

COMP9444 22T2 ASS1

z5380457 | XiangZheng

Part 1

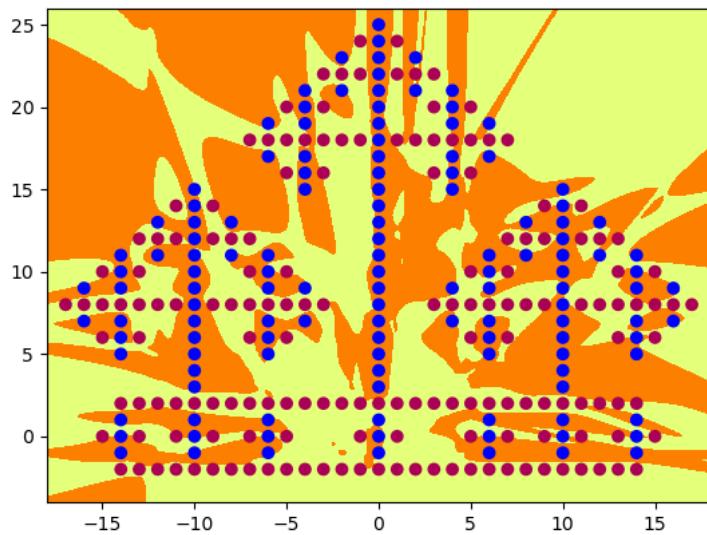
1. see cdoe

2. Full3Net

Hidden nodes	lr	init	acc	ep
10	0.001	0.3	~79%	11400
15	0.001	0.3	~92%	15500
16	0.001	0.3	~94%	41800
17	0.001	0.3	~90%	20000
18	0.001	0.3	~94%	32900
19	0.001	0.3	~97%	33600
20	0.001	0.3	~93%	33900
21	0.001	0.3	~98%	33400
22	0.001	0.3	~96%	21700
23	0.001	0.3	100%	18200

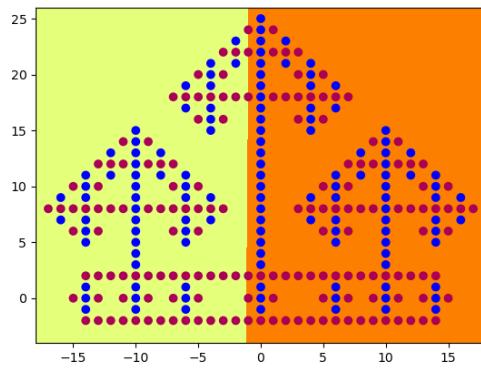
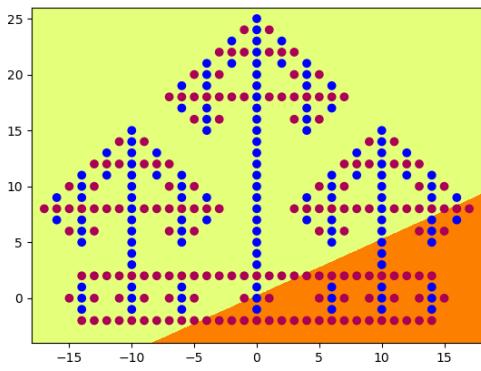
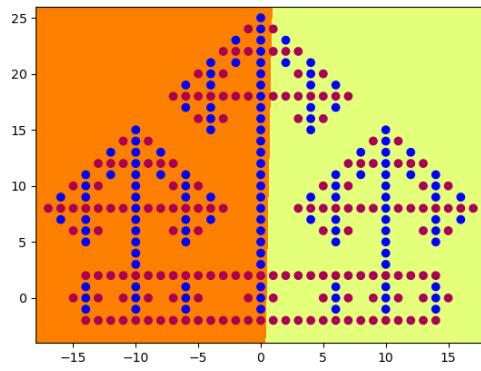
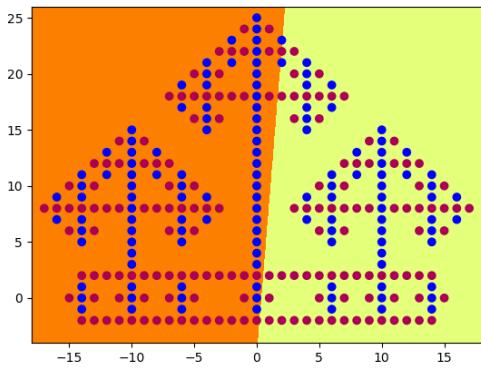
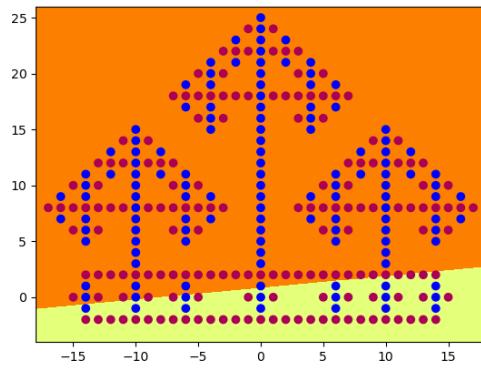
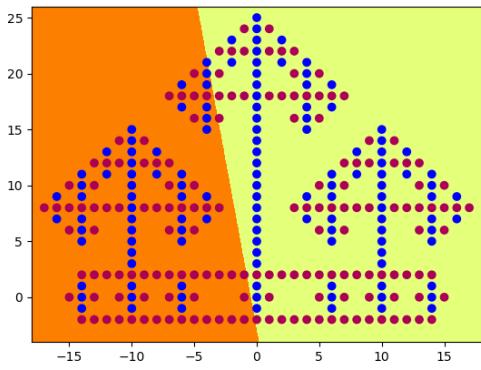
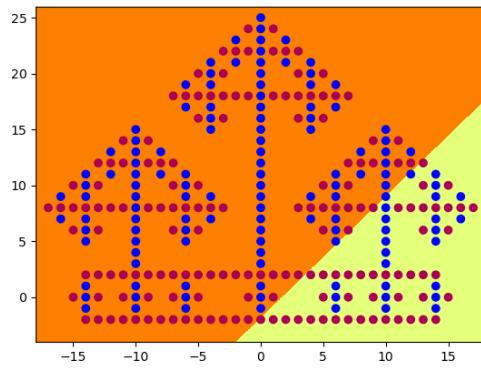
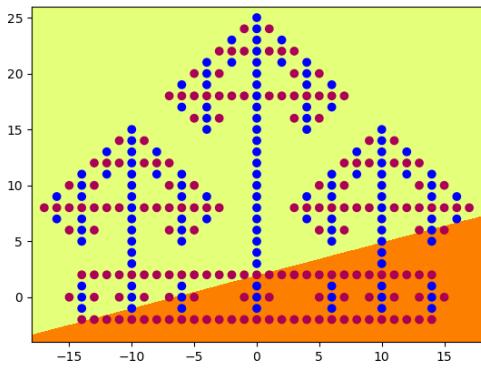
The number of hidden nodes close to the minimum required is 23

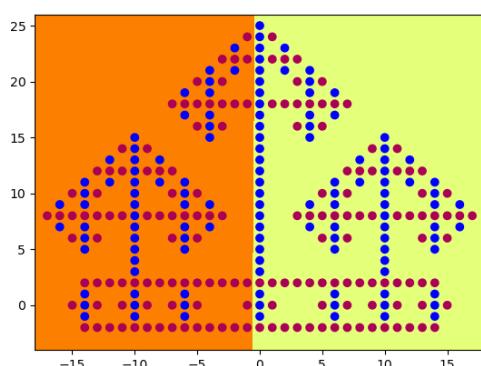
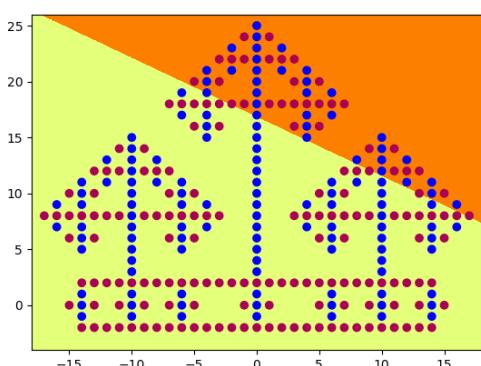
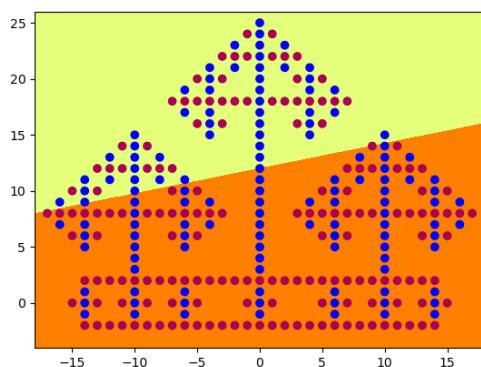
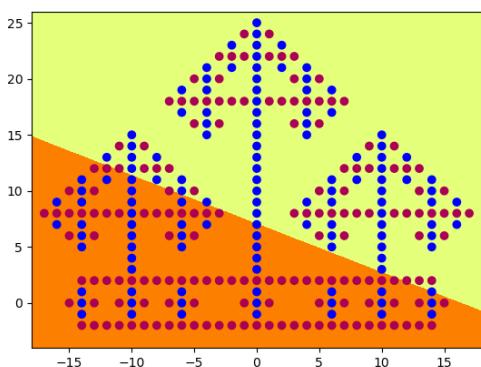
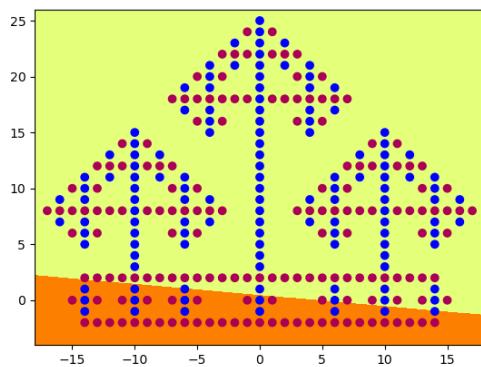
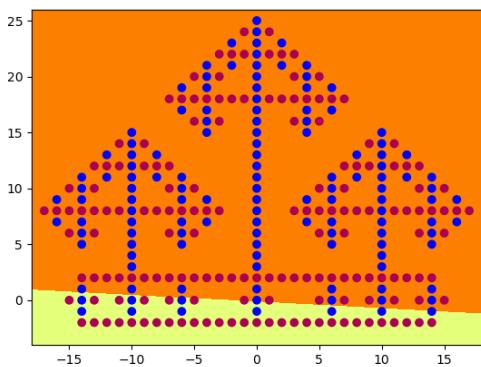
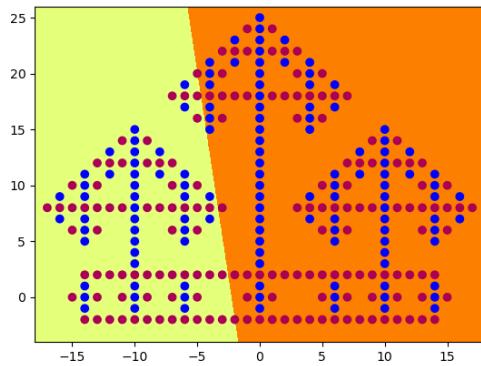
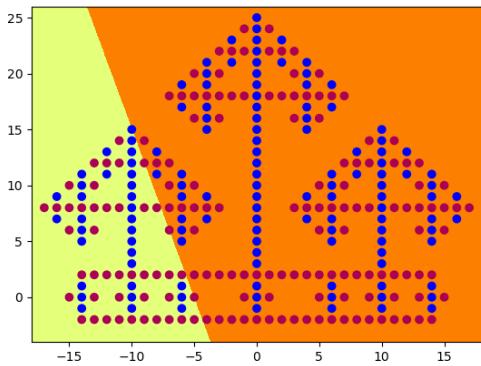
The picture of node 23 is:

