Part 1: share your thoughts on the following questions:

1. **You are a data scientist in ABC company. Your task is to build a model using deep learning. It is about classification of 5 type of customers. your model will decide whether a customer is type 1 or type 2 ... or type 5. Each customer data is composed of 100 features that is already captured. we can supply as many numbers of customer information you want (e.g. 100, 10000 or more data samples). What is your strategy on requesting the number of data samples? is more and more data always good?**

   Starting with a Small Set of Data, Doing Training and Testing and Gradually Increasing the Dataset Comparing the Training and Testing Results Will Be The Most Optimum Process.

2. **You build your model, what is your strategy on selecting the number of epochs?**

   The Number of Epochs Initially Should be Set on a Higher Value and Keep Checking the Validation and Training Error, whether it is Decreasing or Not. Terminate the Training of Model when It Reaches Convergence.

3. **How you train a model with unbalanced data? for example, you are given 100000 data sample regarding to Mastercard transactions where .01% of them are flagged to be fraud. What is your strategy for training your model to detect the fraud?**

   Use the Strategy of Synthetic sampling (SMOTE)-In this Technique we Synthetically try to Manufacture Observations of unbalanced Classes which are like the existing Classes using nearest neighbour classification.

What are the available optimizers in Keras and what is their difference?

SGD – It Includes support for momentum, learning rate decay, and Nesterov momentum.
Adagrad - It is an optimizer with parameter-specific learning rates, which are adapted relative to how frequently a parameter gets updated during training.
Adadelta - It is a more robust extension of Adagrad that adapts learning rates based on a moving window of gradient updates, instead of accumulating all past gradients.
RMSprop - Divides the gradient by a running average of its recent magnitude
Adam - It is a Method for Stochastic Optimization
AdaMax - It is a variant of Adam based on the infinity norm.
Nadam - It is essentially RMSprop with momentum, Nadam is RMSprop with Nesterov momentum.