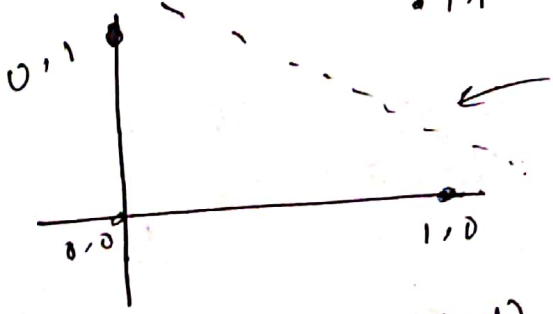
1820057.

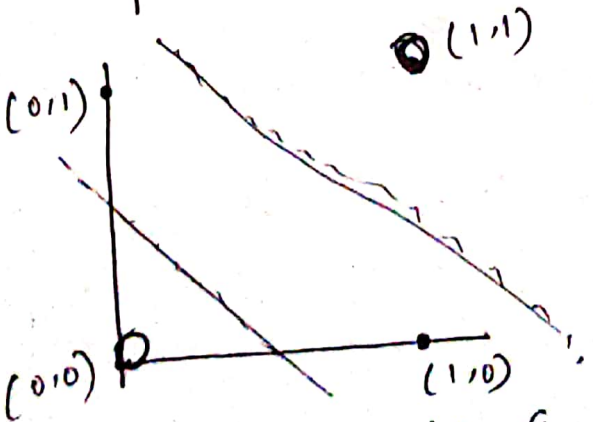
Final

Q-1 → Ans:- The training patterns A and B are not linearly separable and it cannot perfectly classify all of the training examples with any linear classifier. The datas will be linearly inseparable though the composition of linear layers are still linear.

In linear classifier, once the weights and biases are set, we can divide the data set into a region where the points are classified as positive and a region where it is negative. The training set is said to be linearly separable if it can choose a linear decision boundary that classifies all of the training sets. Like for AND gate:- ~~AND~~



For XOR =>



not linearly separable.

So, our patterns A and B, in the figure can be placed in any possible translation with wrap-around. Let's represent it with 16-dimensional binary vectors to distinguish the two patterns, A and B.

Therefore, if we shift the pattern right, then whatever falls off the right side reappears on the left. So, our classifier need to distinguish 16 examples of A and B.

Convexity refers to a geometric concept where a set S is convex if the line segment connecting two points lie within S. So, any weighted average of points in S must also lie within S.

So, By convexity if our classifier is to correctly classify all 16 instances of A, then it must also classify ~~all 16~~ the average of all 16 instances of A. Since 4 out of the 16 values are on the average of all instances, is simply the vectors $\begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ \vdots \\ 0.25 \end{bmatrix}$

Similarly, for it to classify all 16 instances of B, it must also classify the average of B. But the average is also $(0.25, 0.25 \cdots 0.25)$. Since, the dataset and the vector can't possibly classify both A and B, this dataset must not be linearly separable.

So, more generally, any linear classifier cannot detect a pattern of all possible translations of A and B since their average vector is the same.