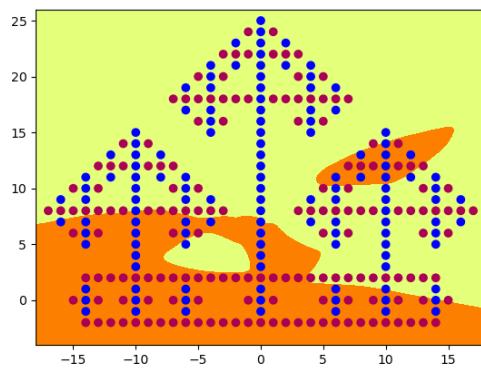
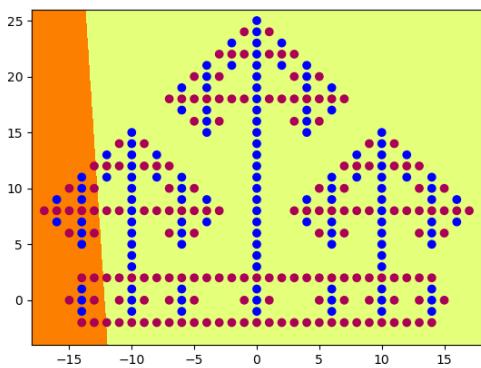
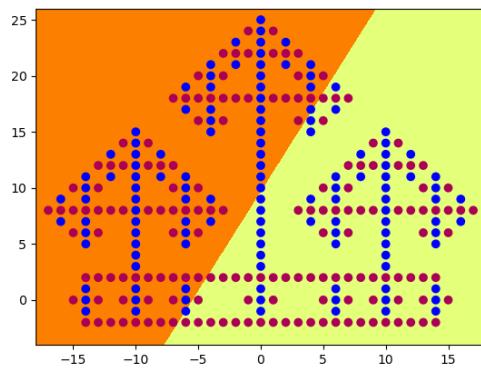
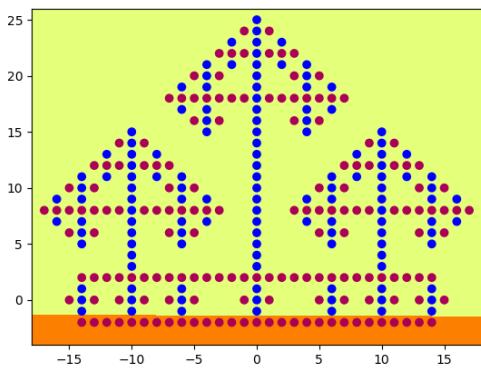
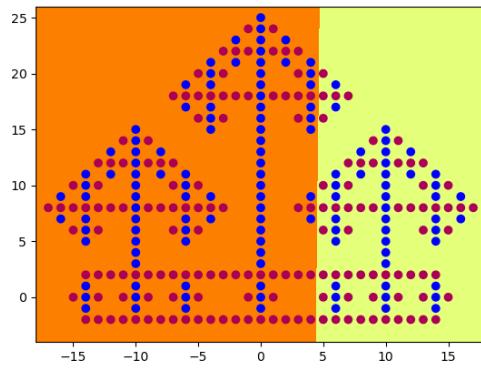
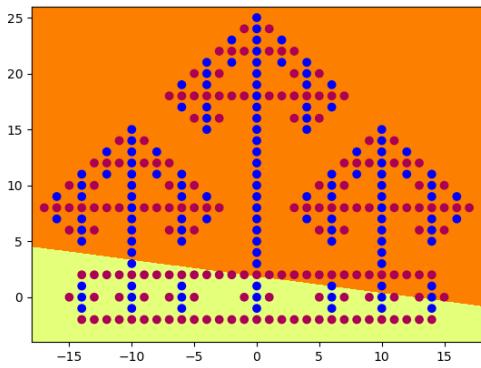
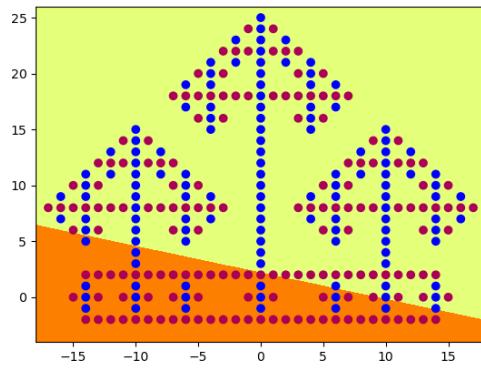
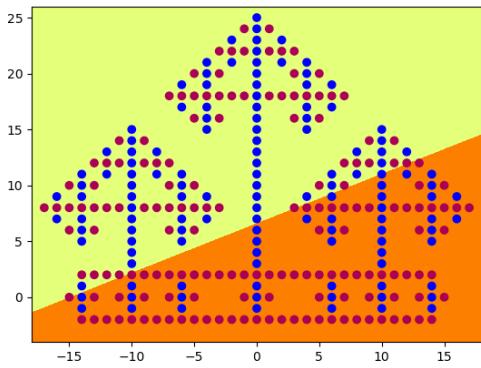


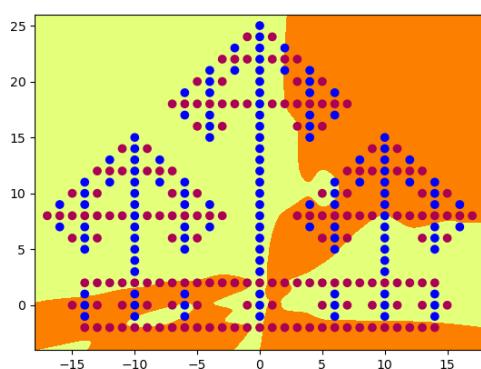
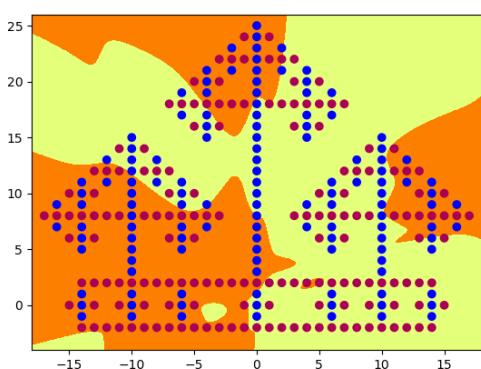
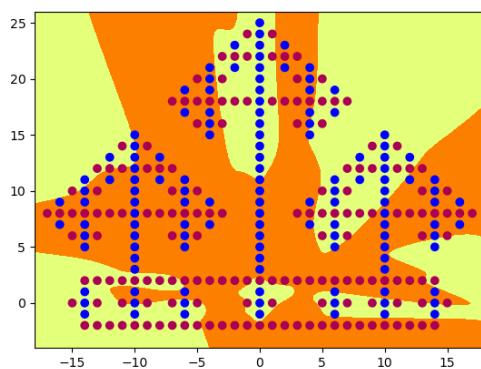
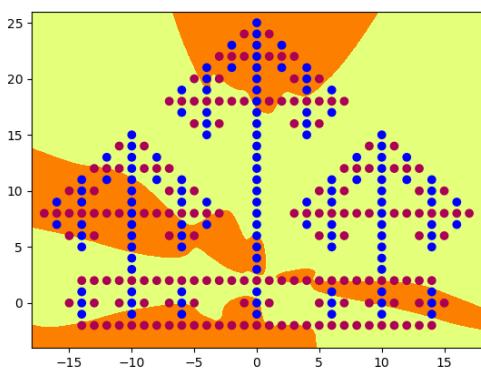
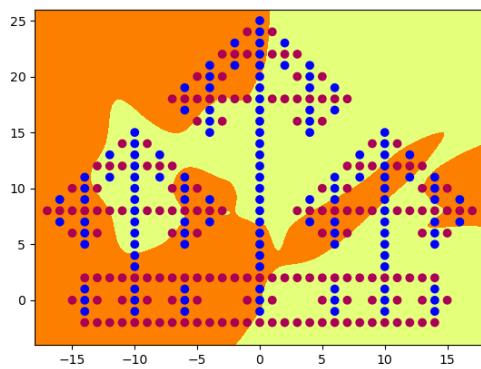
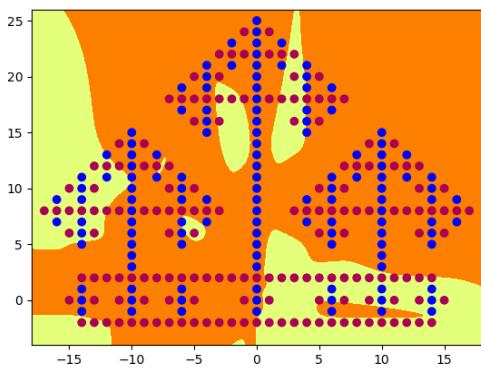
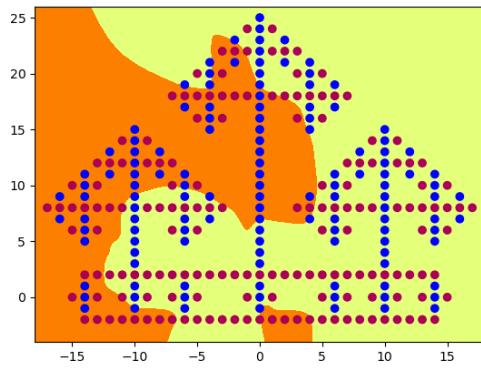
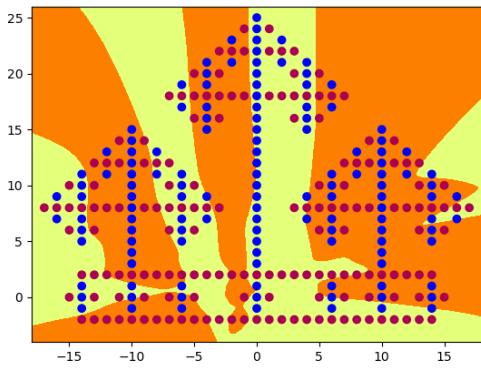
out_full3_23

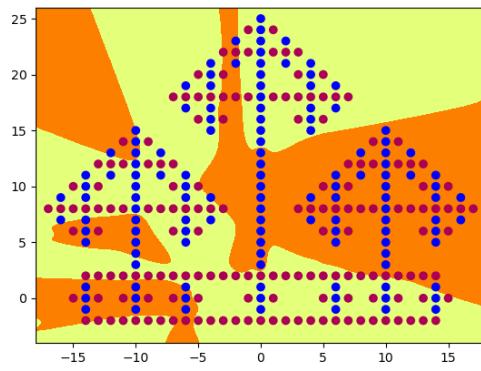
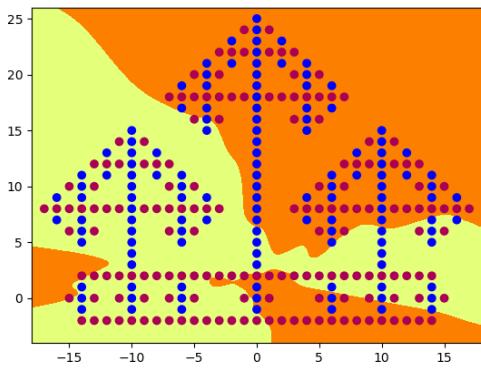
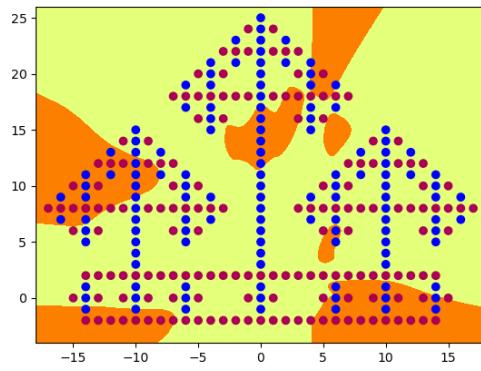
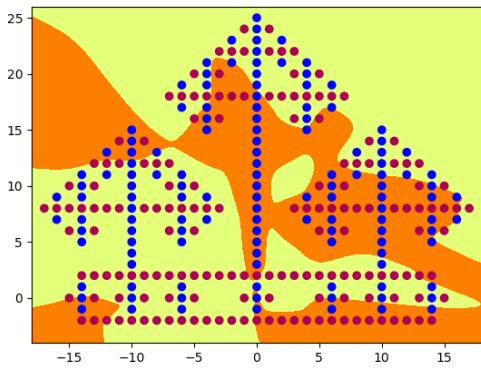
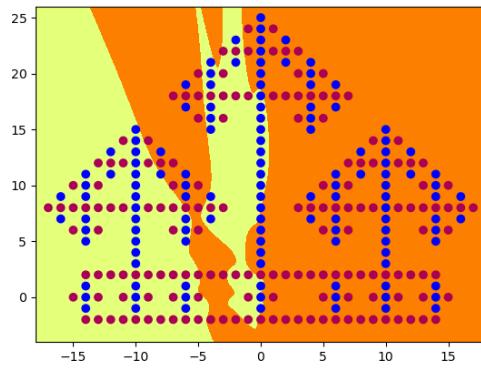
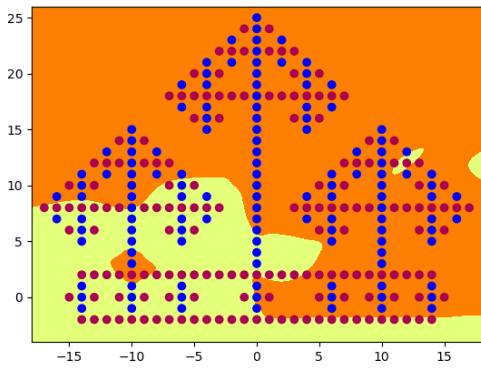
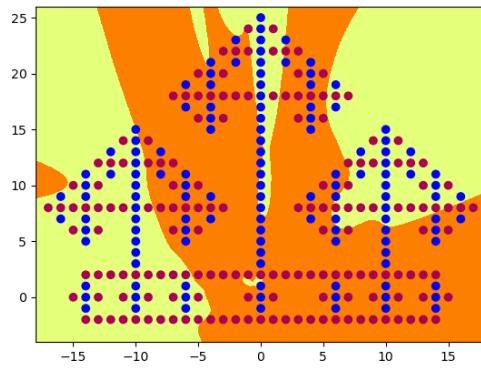
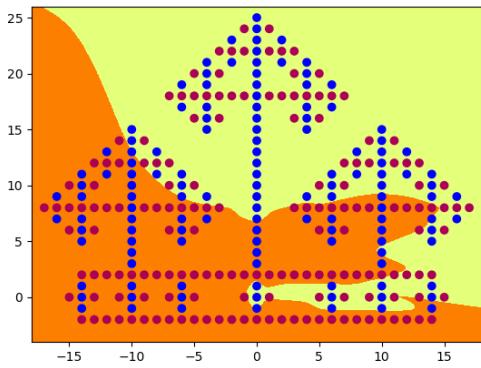
All hidden units:

