1. Explain convolutional neural network, and how does it work?

Ans: A convolutional neural network (CNN) is a type of artificial neural network used in image recognition and processing that is specifically designed to process pixel data. CNNs use a combination of convolutional layers, pooling layers, and fully connected layers to identify patterns in the data. Convolutional layers use a filter to scan across the image and detect specific features, such as lines, corners, and edges. Pooling layers reduce the dimensionality of the image by combining pixels together. The fully connected layers then combine these features to identify objects and classify images.

2. How does refactoring parts of your neural network definition favor you?

Ans: Refactoring parts of your neural network definition can make the code easier to understand, debug, and maintain. It also allows you to reuse code, making it easier to experiment with different model architectures. By making the code more modular, it is also easier to add or remove layers, adjust hyperparameters, or even extend the network with additional features. Refactoring can also help make your code more efficient, as it allows you to reuse code and optimize calculations.

3. What does it mean to flatten? Is it necessary to include it in the MNIST CNN? What is the reason for this?

Ans: To flatten is to transform a multi-dimensional array into a single dimensional array. It is necessary to include it in the MNIST CNN to transform the output of the convolutional layers into a single dimensional array before the dense layer. This is because the dense layer requires a one-dimensional array as input.

4. What exactly does NCHW stand for?

Ans: NCHW stands for "Nested Channel-wise High-dimensional Weighting" and is a type of deep learning architecture. It is used to improve the accuracy of predictive models by utilizing a multi-level hierarchical structure to better understand complex data.

5. Why are there 7*7*(1168-16) multiplications in the MNIST CNN's third layer?

Ans: The 7x7x(1168-16) multiplications in the MNIST CNN's third layer is the result of the number of parameters in the layer. The first two layers of the MNIST CNN have 1168 input neurons and 16 convolutional filters. Each of these filters has a 7x7 parameter matrix, resulting in 1168 x 16 = 7 x 7 x (1168 - 16) multiplications.

6. Explain definition of receptive field?

Ans: A receptive field is a specific region in the visual field of an individual (or a neuron) that responds most strongly to a particular stimulus. It is the area within which a stimulus produces a detectable response in the neuron or individual, and is usually defined by the size and shape of the stimulus that produces the greatest response. The receptive field is often used to analyze the visual processing of the brain, as well as to understand the physiology of the visual system. Receptive fields can also be used to study the effects of visual stimuli on behavior, such as how a person responds to an object in their visual field.

7. What is the scale of an activation's receptive field after two stride-2 convolutions? What is the reason for this?

Ans: After two stride-2 convolutions, the scale of the receptive field is 4. This is because stride-2 convolutions reduce the size of the receptive field by half each time, so two stride-2 convolutions reduces the size of the receptive field by a factor of 4.

8. What is the tensor representation of a color image?

Ans: A tensor representation of a color image is a three-dimensional array that contains the values of each pixel in the image. The first two dimensions represent the height and width of the image, while the third dimension represents the color channels (e.g. Red, Green, Blue).

9. How does a color input interact with a convolution?

Ans: A color input can interact with a convolution by using a color filter. The convolution will apply the filter to the input image, transforming the input into a new image with the same dimensions as the original but with the colors altered according to the filter. The convolution will essentially process the input image and create a new image with the desired color effects.

I