Try different weight initialization methods (random, Xavier, He)

Tested the Batches with different weight initialization methods.

# Commenting out code related to Xavier Initialization

 # Since we are using ReLU activation we will go with He which is the default above

 # Apply Xavier initialization to all Conv2d layers

 #       for layer in self.conv\_layer:

 #           if isinstance(layer, nn.Conv2d):

 #               init.xavier\_uniform\_(layer.weight)

A screenshot of a table

Description automatically generated

**Interpretation**

1. **Batch Accuracy:** The model with He initialization achieved a higher batch accuracy (0.76563) compared to Xavier initialization (0.70313). This suggests that He initialization might have led to better convergence on a smaller batch level during the training process.
2. **Batch Loss:** The model with He initialization had a lower batch loss (4.25012) compared to Xavier initialization (5.28466). This indicates that He initialization might have resulted in a smoother optimization process with smaller errors on each batch.
3. **Epoch Accuracy:** The model with Xavier initialization achieved a slightly higher epoch accuracy (0.69354) compared to He initialization (0.67588). This suggests that Xavier initialization might have led to a better generalization performance on the entire training set after one complete epoch.
4. **Epoch Loss:** The model with He initialization had a lower epoch loss (0.0050501) compared to Xavier initialization (0.006277). This indicates that, overall, He initialization resulted in a lower average error across the entire training dataset after one epoch.
5. **Time Taken:** The training times for both models were relatively similar, with He initialization taking slightly longer (27m 20s) compared to Xavier initialization (26m 12s). However, the difference in training time is not significant enough to draw any major conclusions.

**Inference:**

Based on the tested data, both He and Xavier initialization methods yielded comparable results. However, He initialization seemed to be slightly better at converging on smaller batch sizes and minimizing the overall epoch loss. On the other hand, Xavier initialization seemed to slightly improve the epoch accuracy.