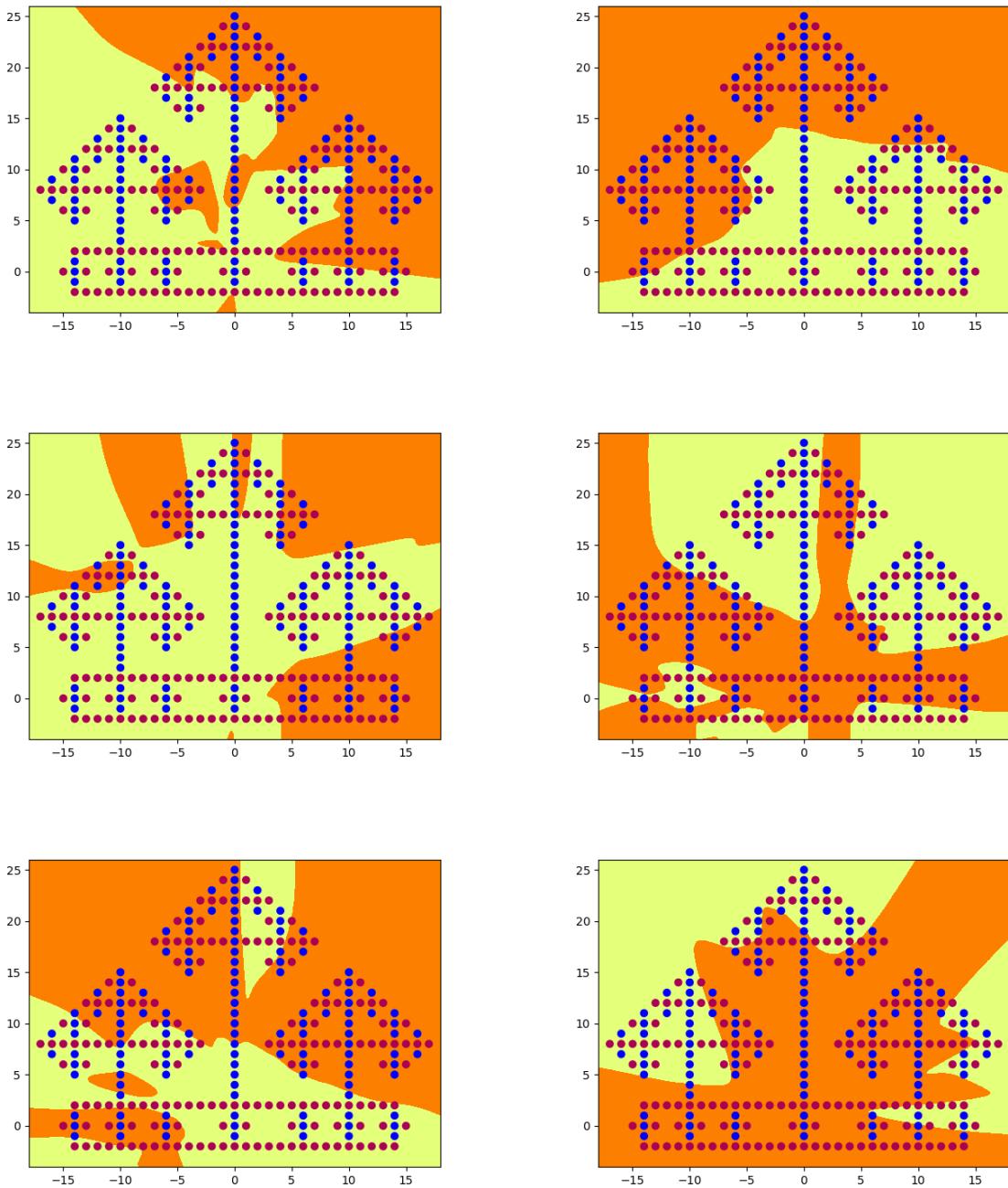












The total number of independent parameters:

$$(23 \times 1 + 1) + (23 \times 23 + 23) + (23 \times 2 + 23) = 645$$

3. see code

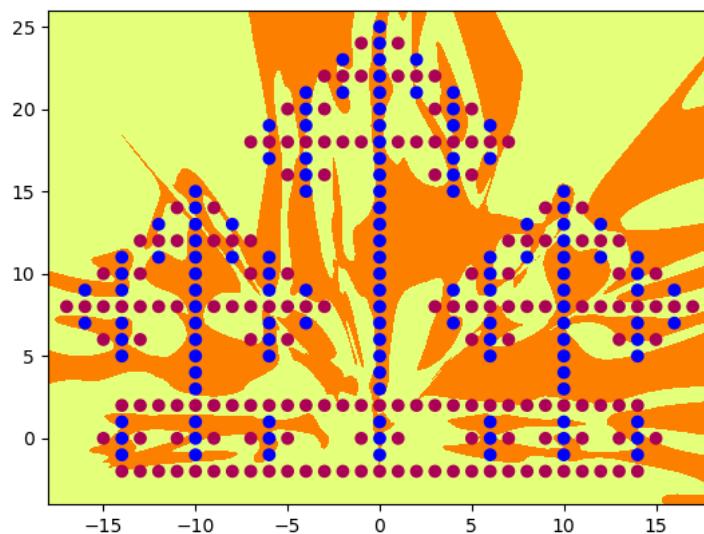
4. Full4Net

Hidden nodes	lr	init	acc	ep
20	0.001	0.3	100%	15800
19	0.001	0.3	~89%	29600
18	0.001	0.3	~85%	25100
17	0.001	0.3	~84%	20800

16	0.001	0.3	~84%	20200
15	0.001	0.3	~92%	18600
10	0.001	0.3	~90%	30000

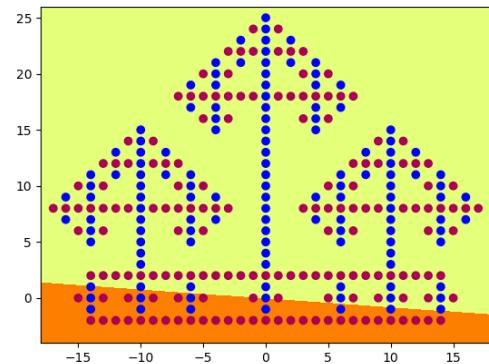
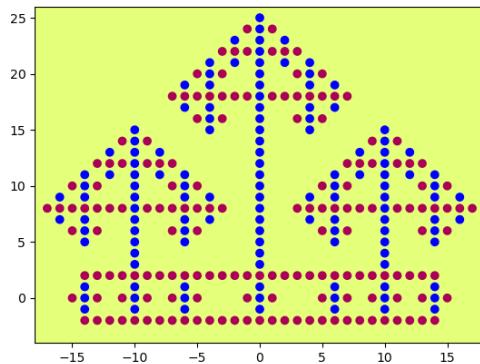
The number of hidden nodes close to the minimum required is 20

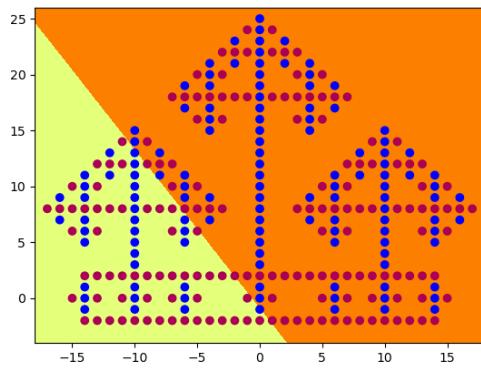
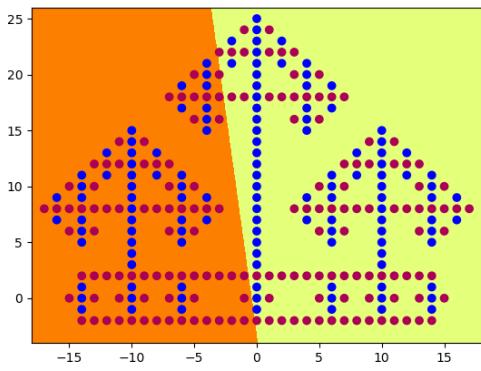
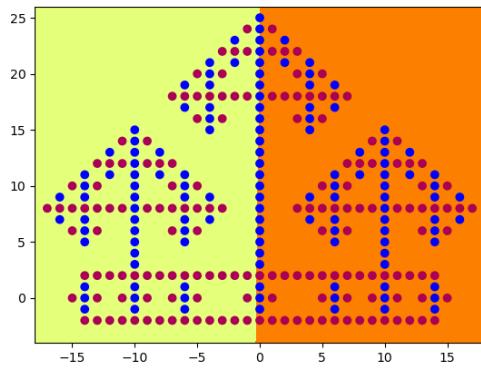
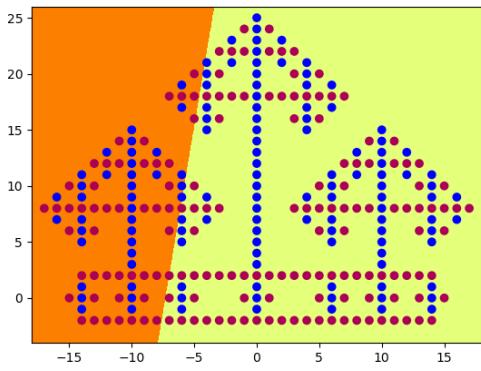
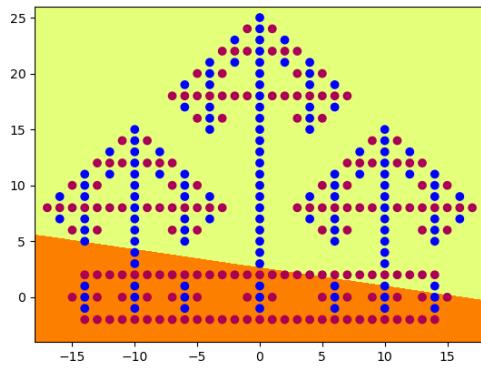
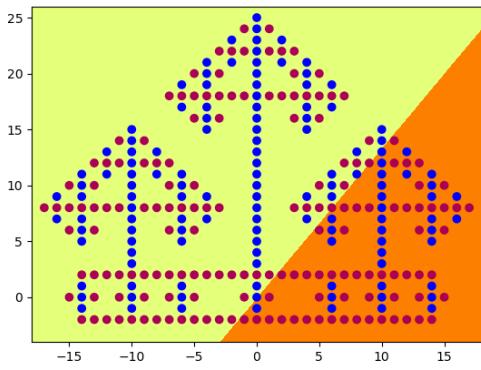
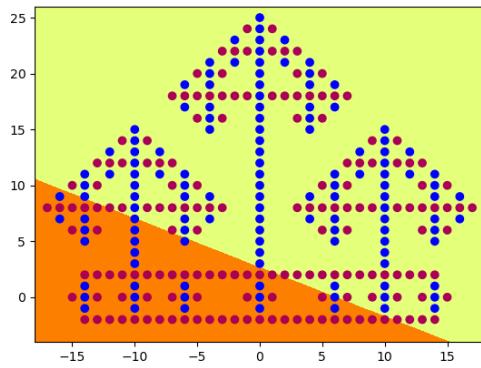
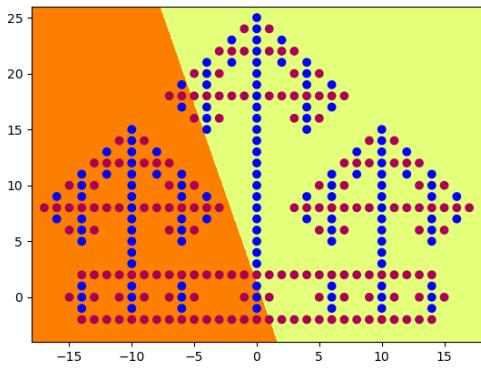
The picture of node 20 is:

