DL- Lab3-Answers

Q1) The signal transitions from 1 to 0.5 after index 6 and transitions back to 1 after 9. Those numbers are highlighted in the resulting convoluted signal. In a similar way edges can be detected using the difference of strengths of adjacent pixels in an image. However, in most cases 2D filters are used.

Q3) Validation error increased because of overfitting. Overfitting can be avoided with use of following techniques.

- Early stopping
- Simplifying the model
- Data augmentation
- Dropouts
- Regularization

There are 2 main optimization algorithms used for training ML models namely.

- Stochastic Gradient Descent (SGD)
- Batch Gradient Descent (BGD)

In BGD, gradient of the cost function is calculated using the entire dataset. Model params are updated once per epoch. This is more stable and accurate but computationally expensive (especially for large data sets since memorizing and performing calculations on the whole data set is required) compared to other approaches.

In SGD a single example is used to compute gradient and update the model params. This is the fastest of all approaches, however the randomness and noise caused by individual samples can lead to oscillations around the global minimum making the convergence less smooth.

Mini Batch Stochastic Gradient Descent (MBSGD) is a combination of those two. It uses a small batch (hence the name mini batch) for each parameter update. This is computationally efficient compared do BGD (since lesser data is used for computation) but also has the benefit of noise induced faster convergence times similar to SGD.