PadhAI: Backpropagation - the light math version

One Fourth Labs

Partial Derivatives with respect to a

Part 2

How do we compute partial derivatives

1. Let us continue calculating the partial derivative of L w.r.t w_{212}
2. Solving the equation sequentially
   a. Let's look at the second partial derivative $\frac{\partial \hat{y}_1}{\partial a_{21}}$
      i. Here, $\hat{y}_1 = (\frac{e^{a_{21}}}{e^{a_{21}} + e^{a_{22}}})$, this is the softmax applied on a_{21}
      ii. To make it easier to compute, multiply both numerator and denominator by e^{-a21}
      iii. $\hat{y}_1 = (\frac{e^{-a_{21}}}{e^{-a_{21}}})(\frac{e^{a_{21}}}{e^{a_{21}} + e^{a_{22}}}) = \frac{1}{1 + e^{-(a_{21} - a_{22})}}$
      iv. $\frac{\partial \hat{y}_1}{\partial a_{21}} = \frac{\partial}{\partial a_{21}}(\frac{1}{1 + e^{-(a_{21} - a_{22})}})$
      v. $\frac{\partial \hat{y}_1}{\partial a_{21}} = (\frac{-1}{(1 + e^{-(a_{21} - a_{22})})^2}).(1).(e^{-(a_{21} - a_{22})}).(-1) = (\frac{1}{1 + e^{-(a_{21} - a_{22})}}).(\frac{e^{-(a_{21} - a_{22})}}{1 + e^{-(a_{21} - a_{22})}})$
      vi. Rewriting the terms $\frac{\partial \hat{y}_1}{\partial a_{21}} = \hat{y}_1(1 - \hat{y}_1)$