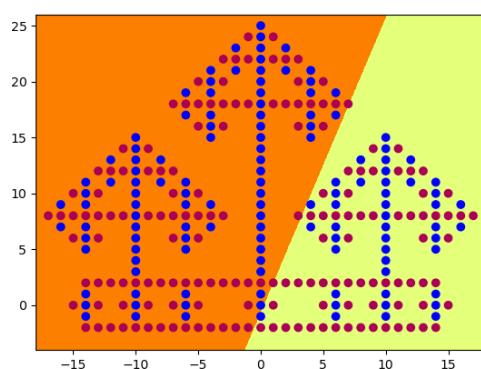
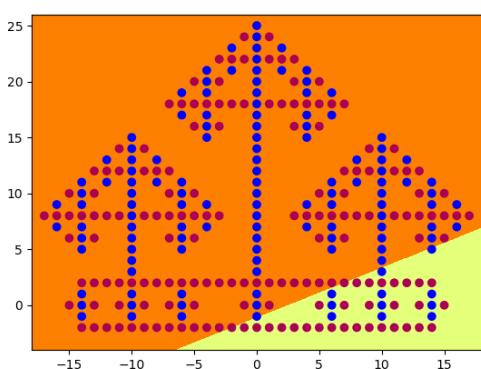
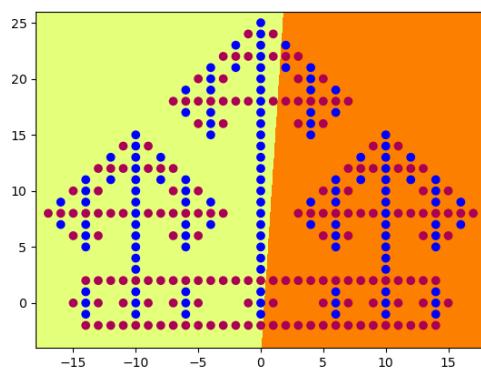
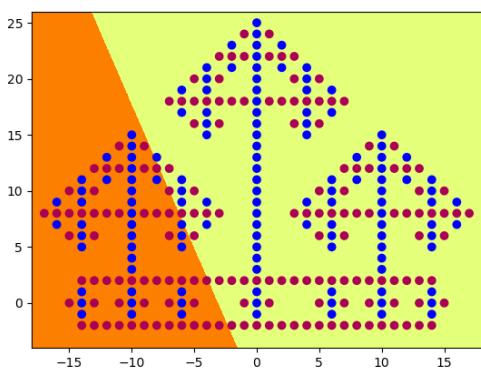
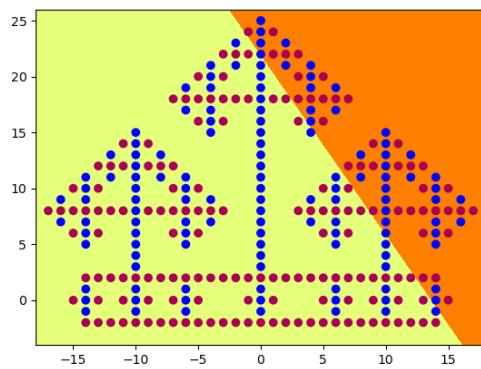
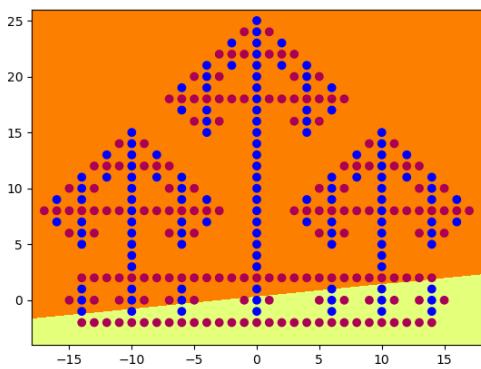
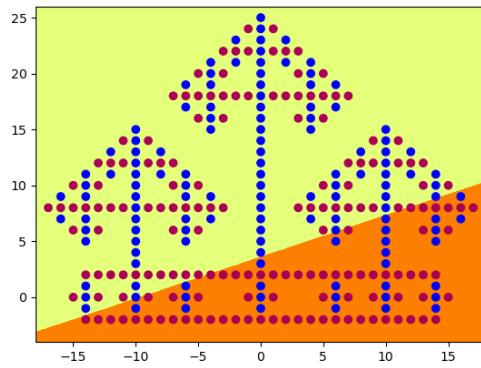
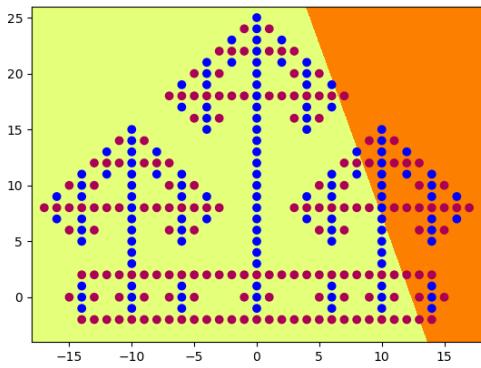


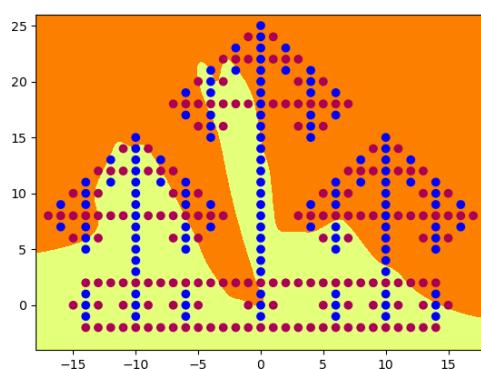
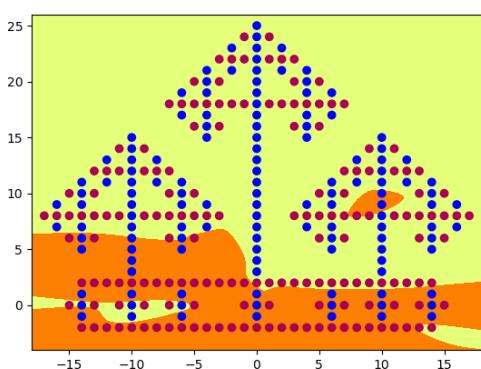
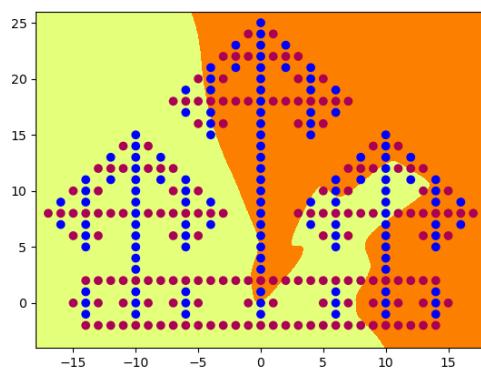
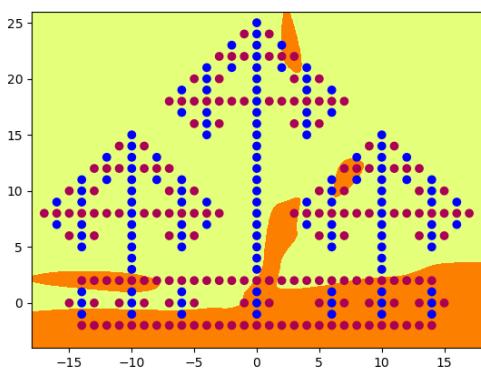
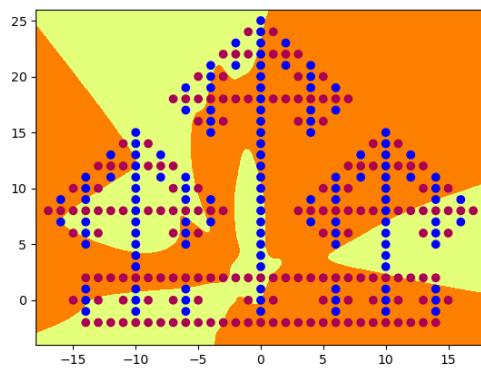
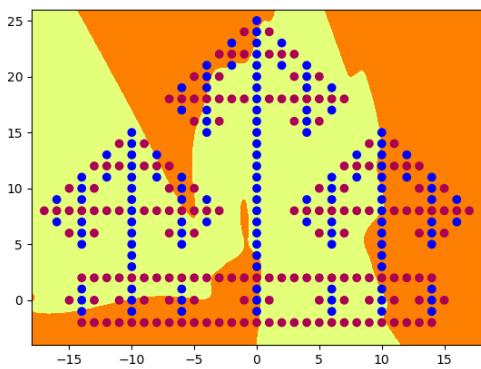
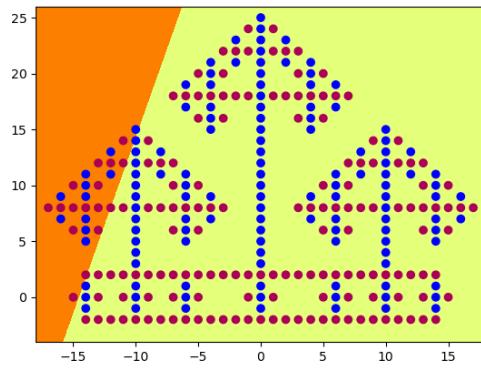
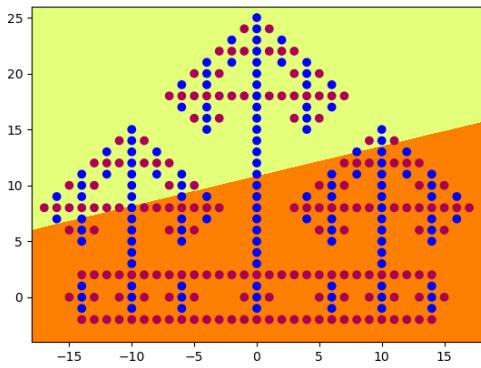
out_full4_20

