LOSS FUNCTIONS IN DEEP LEARNING

Loss functions play a fundamental role in the training process of machine learning models, particularly in deep learning. They serve as a measure of how well the model is performing on the given task, guiding the optimization process to update model parameters in a direction that minimizes the error between predicted and actual values

Importance of Loss Functions:

1. **Evaluation of Model Performance**: Loss functions provide a quantitative measure of how well the model is performing on the training data. By assessing the discrepancy between predicted and actual values, they offer insights into the model's accuracy and effectiveness.
2. **Optimization Guidance**: During the training phase, the primary objective is to minimize the loss function. Optimization algorithms such as gradient descent use the gradient of the loss function with respect to model parameters to iteratively update these parameters, aiming to reduce the loss and improve model performance.

Types of Loss Functions:

- **Regression Loss Functions:**

Regression tasks involve predicting continuous numerical values. Commonly used regression loss functions include:

1. **Mean Squared Error (MSE)**:
    - Computes the average of the squared differences between predicted and actual values.
    - Penalizes larger errors more heavily due to squaring, making it sensitive to outliers.
    - Easy to interpret and differentiate due to its smooth, differentiable nature.
2. **Mean Absolute Error (MAE)**:
    - Calculates the average of the absolute differences between predicted and actual values.
    - Less sensitive to outliers compared to MSE, as it does not square the errors.
    - Provides a more intuitive measure of error, but lacks smoothness, requiring alternative optimization techniques.

3. **Huber Loss**:
    - Combines characteristics of MSE and MAE to create a robust loss function.
    - Less sensitive to outliers compared to MSE while maintaining differentiability.
    - Contains a hyperparameter (δ) that controls the transition from quadratic to linear behaviour, offering flexibility in adjusting to different datasets.

• Classification Loss Functions:

Classification tasks involve categorizing inputs into discrete classes. Common classification loss functions include:

1. **Binary Cross Entropy**:
    - Used in binary classification tasks with two classes.
    - Compares predicted probabilities to actual class labels and penalizes deviations from the true values.
    - Differentiable and suitable for optimization using gradient-based methods.
2. **Categorical Cross Entropy**:
    - Applied in multi-class classification tasks with more than two classes.
    - Measures the dissimilarity between predicted class probabilities and true class labels.
    - Encourages the model to assign high probabilities to the correct classes and penalize incorrect classifications.

Activation Functions:

Activation functions determine the output of individual neurons in a neural network. The choice of activation function at the output layer depends on the nature of the task:

- For regression tasks, linear activation functions are typically used.
- For binary classification, sigmoid activation functions are employed.
- For multi-class classification, SoftMax activation functions are recommended.

Submitted by: Mrudula A P