Answers provided through text replies	
A) About conduct of the course	
How to send questions to presenter ?	On screen on laptop there is a pane on the right hand side of the browser. Click on the question mark icon. It will open a small window at the bottom right hand side. Type question text and click on 'Send'
Questions from Slack channel also answered here?	No, slack channel is not monitored by the presenter during the lecture
how to access the slack channel	The link is given in the mail having details about joning the webinar
where can we find the link of articles to be read?	Will be shared on slack channel
Would we be getting the slides in advance to take notes on them as we learn?	Slides will be shared on slack channel after the lecture is over
Any reference material or books for reference? For these lectures?	All will be provided as part of slides which will be posted in the slack channel
Can we credit this course like the NPTEL course transfer credits?	No, that is not possible.
no audio pl	Please refresh your browser. We are monitoring and the audio is coming perfectly
sir audio is not clear	Please refresh browser
will the recording of class be available for later use	Please refer to FAQ section Q#8 https://iitgoa.ac.in/aishikshaai/faq.php
Sir today i missed the class how can I see today's lecture	please refer to FAQ section Q#8 https://iitgoa.ac.in/aishikshaai/faq.php
Hi, Where do/can the participants refer to the questions that the presenter is discussing?	Questions are sent to Presenter by participants, they are not available for other participants to view
Can I join from mobile? It is very slow in PC with frequent disconnection ...	Yes you can
Is there a possibility to get a transcript of the QnA discussed by the presented?	Sorry, this is not possible.
Sir, could you provide you contact no ?	Sorry, this is not possible.
will we receive a grade sheet and certificate after course completion?	please refer to FAQ section https://iitgoa.ac.in/aishikshaai/faq.php You need to use a stable network connection. We would request you to pay attention to the lecture, such questions can wait
B) About doubts asked during the course	
What is GOPS?	Giga (10^9) operations
Would we get video of the lecture after it gets over, as its getting interrupted multiple number of times..	please refer to FAQ section Q#8 https://iitgoa.ac.in/aishikshaai/faq.php You need to use a stable network connection.

since AI needs a lot of data, does it access the higher levels of memory?	Memory is first stored in RAM and then it gets via the hierarchy like CAche to reach the cores. But it ways gets allocated in RAM first
Reducing feature size: does it means we need to reduce feature detection variables in each step of computation, without compromising accuracy.	Feature size, typically measured in nanometers, is thickness of transistor which is inside a chip. Smaller the feature size, more the transisitors that can be accomodated in a chip making it more complex. This is much different from the feature detection in AI alogrithms, as would have realised
What is SIMT execution	SIMD, not SIMT, refers to Single Instruction Multiple Data. A feature that instruction of modern processor supports. See https://en.wikipedia.org/wiki/Single_instruction,_multiple_data for some details
How does using tensor cores affect accuracy? Since part of the multiplication is done in FP16, wouldn't this affect the model accuracy?	Now a days there are trainings done even in int 8 . The iterative algorithms used in deep Learning . Can handle lower precisions very well. There is well defined mathematics behind these.
Does one need specialized Hardware to do Sparse Matrix Multiplication?	You dont need specialized hardware. But yes A100 has special hardware which exposes sparcity to improve your training time by many folds. So using a specialized core meant to expose sparsity can help you reduce overall time.
Does the programmer need to specify which computations can use tensor cores/mixed precision? Or is this decision made by libraries/compiler?	Primarily NVIDIA libraries already do that behind the scenes. Users can enable it via flags from choice of your frameworks like PyTorch and TensorFlow
How to detect the number of GPU nodes at a particular node?	If you have installed a NVIDIA driver. You can use command nvidia-smi . We will be showing demos in subsequent lectures
With reference to the memory hierarchy slide, how high are the overheads in copying data from CPU RAM to GPU RAM? Also, how does this change with an integrated GPU where RAM is shared between CPU and GPU?	You are right there is initial transfer required. But onc eit is in GPU and if the whole training which is iterative in nature will not involve any data transfer. Integgrated GPU and CPU does share the RAM and newer architectures will evolve to solve some challenges related to cache which is also shared. GPU currently have higher bandwidth in range of few TB/s which CPU RAM is still in GB/sec . So there is a hige difference.
This may be a stupid question, but, on a PC, if there are multiple CPU cores, and one GPU unit, can both be utilized simultaneously?	It is not a stupid question. There are frameworks that allow mechanism for such use