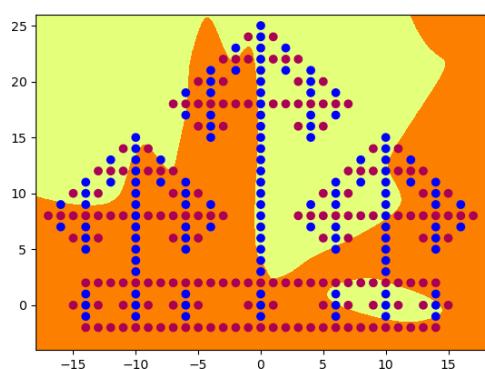
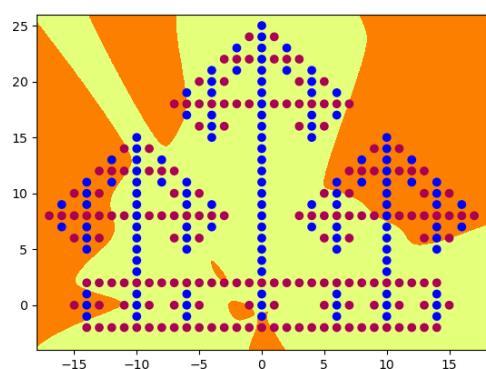
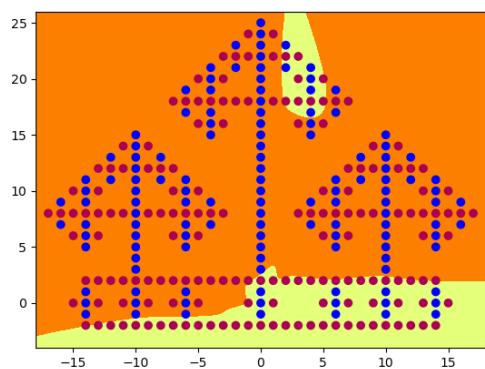
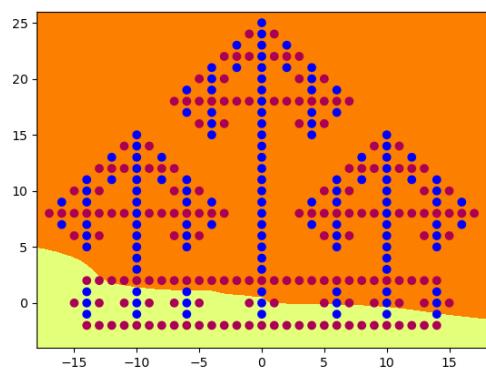
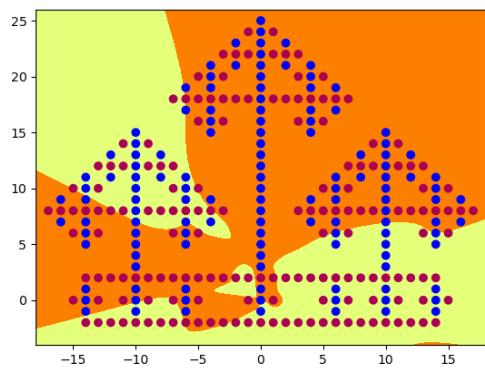
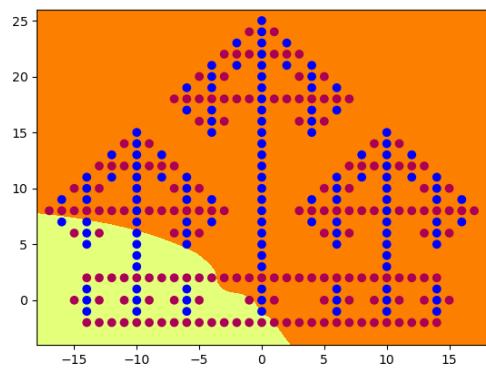
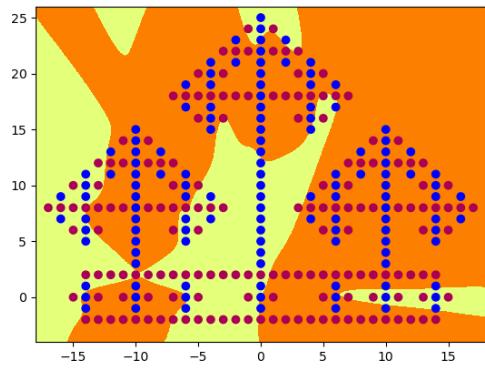
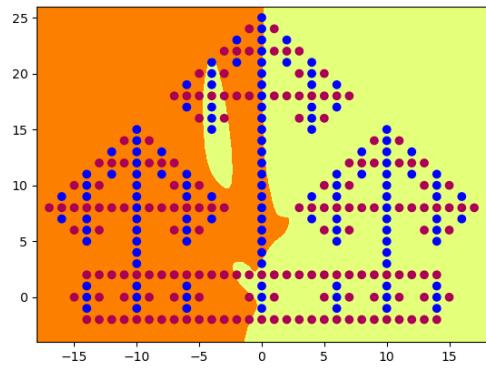
All hidden units:

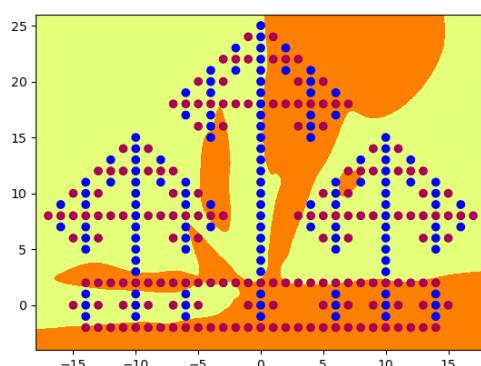
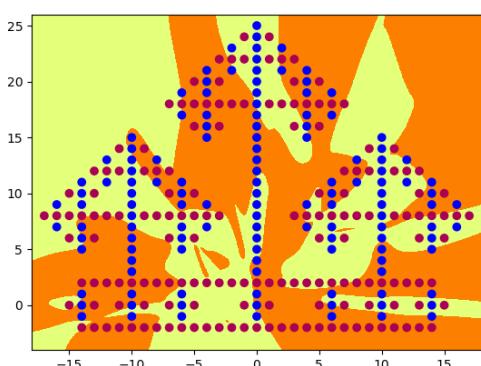
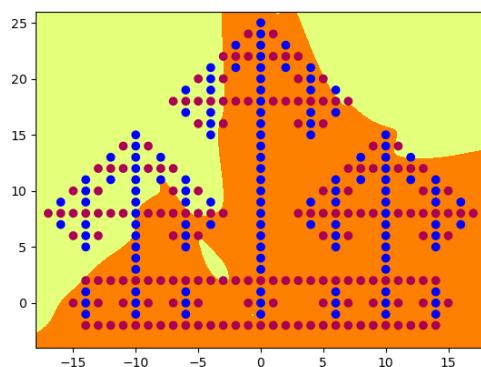
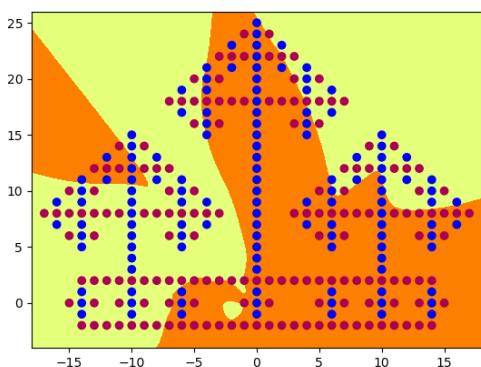
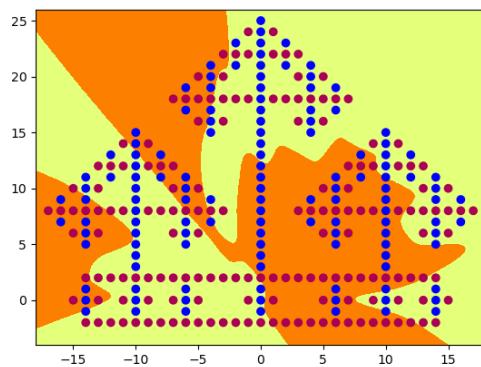
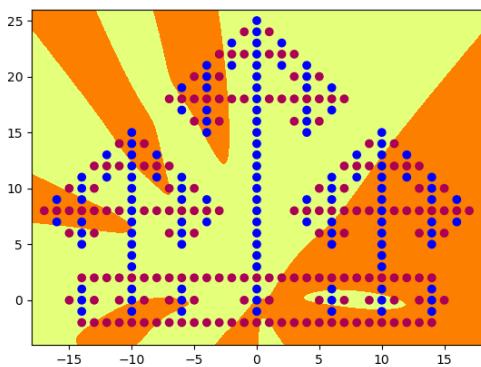
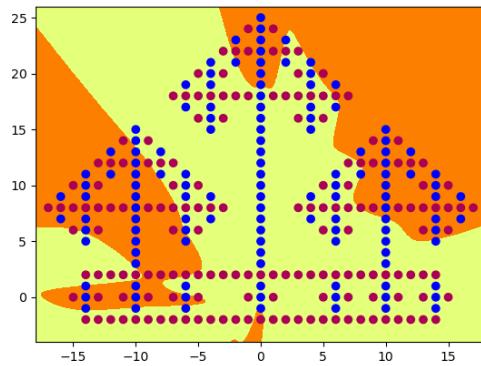
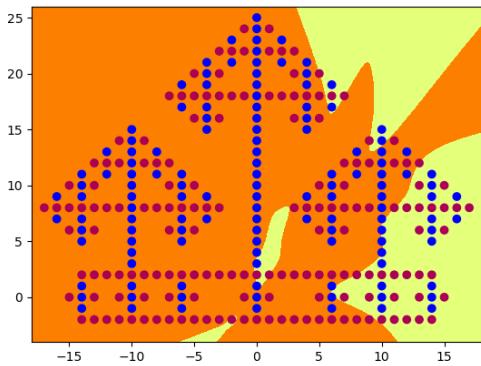


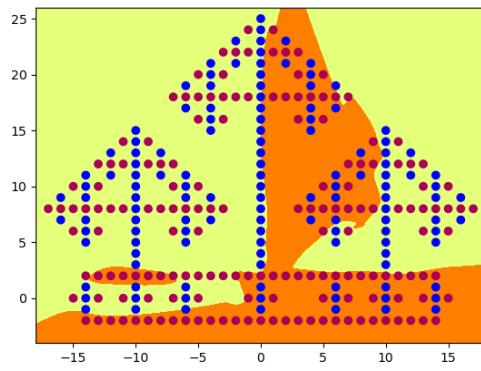
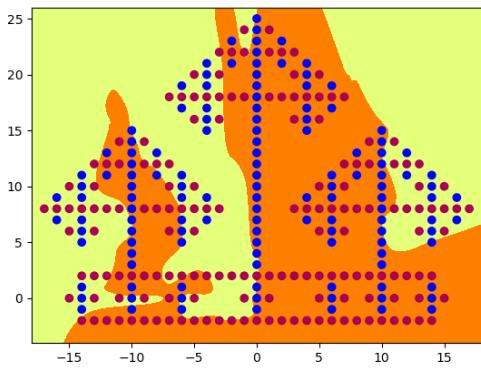
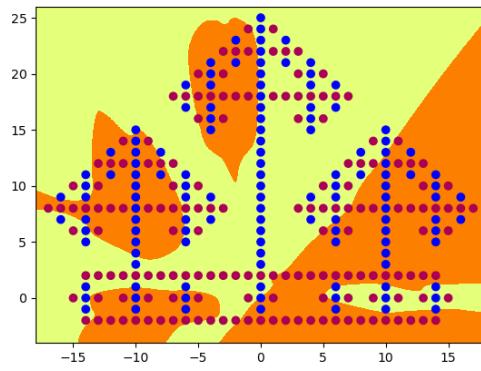
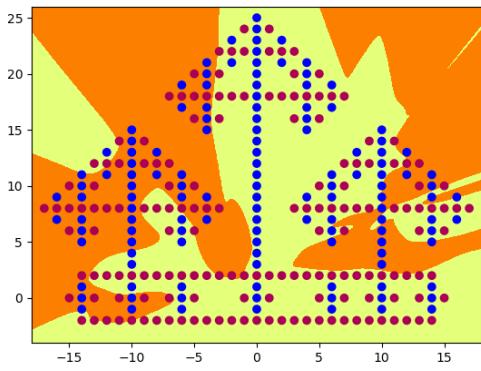
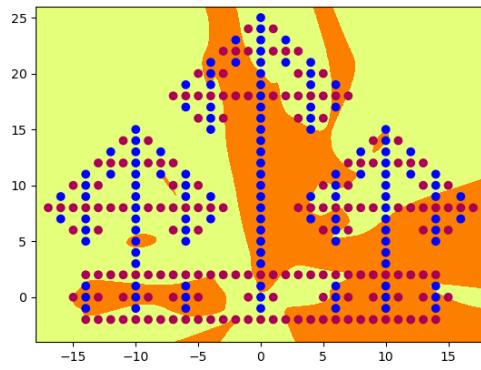
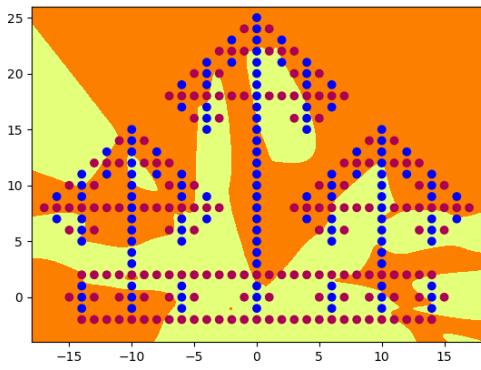
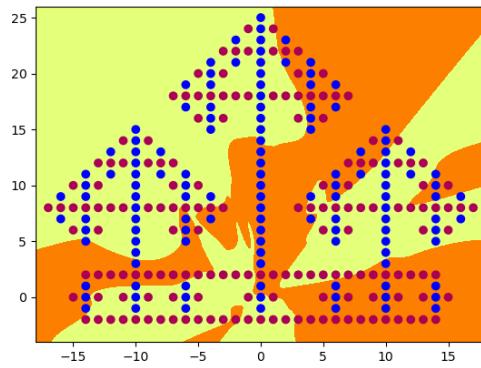
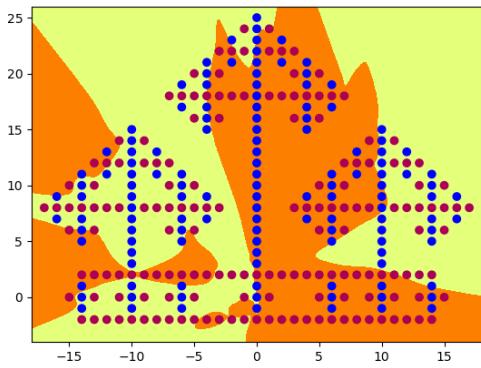


