CS – Machine Learning

Muhammad Hamza
Lecturer, CS
GDC KTS, Haripur

Overview

- Batch (Offline) Machine Learning
- Online Machine Learning
- Instance Based Learning
- Model Based Learning

Batch (Offline) vs Online Machine Learning

1. Batch/Offline Machine Learning

Definition:

The model is trained on the entire dataset or large batches of data at once. This means that the model does not learn continuously from new data but rather in a series of steps or updates. The model is only updated when the entire batch of data is available.

1. Batch/Offline Machine Learning

Disadvantages:

- Lots of data
- Hardware Limitation
- Availability

2. Online Machine Learning

Definition:

Online Machine Learning is a method of machine learning where the model learns incrementally, one data point at a time, as new data becomes available. Unlike batch learning, where the model is trained on the entire dataset at once, online learning updates the model continuously with each new piece of data.

2. Online Machine Learning

Key Features:

1. **Incremental Learning**: The model is updated continuously as new data arrives, rather than waiting for a large batch of data. This makes online learning ideal for applications where data is generated in real-time, such as sensor readings, user interactions, or streaming data.
2. **Real-Time Adaptation**: Online learning allows models to adapt quickly to changes in the data distribution. If the underlying patterns in the data shift over time, the model can adjust without needing to be retrained from scratch.

2. Online Machine Learning

Key Features:

3. **Lower Latency**: Since the model is updated incrementally, it can be deployed and updated with minimal delay, making it suitable for time-sensitive applications.
4. **Efficiency**: Online learning typically requires less memory and computational power compared to batch learning because it processes one data point at a time, rather than the entire dataset.

Applications of Online Machine Learning:

1. **Recommendation Systems**: Where user preferences change over time and the system needs to adapt quickly to provide relevant recommendations.
2. **Stock Market Predictions**: Where the model needs to update its predictions continuously based on new market data.
3. **Fraud Detection**: Where the model needs to adapt to new types of fraud as they are detected.

Disadvantages of Online Machine Learning:

1. **Risk of Instability**: Since the model updates frequently, it may become unstable or overly sensitive to noise in the data if not managed properly.
2. **Complexity in Implementation**: Implementing online learning can be more complex due to the need for continuous monitoring and updating.
3. **Potential for Overfitting**: The model might overfit to recent data points if not properly regularized, leading to poor generalization.

Instance Based vs Model Based Learning

Instance Based Learning

- Works by memorizing things.
- Instance-based learning, also known as **memory-based learning**, is a type of machine learning where the model memorizes the training instances rather than constructing an explicit general model.
- The algorithm makes predictions by comparing new instances with the instances stored in memory (the training data) and uses a similarity measure to determine the output.

Instance Based Learning
Key Characteristics

1. **No Explicit Model:** Instead of generating a model based on the training data, instance-based algorithms keep all or most of the training data and use it directly to make predictions.
2. **Lazy Learning:** Instance-based learning is often referred to as "lazy learning" because it defers processing until a query is made. There's no significant training phase; the algorithm only generalizes when it needs to make a prediction.
3. **Similarity Measure:** To make a prediction, the algorithm finds the most similar stored instances to the new input instance. The similarity is typically measured using distance metrics like Euclidean distance.

Instance Based Learning Examples

1. **k-Nearest Neighbors (k-NN):** One of the most common instance-based learning algorithms, which classifies a new instance based on the majority class among its nearest neighbors.
2. **Locally Weighted Learning:** Where predictions are made by fitting a model locally using instances that are close to the query point.
3. **Case-Based Reasoning:** A method where a new problem is solved by finding a similar past case and reusing it.

Instance Based Learning

Advantages:

- **Flexibility:** Can adapt quickly to new data without retraining.
- **Simplicity:** Easy to implement and understand, particularly k-NN.

Disadvantages:

- **Storage:** Requires storing all training data, which can be inefficient in terms of memory.
- **Speed:** Making predictions can be slow for large datasets since the algorithm needs to compare the new instance with many (or all) stored instances.
- **Noise Sensitivity:** Prone to overfitting and affected by noisy data since it memorizes all instances, including outliers.

Model based Learning

Model-based learning is a way of learning where the system creates and uses an internal model to understand how things work in the environment. This model helps the system predict what will happen next and plan its actions accordingly.

Advantages of Model based Learning

Better Planning: The system can think ahead by predicting future outcomes, leading to smarter decisions.

Fewer Trials Needed: Because it can simulate different scenarios, it doesn’t need to try as many things in the real world, saving time and resources.

Adaptability: If the environment changes, the system can update its model and quickly adjust its actions.

Disadvantages of Model based Learning

Complex to Build: Creating an accurate model can be hard and requires a lot of effort.

Depends on Accuracy: If the model is wrong or incomplete, the system may make bad decisions.

High Computational Cost: Running simulations and updating the model can be computationally expensive, especially in complex environments.

Thank You