points 2/5 ?	
il *	
In most of the tasks, which activation function is used in a model (not the activation function at the last layer) *	1/
ReLU	~
Sigmoid	
TanH	
	ReLU Sigmoid

!

SHA2205050

×	Why is momentum faster than normal gradient descent? *	0/1		
0	Momentum calculates the minima and directly jumps to the minima			
0	Momentum helps the gradient move towards the minima faster by averaging out to gradient values and hence smoothening the gradient path	he		
•	Momentum accelerates by automatically choosing the best learning rate and hence helps it go faster	×		
Correct answer				
•	Momentum helps the gradient move towards the minima faster by averaging out t gradient values and hence smoothening the gradient path	he		
×	Adam is a combination of which 2 optimisers? *	0/2		
	RMSProp			
	Stochastic Gradient Descent	×		
~	Momentum	✓		
	NAG			
Correct answer				
~	RMSProp			
	Momentum			

The sigmoid activation function is used for *	1/1
Multi Class Classification Task	
Regression task	
Binary Classification Task	✓
Name *	
Nag Arvind Gudiseva	

This form was created inside of Indian Institute of Technology Madras.

Google Forms