

BATCH (OFFLINE) AND ONLINE MACHINE LEARNING:

Batch machine learning is a type of learning where the model is trained all at once using a large batch of data, instead of learning continuously .

How it works:

Collect a large dataset -> train the model (using all the data at once) -> deploy the model -> when new data arrives -> retrain the model from scratch

Disadvantages :

Cannot learn from new data instantly

Retraining from scratch can be time consuming and computationally costly too

Not suitable for real time systems.

Online Machine Learning:

Online machine learning is a type of learning where the model learns continuously , updating itself one small batch at a time as new data arrives, it don't need to be retrained from scratch for updation (biggest advantage)

How it works:

Model starts with normal training -> new data arrives -> model updates its parameters immediately. -> learning continues without retraining from scratch .

Examples: Real time systems like Stock price prediction, recommendation systems like YouTube, Netflix etc.

What were the disadvantages of batch learning that online learning solved????

1. Cannot handle new data immediately

BATCH LEARNING:

Model cannot learn from new data unless it is retrained completely

ONLINE LEARNING:

Model updates itself instantly as new data arrives.

Solved by continuous learning.

2. Expensive retraining

- Batch Learning:**

Retraining needs **entire dataset + high computation cost**

- Online Learning:**

Updates model using **only new data points** , therefore no need for retraining from scratch

- Saves time and computational resources

3. High latency

- **Batch Learning:**

Long delay between data collection and model update 

- **Online Learning:**

Low latency – model improves immediately 

 Useful for **real-time decision systems**