模型输入输出说明

I-BRT

输⼊数据X：m x n（数据数\*变量数）的np.array

输⼊数据y：m x 1的np.array

模型参数1——n\_iter：定义模型基本树个数的int

模型参数2——\_gamma：定义模型的某正则化参数float

模型参数3——\_lamda：定义模型的某正则化参数float

模型参数4——max\_depth：定义模型基本树最⼤深度int

输⼊数据testX：m x n（数据数\*变量数）的np.array

输⼊数据testy：m x 1的np.array

RF-ANFIS

输⼊数据train\_data：包括X和y的训练数据torch.DataLoader

输⼊数据test\_data：包括X和y的训练数据torch.DataLoader

输出数据y\_pred：预测结果torch.tensor

假定数据均以 Excel 的形式存储。

GBM 输入输出数据说明

输入数据说明

输入一个数据文件，第一行为表头；

第一列为样本序号，最后一列为因变量 (腐蚀率)，中间的列为可能影响到腐蚀率的各种因素。

除第一列样本序号外，其余列的值均为 \*\*浮点类型\*\*。

输出数据说明

输出数据为 (N, 1) 的向量，分别表示每一行的预测结果 (浮点型)。

HLM 输入输出数据说明

输入数据说明

输入分为 \*\*两个\*\* 数据文件，一个为某站点土壤的 \*\*固定因素\*\* 数据 ，假设为 `soid\_data`；另一个为自变量和因变量数据，设为 `erosion\_data`；

- `soil\_data` ：第一行为表头；第一列为站点名称，之后的列为该站点土壤的固定因素，至少要包含 土壤 PH 值、全氮含量(%)、碳酸根离子浓度(%)、硫酸根离子浓度(%)、镁离子浓度(%)、钾离子浓度(%)、钠离子浓度(%)、土壤电阻(Ω) 8 种主要影响因素。

除站点名称为 string 类型外，其余均为浮点型。另请保证各站点的土壤因素数据 \*\*唯一\*\*。

- `erosion\_data`：第一行为表头；数据共 \*\*三列\*\*，第一列为站点名称，第二列为埋存时间 (单位年，最好为整数)，第三列为腐蚀率 (浮点型)。

输出数据说明

输出为 (N, 1) 的向量，分别表示 对应站点不同埋存时间的 腐蚀率的 预测结果 (浮点型)。

Mert

输入：

1. 训练集，要求第一列为因变量Y,之后N列为自变量X1,X2,...,XN。
2. 测试集，要求第一列为因变量Y,之后M列为自变量X1,X2,...,XM。
3. n(int，默认值为1)，表示训练集及测试集所包含的观察对象(即Y)的种类数量，该模型要求每类所包含的样本量相同。
4. k(int，默认值为1)，表示选择自变量Xk为与随机效应相关的自变量。
5. Epoch(int，默认值为200)，表示迭代次数。

输出：

1. 根据测试集得出的预测值Y0。

Rebet

输入：

1. 训练集，要求第一列为因变量Y,之后N列为自变量X1,X2,...,XN。
2. 测试集，要求第一列为因变量Y,之后M列为自变量X1,X2,...,XM。
3. n(int，默认值为1)，表示训练集及测试集所包含的观察对象(即Y)的种类数量，该模型要求每类所包含的样本量相同。
4. k(int，默认值为1)，表示选择自变量Xk为与随机效应相关的自变量。
5. Epoch(int，默认值为200)，表示迭代次数。
6. M(float，默认值为10)，表示迪利克雷过程的离散程度。

输出：

1. 根据测试集得出的预测值Y0。

Ga(遗传算法)

输入：

1. c(int，必选)，指定求函数的最大值或最小值，‘1’为求最大值，‘0’为求最小值。
2. F(function，必选)，指定所要求最值的函数表达式。
3. N(int，必选)，指定函数所含变量个数。
4. Ranges(array，必选)，指定各个变量的取值范围。
5. Precisions(int，可选，默认值=24)，指定精度。
6. N\_GENERATIONS(int，可选，默认值=50)，指定迭代轮数。
7. POP\_SIZE(int，可选，默认值 = 200)，指定种群大小。
8. MUTATION\_RATE(float，可选，默认值 = 0.005)，指定变异概率。
9. CROSSOVER\_RATE(float，可选，默认值 = 0.8)，指定交叉概率。

输出：

1. 最优值。
2. 取最优值时，各个变量的值。