New chat

Certainly! The equation for the cross-entropy loss in the case of multi-class classification is as follows:

Given:

- N : the number of data points
- C : the number of classes
- y_{ij} : the indicator function (1 if data point i is of class j, 0 otherwise)
- p_{ij} : the predicted probability that data point i belongs to class j

The cross-entropy loss is calculated as:

$$L(y,p) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij}\log(p_{ij})$$

Here's an expansion of the terms:

- $\sum_{i=1}^{N}$: This represents the sum over all data points.
- $\sum_{j=1}^{C}$: This represents the sum over all classes.
- y_{ij} : This is 1 if the true label of data point i is class j, and 0 otherwise. It serves as an indicator function.
- p_{ij} : This is the predicted probability that data point i belongs to class j.
- $\log(p_{ij})$: This is the natural logarithm of the predicted probability. It penalizes the model when it's confident but wrong, i.e., when it predicts a high probability for the incorrect class.
- $-$: The negative sign is applied to make it a loss function that we want to minimize.

The overall loss is the average over all data points, which is why we divide by N at the beginning.

This loss function essentially quantifies how well the predicted probabilities align with the

Send a message