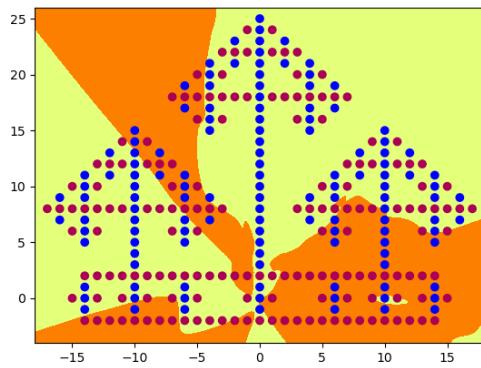
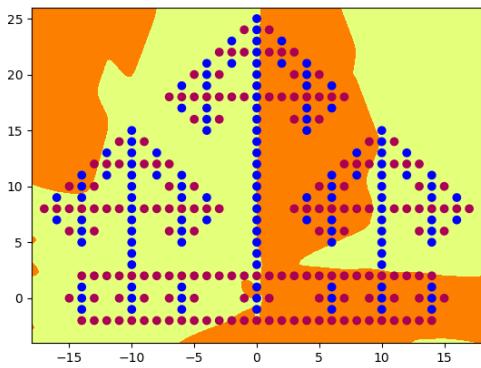
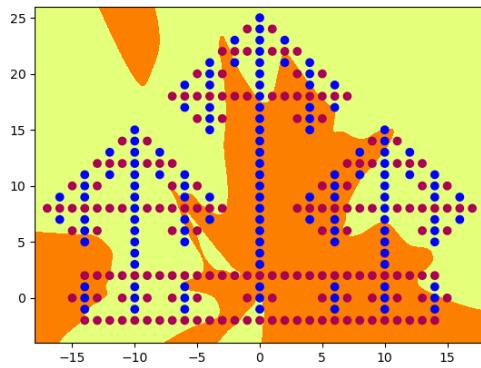
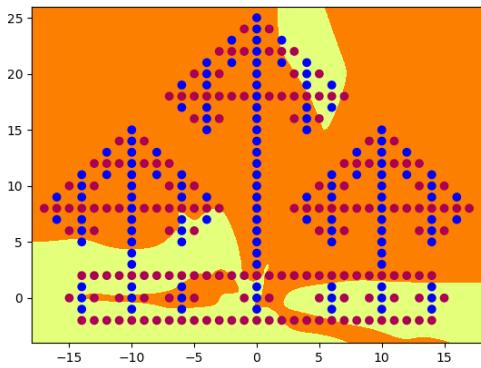
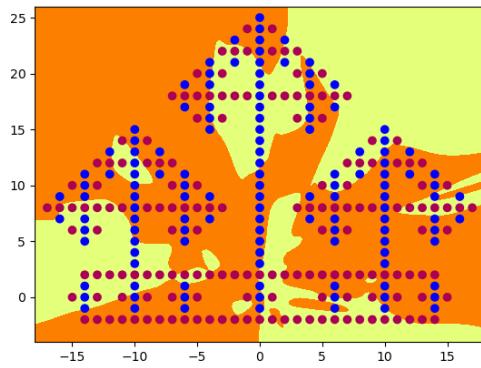
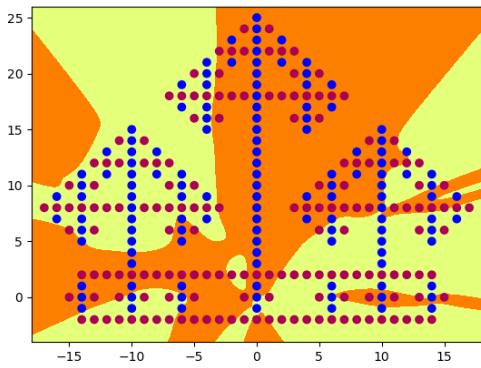
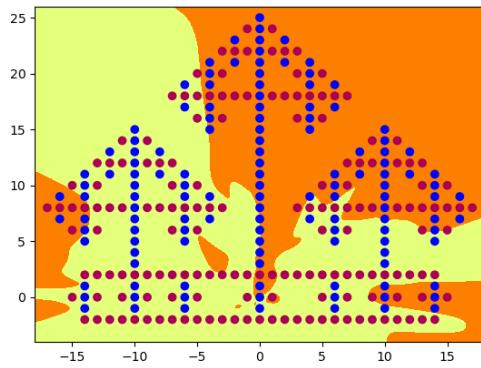
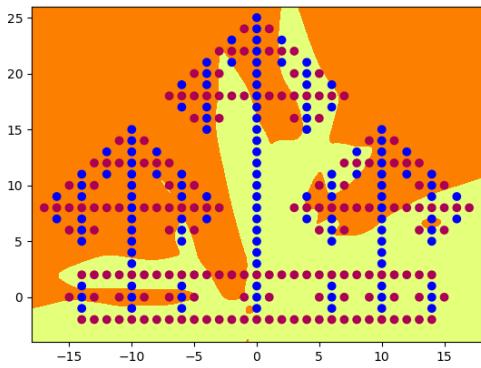


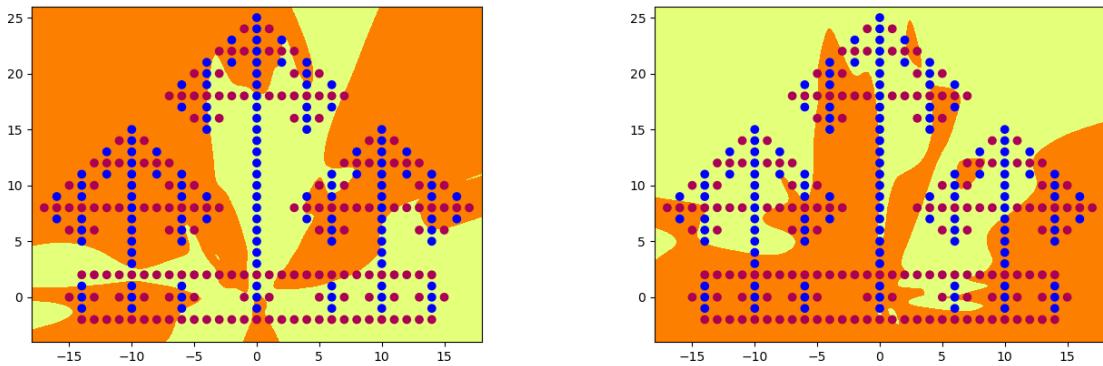












The total number of independent parameters:

$$(20 \times 2 + 20) + (20 \times 20 + 20) \times 2 + (20 \times 1 + 1) = 921$$

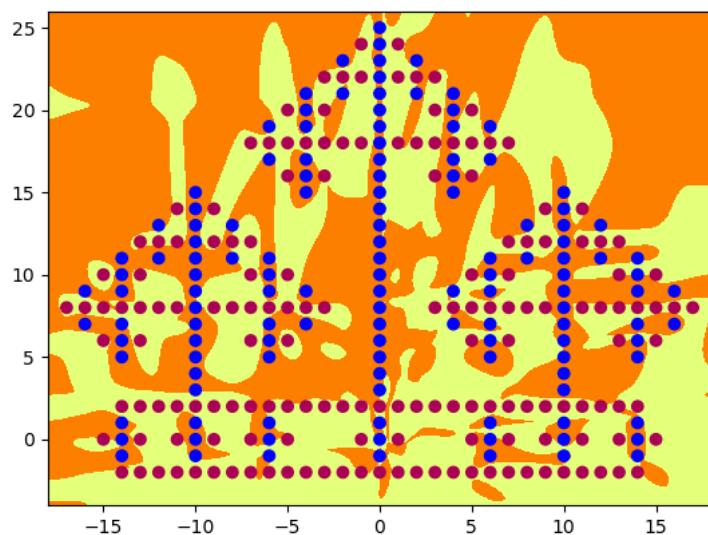
5. see code

6. DenseNet

Hidden nodes	lr	init	acc	ep
17	0.001	0.3	100%	24600
16	0.001	0.3	~98%	19400
15	0.001	0.3	~98%	36500
10	0.001	0.3	~90%	35000

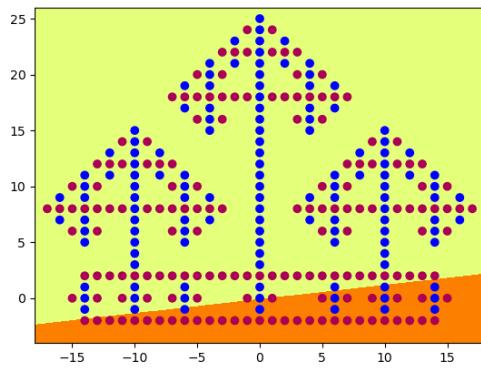
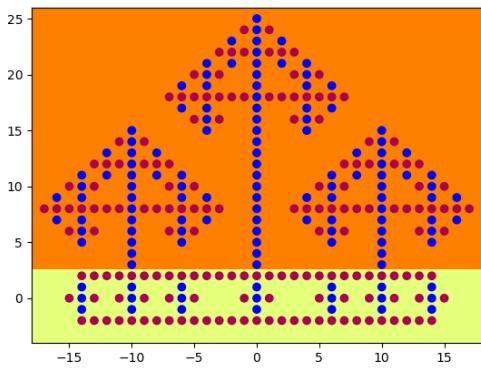
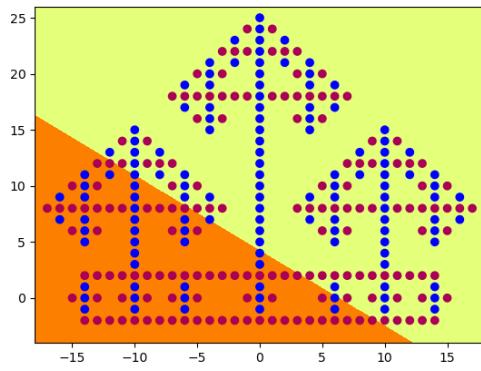
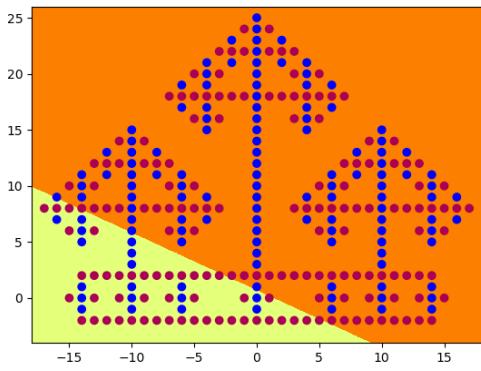
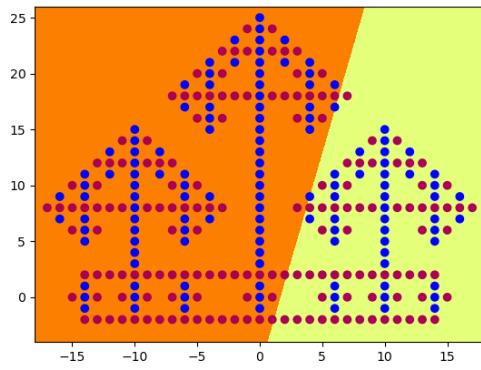
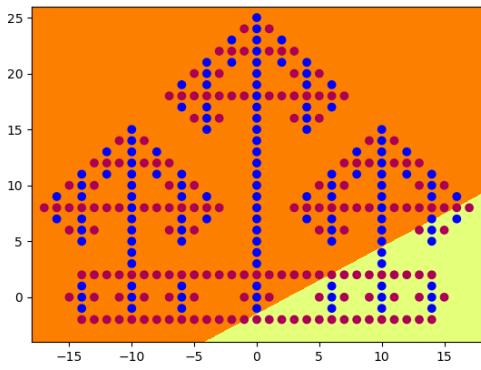
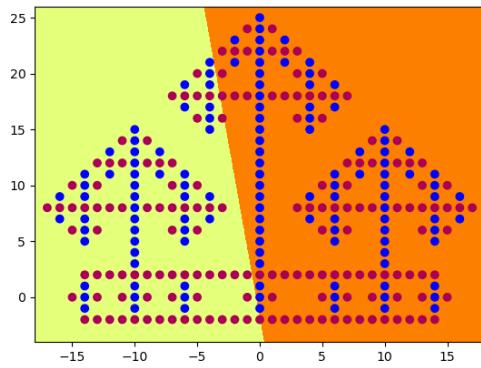
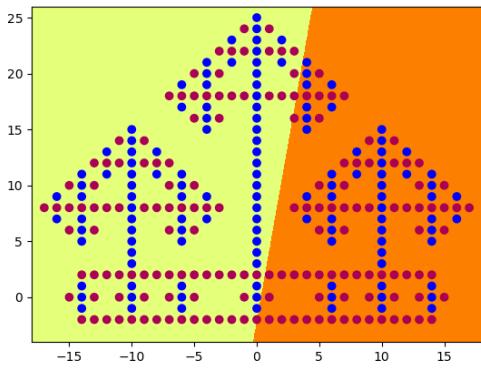
The number of hidden nodes close to the minimum required is 17

