Assignment 01: Deep Learning Essentials

Instructors: Sir Irfan Malik , Sir Sheraz Nasser

Student Name: Shanza Khan

Question 1: Exploring Neural Network Architectures

1. Convolutional Neural Networks (CNNs)

- CNNs are primarily used for image-related tasks, employing convolutional layers to extract spatial features.
- Filters (kernels) help detect edges, textures, and patterns, reducing parameter count compared to fully connected networks.
- Pooling layers further reduce dimensions, enhancing computational efficiency.
- **Difference from Fully Connected Networks**: CNNs leverage spatial correlation, improving efficiency, whereas fully connected networks treat each input independently.
- **Real-world Application**: Image classification (e.g., self-driving cars, medical imaging).

2. Recurrent Neural Networks (RNNs)

- RNNs process sequential data, maintaining a hidden state to carry information across time steps.
- Loops in the network allow memory retention, making them suitable for context-dependent tasks.
- **Difference from Fully Connected Networks**: Fully connected networks handle inputs independently, while RNNs maintain context.
- **Real-world Application**: Speech recognition (e.g., Siri, Google Assistant).

Question 2: Beyond Sigmoid - Activation Functions in Neural Networks

1. Rectified Linear Unit (ReLU)

- **Formula**:
- **How it Works**: If the input is positive, it remains unchanged; if negative, it becomes zero.
- **Advantages**:
    - Mitigates vanishing gradient issues seen in sigmoid functions.
    - Computationally efficient.
- **Common Usage**: Deep neural networks due to its simplicity and effectiveness.

2. Hyperbolic Tangent (Tanh)

- **Formula**:
- **How it Works**: Outputs values between -1 and 1, centering data around zero.
- **Advantages**:
    - Symmetric around zero, improving convergence speed.
    - Helps in learning both positive and negative inputs.
- **Common Usage**: Frequently used in RNNs as it retains better gradient flow than sigmoid.

Question 3: Exploring Loss Functions

1. Mean Squared Error (MSE)

- **Formula**:
- **Usage**: Applied in regression problems with continuous outputs.
- **Why Suitable**: Penalizes large errors more, ensuring smoother gradient updates.

2. Cross-Entropy Loss (for Multi-Class Classification)

- **Formula**:
- **Usage**: Applied in multi-class classification problems.
- **Why Suitable**: Optimizes softmax outputs, ensuring accurate probability distributions.

Bonus Activity: Interactive Practice

Using **TensorFlow Playground**, the following observations were made:

- Increasing hidden layers improves accuracy but increases training time.
- ReLU enables faster learning, whereas Sigmoid struggles with vanishing gradients.
- Overfitting occurs with excessive neurons, while too few lead to underfitting.

These findings highlight the significance of choosing appropriate activation functions and network architectures.

References

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
2. LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning*. *Nature*, 521(7553), 436-444.
3. TensorFlow Playground: https://playground.tensorflow.org/