1. What is the concept of cyclical momentum?

2. What callback keeps track of hyperparameter values (along with other data) during training?

3. In the color dim plot, what does one column of pixels represent?

4. In color dim, what does "poor teaching" look like? What is the reason for this?

5. Does a batch normalization layer have any trainable parameters?

6. In batch normalization during preparation, what statistics are used to normalize? What about during the validation process?

7. Why do batch normalization layers help models generalize better?

8.Explain between MAX POOLING and AVERAGE POOLING is number eight.

9. What is the purpose of the POOLING LAYER?

10. Why do we end up with Completely CONNECTED LAYERS?

11. What do you mean by PARAMETERS?

12. What formulas are used to measure these PARAMETERS?