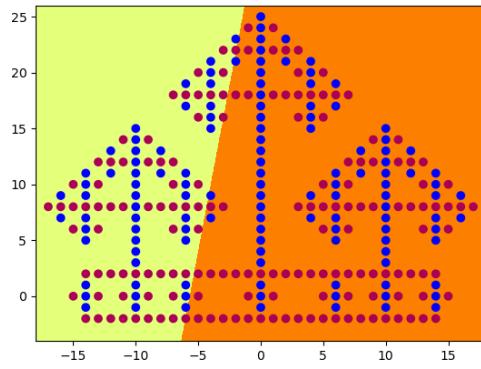
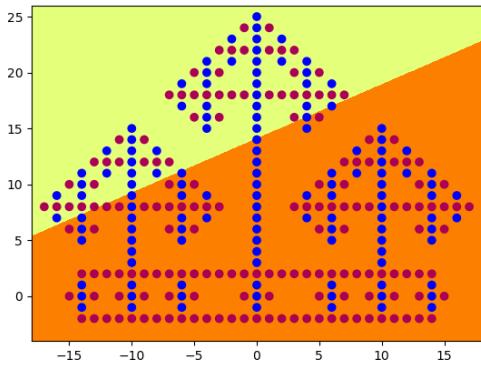
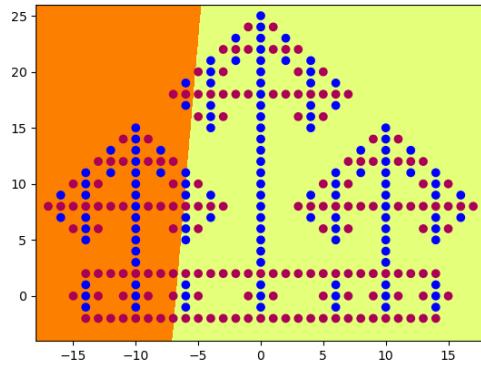
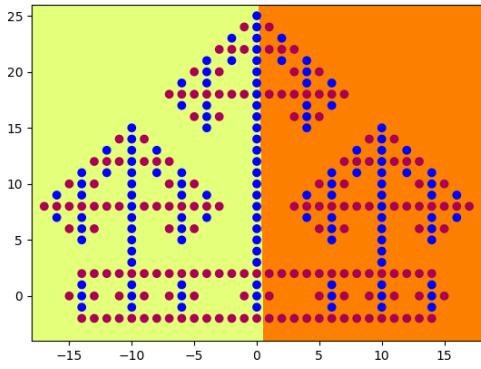
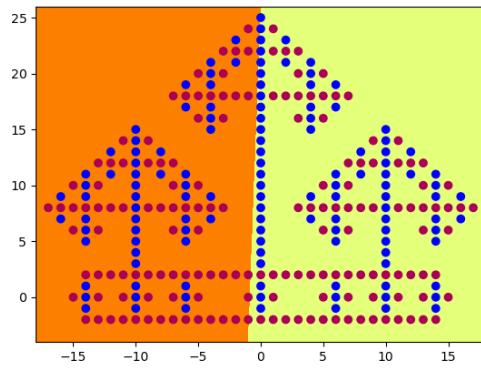
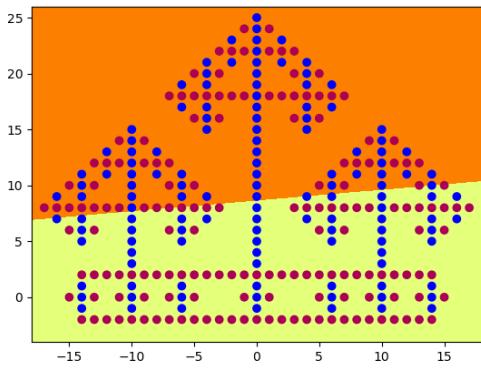
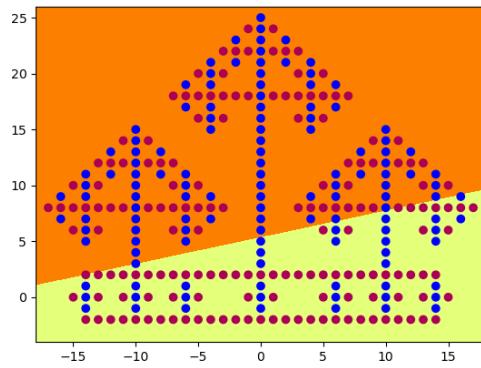
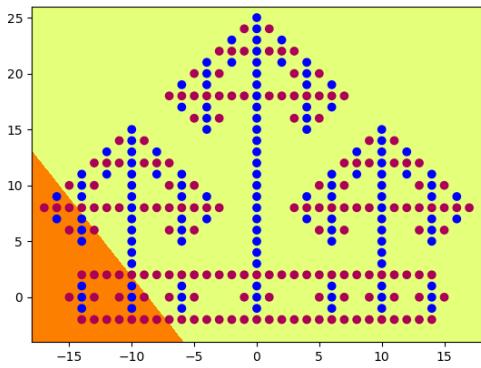
The picture of node 17 is:

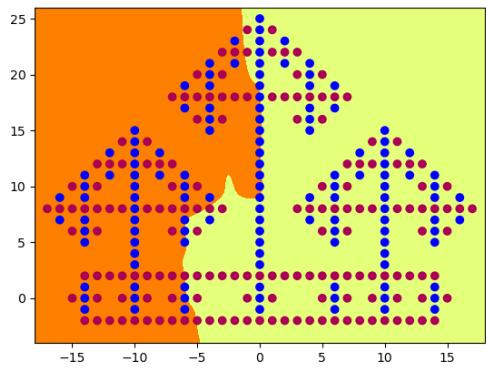
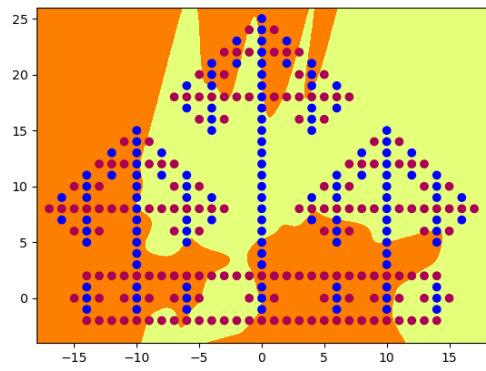
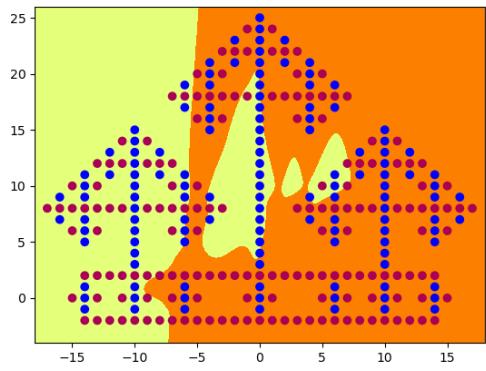
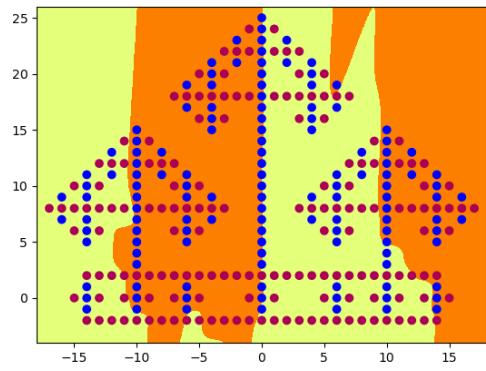
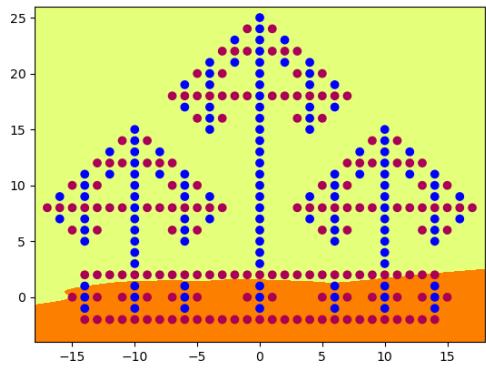
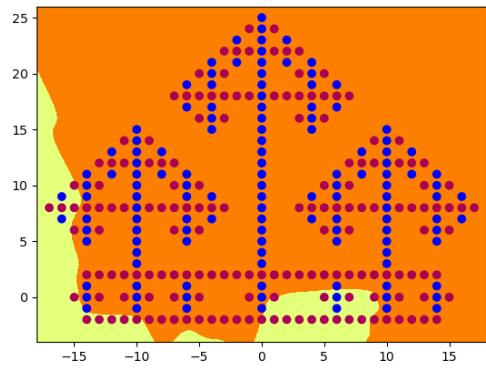
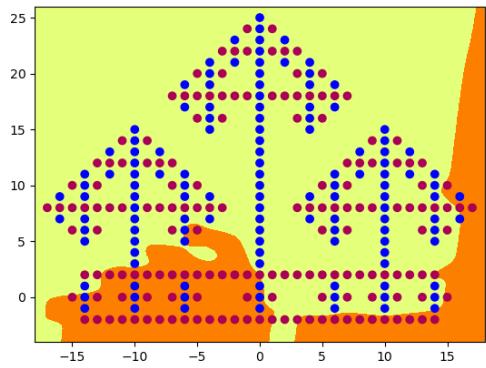
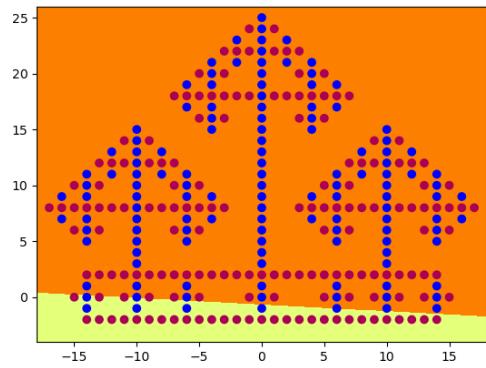


