

随机森林的增量学习：

Saffari A , Leistner C , Santner J , et al. On-line Random Forests[C]// Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on. IEEE, 2009.

一棵树的决策节点的形式为 $g(x) > \theta$ 。它包含两个部分，一个是返回标量值的函数 g ，一个是阈值 θ 。在离线模式，随机森林随机选取一些这样的测试函数，根据度量得到最优的函数。如果阈值也是随机选取的，那么称这样的随机森林为 Extremely Randomized Forest。

在线模式。采用了这种 ERF，随机生成测试函数和阈值。在随机树的生长过程中，每个决策节点随机生成一些测试并根据度量选取最好的测试函数。通常选取信息熵或者 gini 进行度量，

特别地，当产生一个节点的同时，这个节点产生 N 个随机测试 S 。之后，

$S = \{(g_1(x), \theta_1), \dots, (g_N(x), \theta_N)\}$ 。这个节点收集所有落入这个

节点的样本统计信息。 $\mathbf{p}_j = [p_1^j, \dots, p_K^j]$ 。一共有 K 类样本， \mathbf{p} 表示落入节点 j 的类标签统计。对一个随机测试 s ，收集两组统计。

$\mathbf{p}_{jls} = [p_1^{jls}, \dots, p_K^{jls}]$ $\mathbf{p}_{jrs} = [p_1^{jrs}, \dots, p_K^{jrs}]$ 。表示在测试 s ，落入左边和右边的样本的统计。所以根据测试 s ，可以得到增益

$$\Delta L(\mathcal{R}_j, s) = L(\mathcal{R}_j) - \frac{|\mathcal{R}_{jls}|}{|\mathcal{R}_j|} L(\mathcal{R}_{jls}) - \frac{|\mathcal{R}_{jrs}|}{|\mathcal{R}_j|} L(\mathcal{R}_{jrs}),$$

。一般选

择增益最大的测试进行分裂。

离线模式，每个节点能得到落入这个节点的所有数据，所以可以根据统计得到一个比在线模式更加鲁棒的模型。而在在线模式，所有统计都是实时获取，所以依靠以下两点进行分裂决策。1) 当有足够的样本落入这个节点（保证鲁棒），2) 对于分类来说，这个分裂会有足够帮助。

一棵新生成的树只包含根节点，这个节点上有随机选取测试的集合。我们在线手机这个几点上每一个测试的统计。在这里，有两个超参数：1) 在一个节点分裂前，需要最少收集 α 个样本，2) 最小的分裂增益为 β 。因此，当 $|\mathcal{R}_j| > \alpha$ 和 $\exists s \in S : \Delta L(\mathcal{R}_j, s) > \beta$ ，节点分裂。

分裂后，将 \mathbf{p}_{jls} 和 \mathbf{p}_{jrs} 传播到新生成的左右叶节点。这里，一个新生成的节点是由父节点现有的统计得到，所有新节点可以直接保留分类的性能。