Q-1 – 4

Q-2 – Faster / less mem usage ⇒ accuracy ↑

Q-3 why?

Q-4 → 1.05 –

Q-5 → 0-5s

Q-6 → No –

Q-7 – No –

Q-8 – No –

9 – [(4)(7)(7)]

Sum = 18

10 – ~~False~~ True

11 – ~~5~~ grad is 5 – because that's the slope –

12 – 2(y_pred-y) is gradient of ~~loss~~ mean square error

grad[h<0] = 0 ⇒ ~~sigmoid~~ relu

13 → epochs → ? –

64 → 1000 → $\frac{1000}{64} = 15.22$

32 epochs $\frac{1000}{64} \times \frac{500 \times 64}{1000} = 32$

Q. 15 → No → this is random number – so No –

16 → No –