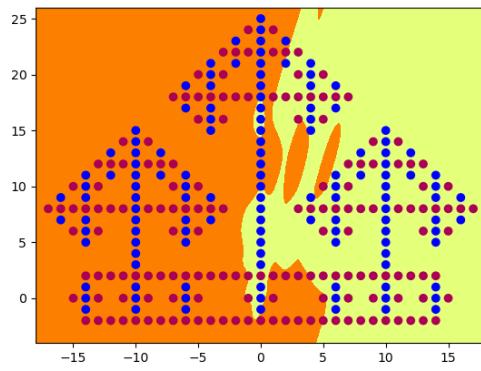
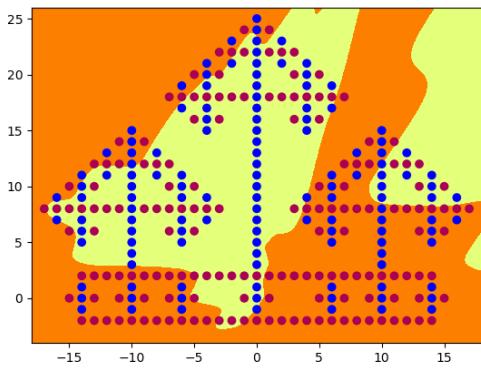
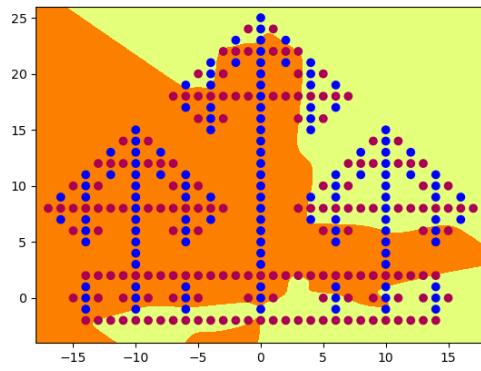
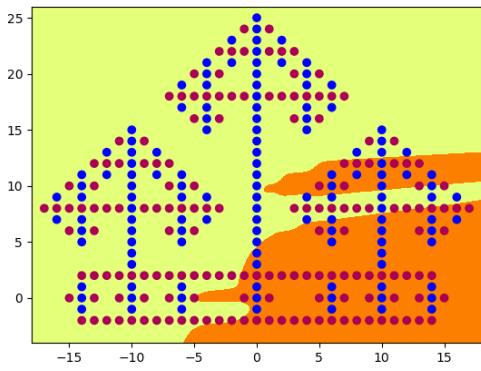
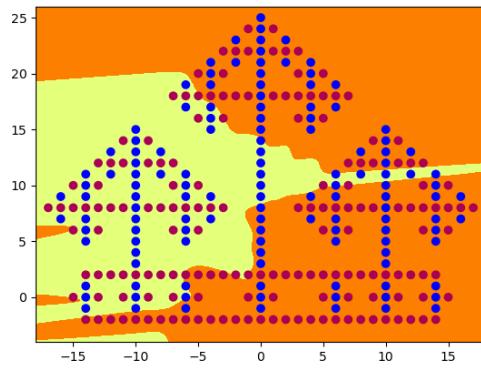
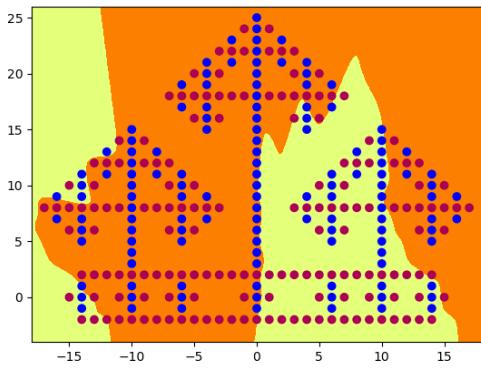
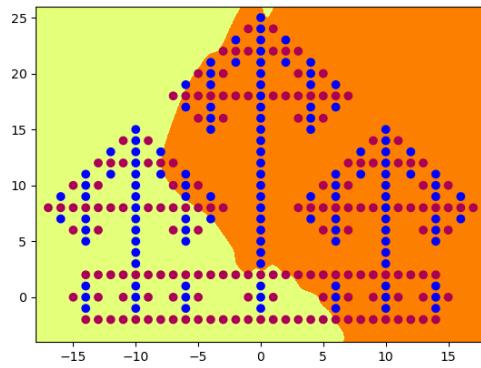
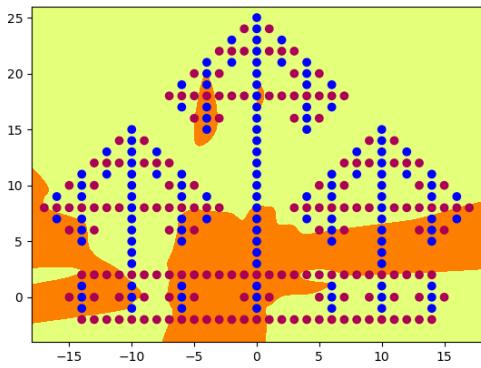
out_dense_17

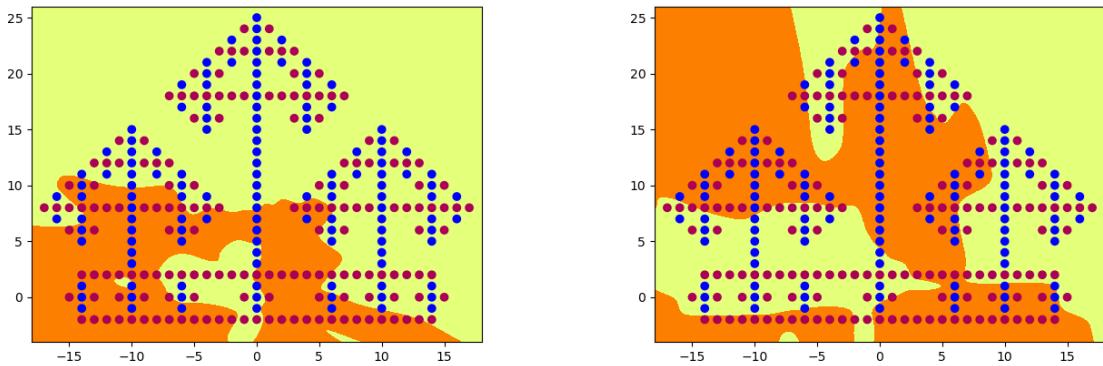
All hidden units:











The total number of independent parameters:

$$(17 \times 1 + 1) + (17 \times 2 + 17) + (17 \times 17 + 17) + 3 = 378$$

7.

a.

Model	Hidden Nodes	Independent Parameters	Epochs
Full3Net	23	645	18200
Full4Net	20	921	15800
DenseNet	17	378	24600

We can find that the first hidden layer of the neural network only generate liner outcomes, and seems like that the DenseNet is the best mode, because it have less number of hidden nodes and the small number of Independent Parameters

b.

The first layer of both three networks are all liner lines

The boundary complexity of second layer of all three networks increased, and further on dense net, because it owns more connections between each layer

c.

The three networks did not differ significantly in the fractal classification task. The same inputs yielded outputs with little difference

Part 2

Making a 12*11 matrix, assigning each dot a value

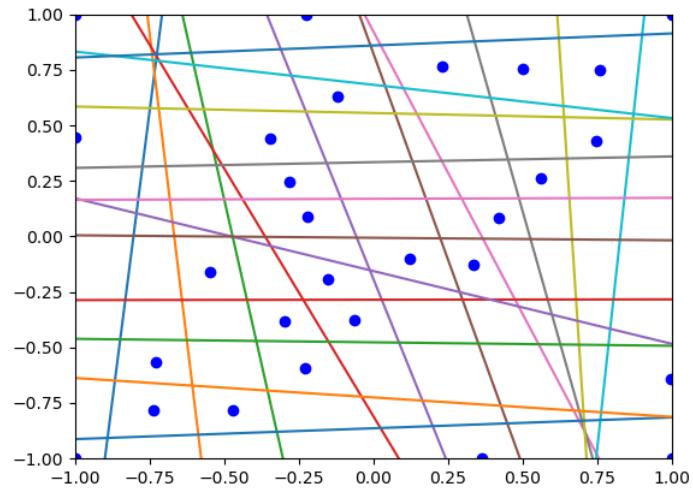
	A	B	C	D	E	F	G	H	I	J	K
T	1					8					7
Z		9	28								
S		10		27							6
4				12	26						
C			11		25						
9						24	23				
/					13			22			
8					14				21		
6	2				15					20	
U						16					
T							17	18	19		
Z	3					4					5

Each line represents a dot, then follow the rules below:

- Dots on the left of a vertical line is 0, right is 1
 - Dots on the top of a horizontal line is 0, below is 1

So we have this:

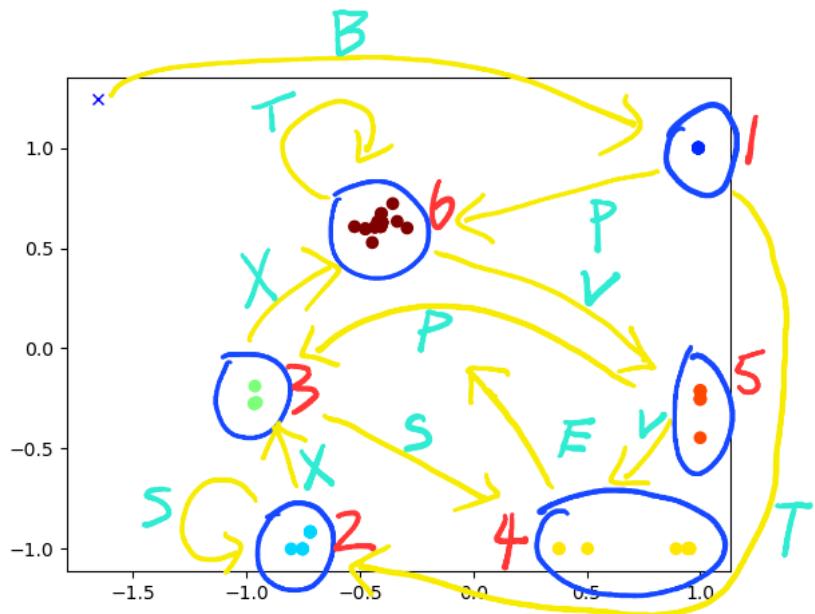
Getting the image produced:



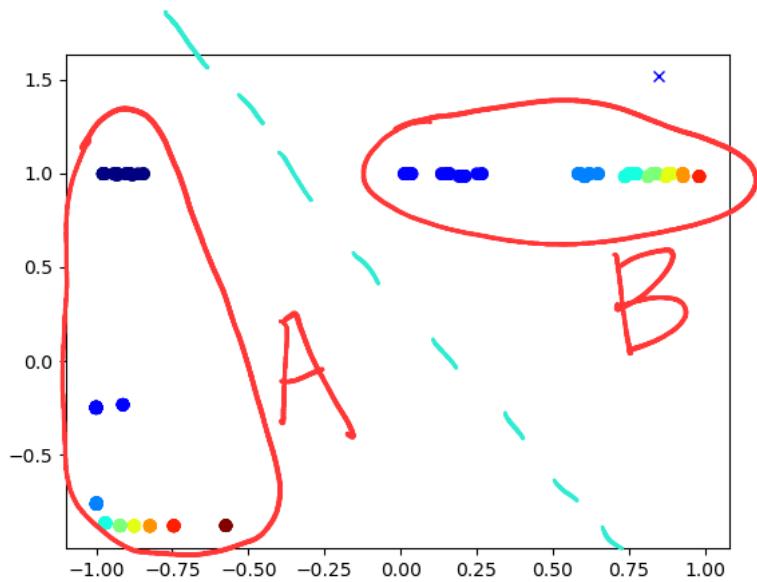
malta28.png

Part 3

1.



2.

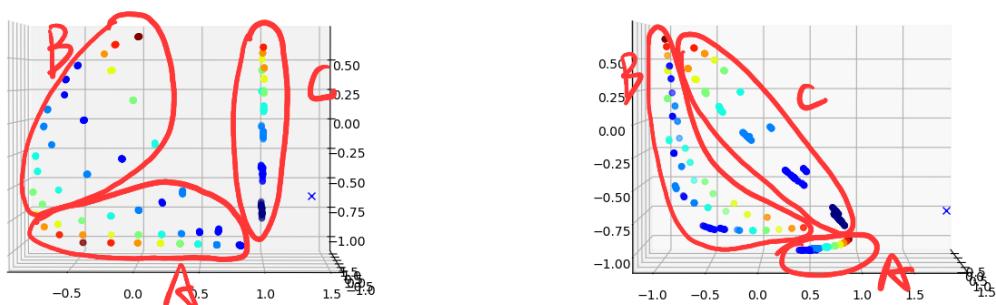


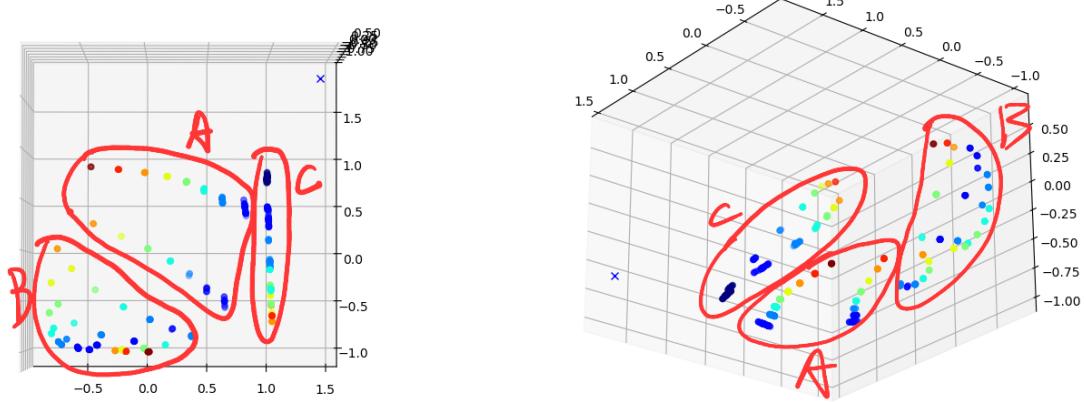
3.

The first A in every sequence is determinate, its activated on the bottom of A part in above figure. A certain model can predict next B according to previous value, its at right of B part show above.

The dark blue dots is the uncertain prediction, the network will decrease the probability of guess A but increase it of B after every A guess.

4.





5.

As figures show above, 3 hidden units activations correspond to A, B, C

Among the 3 letters, C has the least fluctuation, almost set on the plane of $z=1$,

so C is a certain class, so the network can predict C, second and every next B, and the first A.