6.5 NOTE ABOUT Samplers

03 September 2025   02:55 AM   Avinash Yadav

In PyTorch, the sampler in the DataLoader *determines the strategy for selecting samples from the dataset during data loading.* It controls how indices of the dataset are drawn for each batch.

Types of Samplers

PyTorch provides several predefined samplers, and you can create custom ones:

1. **SequentialSampler:**
   - Samples elements **sequentially**, in the order they appear in the dataset.
   - Sometimes useful in time-series tasks.
   - Default when *shuffle=False*.

2. **RandomSampler:**
   - Samples elements **randomly** without replacement.
   - Default when *shuffle=True*.

3. **Custom Sampler With Custom Logic**

for example we have a sample dataset with

Class label-1 → 99 % of data
Class label-0 → 1 % of data

So while making the batches, we also want to make sure the same representation 99:1 to reflect in the batches we can utilize the concept of making our own sampler.