AutoEncoder Notes

INRIA Montpellier, Evergreen Team

Leonardo Russo
leonardo.russo@inria.fr

***Abstract*—these are just notes for now.**

I. Introduction

This report discusses various loss functions commonly used in neural network architectures, particularly focussing on their definitions, mathematical expressions, and interpretations. The loss functions covered include Mean Squared Error (MSE), Mean Absolute Error (MAE), Huber Loss, and Perceptual Loss. Each loss function serves different purposes depending on the specific requirements and nature of the problem at hand.

II. Loss Functions

A. Mean Squared Error

Mean Squared Error (MSE) is a common loss function used for regression tasks, which quantify the average of squares of the differences between the predicted values and the actual values.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

where y_i is the true value, $\hat{y}_i$ is the predicted value, and n is the number of samples.

The MSE value represents the average squared difference between the estimated values and the actual values. A value of 0 indicates perfect predictions. Higher values indicate worse predictions, as the errors are squared, thus exaggerating larger errors.

B. Mean Absolute Error

Mean Absolute Error (MAE) measures the average magnitude of errors in a set of predictions, without considering their direction.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

MAE provides a straightforward measure of average error magnitude. A value of 0 indicates that there is no error (perfect predictions). It is less sensitive to outliers compared to MSE, as it does not square the error terms.

C. Huber Loss

Huber loss combines the properties of MSE and MAE. It is quadratic for small errors and linear for large errors, controlled by a parameter δ.

$$L_\delta(a) = \begin{cases} \frac{1}{2}a^2 & \text{for } |a| \le \delta \\ \delta(|a| - \frac{1}{2}\delta) & \text{otherwise} \end{cases}$$

where $a = y_i - \hat{y}_i$ is the error term.

Huber Loss is less sensitive to outliers in data than MSE, making it more robust. It is differentiable at all points, which is beneficial for gradient-based optimization algorithms.

D. Perceptual Loss

Perceptual Loss measures differences in perceptual features between predicted and target images using pre-trained deep neural network feature maps.

This loss is useful in image generation tasks where the goal is not just to match pixel values but to preserve perceptual qualities such as texture and style. Lower values indicate that the predicted image is perceptually closer to the target.

Selecting an appropriate loss function is crucial for training effective neural networks, as it directly influences how the model learns during training. The choice of loss function should be tailored to the specific characteristics and requirements of the data and task.

III. Performance Metrics

A. Peak Signal-to-Noise Ratio

PSNR is a widely used metric to measure reconstruction quality in image compression and other image restoration applications.

$$\text{PSNR} = 10\log_{10}\left(\frac{\text{MAX}_I^2}{\text{MSE}}\right)$$

where MAX_I is the maximum possible pixel value of the image and MSE is the mean squared error between the original and reconstructed images.

Higher PSNR values indicate better quality of reconstruction, implying less distortion.

B. Structural Similarity Index Measure

SSIM is used to measure the similarity between two images. SSIM considers luminance, contrast, and structure components that are relatively independent and can be combined to produce a global measure of similarity.

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

where μ_x, μ_y are the average values of x, y, σ_x^2, σ_y^2 are the variances, and σ_{xy} is the covariance. c_1, c_2 are constants used to maintain stability.

Higher SSIM values (close to 1) indicate greater similarity between the two compared images, suggesting better quality of the reconstruction.