CV ASSIGNMENT 3

Q1) after stride2 con, activation map reduces by half and hence filter size is doubled.

Q2) larger kernel attracts features better and later small kernels would reduce trainable parameters.

Q4) we mention the event and define training loop in callback.

Q5) since output would be binary (0 or 1), having above zero activation would be problematic and arbitrary value which will fail the classification task.

Q6) benefits include better neuron learning and less overfitting/underfitting. Drawbacks include slower computation and hardware capability issues.

Q7) starting with very high learning rate could lead to overfitting.

Q8) faster convergence and could address vanishing gradient problem.

Q9) to prevent overfitting/underfitting and to actually be close to minimum point.