight vector.

Let's calculate the best matching unit using the ~~Dis~~ Distance formula.

[* (Don't need to ~~write~~ draw this)]



1st Node distance $= \sqrt{(x_1 - W_{11})^2 + (x_2 - W_{12})^2 + (x_3 - W_{13})^2}$

$= \sqrt{(0.7 - 0.31)^2 + (0.6 - 0.22)^2 + (0.9 - 0.10)^2}$

$= \sqrt{0.9365}$

$= 0.968$

2nd Node distance $= \sqrt{(x_1 - W_{21})^2 + (x_2 - W_{22})^2 + (x_3 - W_{23})^2}$

$= \sqrt{(0.7 - 0.21)^2 + (0.6 - 0.34)^2 + (0.9 - 0.19)^2}$

$= \sqrt{0.8118}$

$= 0.901$

3rd Node distance $= \sqrt{(x_1 - W_{31})^2 + (x_2 - W_{32})^2 + (x_3 - W_{33})^2}$

$= \sqrt{(0.7 - 0.39)^2 + (0.6 - 0.42)^2 + (0.9 - 0.45)^2}$

$= \sqrt{0.331}$

$= 0.575$

4th Node distance $= \sqrt{(x_1 - W_{41})^2 + (x_2 - W_{42})^2 + (x_3 - W_{43})^2}$

$= \sqrt{(0.7 - 0.25)^2 + (0.6 - 0.32)^2 + (0.9 - 0.62)^2}$

$= \sqrt{0.3593}$

$= 0.599$

5th node distance $= \sqrt{(x_1 - W_{51})^2 + (x_2 - W_{52})^2 + (x_3 - W_{53})^2}$

$= \sqrt{(0.7 - 0.24)^2 + (0.6 - 0.31)^2 + (0.9 - 0.16)^2}$

$= \sqrt{0.8433}$

$= 0.918$

6th node distance $= \sqrt{(x_1 - W_{61})^2 + (x_2 - W_{62})^2 + (x_3 - W_{63})^2}$

$= \sqrt{(0.7 - 0.52)^2 + (0.6 - 0.33)^2 + \cancel{(0.9 - 0.42)^2}}$
$(0.9 - 0.42)^2$

$= \sqrt{0.3357}$

$= 0.579$

7th node distance $= \sqrt{(x_1 - W_{71})^2 + (x_2 - W_{72})^2 + (x_3 - W_{73})^2}$

$= \sqrt{(0.7 - 0.31)^2 + (0.6 - 0.22)^2 + (0.9 - 0.10)^2}$

$= \sqrt{0.9365}$

$= 0.968$

8th node distance $= \sqrt{(x_1 - W_{81})^2 + (x_2 - W_{82})^2 + (x_3 - W_{83})^2}$

$= \sqrt{(0.7 - 0.12)^2 + (0.6 - 0.41)^2 + (0.9 - 0.29)^2}$

$= \sqrt{0.87\cancel{3}66}$

$= 0.936$

9th node distance $= \sqrt{(x_1 - W_{91})^2 + (x_2 - W_{92})^2 + (x_3 - W_{93})^2}$

$= \sqrt{(0.7 - 0.34)^2 + (0.6 - 0.40)^2 + (0.9 - 0.51)^2}$

$= \sqrt{0.3217}$

$= 0.5672$

Since we have calculated all the values of respected nodes. Now it's time to calculate the best match unit.

As we can see, node number "9" is the closest with a distance of 0.5672. So, we can call this node our BMU (Best-matching Unit).

Ans:

CNN - Convolutional Neural Network

Q: Question can be like this:

Given the input size, filter size, padding, stride and dilation you can calculate the output size of the convolution operation.

This math will come with "ID".

0	2	1	0	1
0	1	2	1	0
1	0	2	0	0
1	0	0	1	1
0	1	1	2	2

Input Vector

0	1	-4
0	2	0
-1	4	3

Filter

Soln: calculating the size of the output:

0	2	1	0	1
0	1	2	1	0
1	0	2	0	0
1	0	0	1	1
0	1	1	2	2

5×5

$*$

0	1	-4
0	2	0
-1	4	3

3×3

We are doing this with stride - 1 :.

