Vocabulary

words & phrases	definition	Reference in Wiki
out-of-core_algorithms	In computing, external memory algorithms or out-of-core algorithms are algorithms that are designed to process data that is too large to fit into a computer's main memory at one time. Such algorithms must be optimized to efficiently fetch and access data stored in slow bulk memory (auxiliary memory) such as hard drives or tape drives, or when memory is on a computer network.	Online learning is a common technique used in areas of machine learning where it is computationally infeasible to train over the entire dataset, requiring the need of out-of-core algorithms.
catastrophic interference	Catastrophic interference, also known as catastrophic forgetting, is the tendency of an artificial neural network to completely and abruptly forget previously learned information upon learning new information.	Online learning algorithms may be prone to catastrophic interference, a problem that can be addressed by incremental learning approaches.
Progressive learning	Progressive learning is an effective learning model which is demonstrated by the human learning process. It is the process of learning continuously from direct experience.	Progressive learning technique (PLT) in machine learning can learn new classes (or labels) dynamically on the run.

1. Online Learning

从技术的角度上来看。传统中狭义的machine learning技术，是利用一批已有的数据，学习到一个固化的模型。该模型的泛化能力，不仅依赖于精心设计的模型，更需要一次性灌注海量数据来保证。而 **online learning 则不需要启动数据，或只需少量启动数据，通过探索，反馈，修正来逐渐学习。**相比之下，online learning 对数据的使用更加灵活，由此带来的好处，不仅是能够减轻更新模型时的计算负担，更可以提高模型的时效性，这更加符合人的学习方式。**传统的machine learning，则是侧重于统计分析；在线学习是在哲学上真正模仿人学习过程的研究。**

（1）与Batch Learning的区别

同于 Batch，Online 中每次 W 的更新并不是沿着全局梯度进行下降，而是沿着某个样本的产生的梯度方向进行下降，整个寻优过程变得像是一个"随机"查找的过程 (SGD 中 Stochastic 的来历)，这样 Online 最优化求解即使采用 L1 正则化的方式，也很难产生稀疏解。后面介绍的各个在线最优化求解算法中，稀疏性是一个主要的追求目标。

（2）在线学习经典算法

I. FTL (Follow the Leader)

$$w_t = \arg\min_{w \in S} \sum_{i=1}^{t-1} v_i(w)$$

当前step更新参数时考虑之前所有的最小损失 (least loss over all past rounds)。

公式里的S代表整个参数空间。

II. FTRL (Follow the regularised leader)

正则化 (Regularization) 的意义本质上是为了避免训练得到的模型过度拟合(overfitting) 训练数据。

相较于FTL，正则项的目的是：to stabilize the FTL solutions and obtain better regret bounds.

$$w_t = \arg\min_{w \in S} \sum_{i=1}^{t-1} v_i(w) + R(w)$$

A regularization function $R : S \rightarrow \mathbb{R}$ is chosen and learning performed in round t.

（3）OML 和 Incremental Learning 的关系

online learning 包括了 incremental learning 和 decremental learning等情况，描述的是一个动态学习的过程。前者是增量学习，每次学习一个或多个样本，这些训练样本可以全部保留、部分保留或不保留；后者是递减学习，即抛弃"价值最低"的保留的训练样本。

2. Progressive Learning

Definition

Progressive learning is an effective learning model which is demonstrated by the human learning process. It is the process of learning continuously from direct experience. **Progressive learning technique (PLT) in machine learning can learn new classes (or labels) dynamically on the run.[1]**

Reference

[1] A Novel Progressive Learning Technique for Multi-class Classification

> Venkatesan, Rajasekar, and Meng Joo Er. "A novel progressive learning technique for multi-class classification." *Neurocomputing* 207 (2016): 310-321.
>
> https://arxiv.org/abs/1609.00085