**Machine Learning Algorithms suitable for Online Learning**

Several machine learning algorithms are designed or adapted to process data in real time, also known as **online learning** or **incremental learning**. These algorithms update their models as new data points arrive, without requiring the entire dataset to be loaded into memory. Here are a few commonly used real-time algorithms:

**1. Stochastic Gradient Descent (SGD)**

* **Description**: An online learning version of gradient descent, which updates the model's weights incrementally based on one data point at a time.
* **Use Cases**: Linear regression, logistic regression, and neural networks in real-time learning scenarios.
* **Library Support**: Available in sklearn.linear\_model.SGDClassifier, sklearn.linear\_model.SGDRegressor.

**2. Naive Bayes (Online)**

* **Description**: Naive Bayes classifiers can be updated incrementally. They estimate class probabilities based on the likelihood of features, and these estimates can be updated when new data arrives.
* **Use Cases**: Text classification, spam filtering, sentiment analysis.
* **Library Support**: river.naive\_bayes.GaussianNB (in river), sklearn.naive\_bayes.MultinomialNB (with partial fit).

**3. Decision Trees (Incremental)**

* **Description**: Decision trees like the **Hoeffding Tree** (also known as **Very Fast Decision Tree** or **VFDT**) are capable of growing and adapting incrementally as new data becomes available.
* **Use Cases**: Classification tasks with evolving data, such as online recommendation systems or fraud detection.
* **Library Support**: river.tree.HoeffdingTree (in river), sklearn.tree.DecisionTreeClassifier (with partial\_fit).

**4. Online K-Means**

* **Description**: An adaptation of the classic K-Means clustering algorithm, which allows updating the centroids incrementally with each new data point.
* **Use Cases**: Real-time clustering tasks, such as customer segmentation in marketing.
* **Library Support**: river.cluster.KMeans (in river).

**5. Passive-Aggressive Classifier**

* **Description**: This is an online learning algorithm designed for classification tasks. It makes small updates to the model when predictions are incorrect, and it does nothing when the prediction is correct.
* **Use Cases**: Text classification and real-time prediction systems.
* **Library Support**: sklearn.linear\_model.PassiveAggressiveClassifier.

**6. Online Random Forest**

* **Description**: Random forests can be adapted for online learning, with incremental versions such as **Online Random Forests** or **Hoeffding Random Forests**. These algorithms update trees and their ensemble of models as new data arrives.
* **Use Cases**: Classification or regression in real-time, such as stock market predictions.
* **Library Support**: river.ensemble.HoeffdingAdaptiveTree (a variant in river).

**7. Perceptron**

* **Description**: The perceptron algorithm is a type of neural network that is inherently incremental, allowing for real-time updates to the weights as new data is encountered.
* **Use Cases**: Real-time binary classification tasks.
* **Library Support**: sklearn.linear\_model.Perceptron.

**8. Support Vector Machines (SVM) (with Online Learning)**

* **Description**: Support vector machines can be adapted for online learning by using an algorithm like **LibSVM's** svm\_online or **Linear SVMs** with incremental updates.
* **Use Cases**: Classification tasks in real-time, particularly for text or image classification.
* **Library Support**: sklearn.svm.SVC (with partial\_fit).

**9. Hough Transform (Online)**

* **Description**: Adapted for real-time detection tasks, such as feature detection or line detection in images.
* **Use Cases**: Image processing and real-time computer vision tasks.
* **Library Support**: Custom implementations, but not commonly available in mainstream ML libraries.

**10. Matrix Factorization (Online)**

* **Description**: Algorithms for collaborative filtering, such as **Stochastic Gradient Descent (SGD)** for matrix factorization, can be applied incrementally.
* **Use Cases**: Real-time recommendation systems (e.g., Netflix, Amazon recommendations).
* **Library Support**: river.recommendation.SGD (in river).