$0\times0 + 2\times1 + 1\times(-4) + 0\times0 + 1\times2 + 2\times0 + 1\times(-1) + 0\times4 + 2\times3$

$= 0 + 2 - 4 + 0 + 2 + 0 - 1 + 0 + 6$

$= 5$

$2\times0 + 1\times1 + 0\times(-4) + 1\times0 + 2\times2 + 1\times0 + 0\times(-1) + 2\times4 + 3\times0$

$= 0 + 1 - 4 + 0 + 4 + 0 - 1 + 8 + 0$

$= 8$

$1\times0 + 0\times1 + 1\times(-4) + 2\times0 + 1\times2 + 0\times0 + 2\times(-1) + 0\times4 + 0\times3$

$= 0 + 0 - 4 + 0 + 2 + 0 - 2 + 0 + 0$

$= -4$

$0\times0 + 1\times1 + 2\times(-4) + 1\times0 + 0\times2 + 2\times0 + 1\times(-1) + 0\times4 + 0\times3$

$= 0 + 1 - 8 + 0 + 0 + 0 - 1 + 0 + 0$

$= -8$

$1\times 0 + 2\times 1 + 1\times(-4) + 0\times 0 + 2\times 2 + 0\times 0 + 0\times(-1) + 0\times 4 + 1\times 3$

$= 0 + 2 - 4 + 0 + 4 + 0 - 1 + 0 + 3$

$= \underline{4}$

$2\times 0 + 1\times 1 + 0\times(-4) + 2\times 0 + 0\times 2 + 0\times 0 + 0\times(-1) + 1\times 4 + 1\times 3$

$= 0 + 1 + 0 + 0 + 0 + 0 - 0 + 4 + 3$

$= \underline{8}$

$1\times 0 + 0\times 1 + 2\times(-4) + 1\times 0 + 0\times 2 + 0\times 0 + 0\times -1 + 1\times 4 + 1\times 3$

$= 0 + 0 - 8 + 0 + 0 + 0 + 0 + 4 + 3$

$= \underline{-1}$

$0\times 0 + 2\times 1 + 0\times(-4) + 0\times 0 + 0\times 2 + 1\times 0 + 1\times(-1) + 1\times 4 + 2\times 3$

$\cancel{= 0 + 2 + 0 + 0 + 0 + 0}$

$\cancel{= 2}$

$= 0 + 2 - 0 + 0 + 0 + 0 - 1 + 4 + 6$

$= \underline{11}$

$2\times0 + 0\times1 + 0\times(-4) + 0\times0 + 1\times2 + 1\times0 + 1\times(-1) + 2\times4 + 2\times3$

$= 0 + 0 - 4 + 0 + 2 + 0 - 1 + 8 + 6$

$= 11$

5	8	-4
-8	4	8
-1	11	11

Output

Ans:

RNN - Recurrent Neural Network

Q: Question can be like this:

Find out how the above structure can be used to predict, the fifth letter in the word "hello".

The given vocabulary is {h, e, l, o}.

Forward Propagation in a Recurrent Neuron in Excel. Let's take a look at the inputs first -

1	0	0	0
0	1	0	0
0	0	1	1
0	0	0	0
h	e	l	l

Why		
0.371	0.974	0.830
0.391	0.282	0.659
0.640	0.098	0.334
0.922	0.325	0.144

Randomly initialize the weights as 3 * 4 matrix -

Wxh			
0.287	0.846	0.572	0.486
0.902	0.871	0.691	0.189
0.537	0.092	0.558	0.491

Whh as the weight and bias (1*1 matrix) =

0.427	Whh
0.567	bias

Predict the 5th word of "hello" using RNN.

Soln: For "h"

Step 1:

Now, For ~~the~~ the letter "h", For the hidden state we would need $W_{xh} * x_t$. By matrix multiplication, we get it as -

W_{xh}			
0.287	0.846	0.572	0.486
0.902	0.872	0.692	0.189
0.537	0.092	0.558	0.491

$\times$

1
0
0
0
h

$=$

0.287
0.902
0.537

Step 2:

For the letter "h", the previous state is [0,0,0]. Since there is no letter prior to it.

