CSE 849 Deep Learning

HW 2: Derivatives for Batch Normalization

Due Date: 3/10/2025, 11:59 PM

For the Batch Normalization introduced in class, we discussed (Slide 23, *Tame Your Model Tricks*) that:

$$\frac{du_i}{dz_i} = \frac{\partial u_i}{\partial z_i} + \frac{\partial u_i}{\partial \mu_B}\frac{\partial \mu_B}{\partial z_i} + \frac{\partial u_i}{\partial \sigma_B^2}\frac{d\sigma_B^2}{dz_i}$$

Derive the derivatives for each component and show that your final results match those shown in Slide 24.

Grading Policy:

50 points in total. 10 points for the correctness of the derivative for each component. Show your derivation steps.

- $\frac{\partial u_i}{\partial z_i}$
- $\frac{\partial u_i}{\partial \mu_B}$
- $\frac{\partial \mu_B}{\partial z_i}$
- $\frac{\partial u_i}{\partial \sigma_B^2}$
- $\frac{\partial \sigma_B^2}{\partial z_i}$

Submission:

Submit your work in a **PDF file** through D2L.