* What are the steps in training of a neural network? Discuss with your peers. Pay attention to the terminologies used when describing the steps (slide 7).

* What is vanishing gradient in deep neural networks? Why does this happen? (slide 8)

* What is exploding gradient in deep neural networks? Why does this happen? (slide 8)

* How to prevent vanishing and exploding gradients? (slide 28)

* What are non-saturating activation functions? (slide 9)

* What happens when we add normatisation layer before or after each hidden layer? (slide 10)

* What is gradient clipping? (slide 11)

* What are initialisation strategies? (slide 12)

* What are the different choices of optimisers? Compare them in terms of convergence speed and quality. (slide 16-17)

* What is the idea of early stopping? How to do this in Keras? (slide 18)

* What is learning rate and what factors are dependent on the learning rate? (slide 20)

* What is learning rate scheduling? Explain piecewise constant, power scheduling and exponential scheduling and how to implement them. (slide 22)

* What is cyclical learning? (slide 23)