So, $W_{hh} * h_{t-1} + bias$

Weight (W_{hh})	bias
0.427	0.567

$\times$

0
0
0
h_{t-1}

$=$

0.567
0.567
0.567

((math: $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} * [0.427]$; $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + [0.567]$ w))

Step 3:

Now, we get the current state as –

$$h_t = \tanh(W_{hh} h_{t-1} + W_{xh} X_t)$$

$$\begin{bmatrix} 0.567 \\ 0.567 \\ 0.567 \end{bmatrix} + \begin{bmatrix} 0.287 \\ 0.902 \\ 0.537 \end{bmatrix} = \begin{Bmatrix} 0.854 \\ 1.469 \\ 1.104 \end{Bmatrix}$$

$$H_t = \tanh \begin{Bmatrix} 0.854 \\ 1.469 \\ 1.104 \end{Bmatrix} = \begin{bmatrix} 0.693 \\ 0.899 \\ 0.802 \end{bmatrix}$$

This is "H_{t-1}" for the next state

Step 4:

For "e" $W_{xh} X_t$ will be –

$$\begin{bmatrix} 0.287 & 0.846 & 0.572 & 0.486 \\ 0.902 & 0.871 & 0.697 & 0.189 \\ 0.537 & 0.093 & 0.558 & 0.491 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ e \end{bmatrix} = \begin{bmatrix} 0.846 \\ 0.871 \\ 0.093 \end{bmatrix}$$

Step 4:

Now, calculating h_t for the letter "e".

$$H_t = \tanh(W_{hh} h_{t-1} + W_{xh} x_t)$$

For "e"

$W_{hh} * h_{t-1} + bias$ will be -

$$W_{hh} * h_{t-1} + bias = 0.427 * \begin{bmatrix} 0.693 \\ 0.899 \\ 0.802 \end{bmatrix} + \boxed{0.567}$$

$$\begin{bmatrix} 0.863 \\ 0.950 \\ 0.909 \end{bmatrix}$$

Now,

$W_{xh} x_t$ will be -

$$\begin{bmatrix} 0.287 & 0.846 & 0.572 & 0.486 \\ 0.902 & 0.871 & 0.691 & 0.189 \\ 0.537 & 0.093 & 0.558 & 0.491 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ e \end{bmatrix} = \begin{bmatrix} 0.846 \\ 0.871 \\ 0.093 \end{bmatrix}$$

Step 5:

Now calculating h_t for the letter "e",

$$H_t = \tanh(W_{hh} h_{t-1} + W_{xh} x_t)$$

$$H_t = \text{TANH}\left\{ \begin{bmatrix} 0.863 \\ 0.950 \\ 0.009 \end{bmatrix} + \begin{bmatrix} 0.846 \\ 0.871 \\ 0.093 \end{bmatrix} \right\}$$

$$= \text{TANH}\left\{ \begin{matrix} 1.709 \\ 1.821 \\ 1.002 \end{matrix} \right\}$$

$$= \begin{bmatrix} 0.936 \\ 0.948 \\ 0.762 \end{bmatrix}$$

This is the "h_{t-1}" for the next state.

Step 6:

Calculating y_t for "e":

$$y_t = W_{hy} h_t$$

$$\begin{bmatrix} 0.371 & 0.978 & 0.830 \\ 0.391 & 0.282 & 0.650 \\ 0.649 & 0.098 & 0.334 \\ 0.913 & 0.326 & 0.144 \end{bmatrix} \times \begin{bmatrix} 0.936 \\ 0.948 \\ 0.762 \end{bmatrix} = \begin{bmatrix} 1.904 \\ 1.135 \\ 0.9505 \\ 1.273 \end{bmatrix}$$

W_{hy} $\qquad$ H_t $\qquad$ y_t

Step 7:

(Don't need to write this step if you don't have enough time) ✓

$$\boxed{\text{Classwise Probabilities for the next letter}} = \text{Softmax}\begin{Bmatrix} 0.410 \\ 0.194 \\ 0.162 \\ 0.223 \end{Bmatrix}$$

Ans