1. **已实验算法记录**
   1. **多目标-线性模型（linear）**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 优化目标\算法 | FPA, AAE | FPA, NNZ | FPA, AAE, NNZ | FPA, MSE | FPA, MSE, NNZ | FPA, L1 | FPA, L1, AAE | FPA, L1, MSE |
| NSGA2 | **√** | **√** | **√** | **√** | **√** | **√** | **√** | **√** |
| NSGA2-toZero |  | **√** |  | **√** | **√** | **√** |  | **√** |
| NSGA2-10p-toZero | **√** | **√** | **√** | **√** | **√** | **√** | **√** | **√** |
| NSGA2-20p-toZero | **√** | **√** | **√** | **√** | **√** | **√** | **√** | **√** |
| NSGA2-30p-toZero | **√** | **√** | **√** | **√** | **√** | **√** | **√** | **√** |
| NSGA2\_random10p |  |  |  | **√** | **√** |  |  | **√** |
| NSGA2\_random20p |  |  |  | **√** | **√** |  |  | **√** |
| NSGA2\_random30p |  |  |  | **√** | **√** |  |  | **√** |

* 1. **多目标-MLP**

（模型优化时，未考虑bias，bias不进行优化）

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 优化目标\算法 | FPA, AAE | FPA, NNZ | FPA, AAE, NNZ | FPA, MSE | FPA, MSE, NNZ | FPA, L1 | FPA, L1, AAE | FPA, L1, MSE |
| NSGA2 |  | **√** |  | **√** | **√** |  |  | **√** |
| NSGA2-toZero |  | **√** |  | **√** | **√** |  |  | **√** |
| NSGA2-10p-toZero |  | **√** |  | **√** | **√** |  |  | **√** |
| NSGA2-20p-toZero |  | **√** |  | **√** | **√** |  |  | **√** |
| NSGA2-30p-toZero |  | **√** |  | **√** | **√** |  |  | **√** |
| NSGA2\_random10p |  |  |  | **√** | **√** |  |  | **√** |
| NSGA2\_random20p |  |  |  | **√** | **√** |  |  | **√** |
| NSGA2\_random30p |  |  |  | **√** | **√** |  |  | **√** |

* 1. **多目标-mlp3**

注：mlp3模型，相比MLP模型，考虑bias的优化，且三层感知机模型的隐藏层值被设为3。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 优化目标\算法 | FPA, AAE | FPA, NNZ | FPA, AAE, NNZ | FPA, MSE | FPA, MSE, NNZ | FPA, L1 | FPA, L1, AAE | FPA, L1, MSE |
| NSGA2 |  |  |  |  |  |  |  |  |
| NSGA2-toZero |  |  |  |  |  |  |  |  |
| NSGA2-10p-toZero |  |  |  |  |  |  |  |  |
| NSGA2-20p-toZero |  |  |  |  |  |  |  |  |
| NSGA2-30p-toZero |  |  |  |  |  |  |  |  |
| NSGA2\_random10p |  |  |  |  |  |  |  | **√** |
| NSGA2\_random20p |  |  |  |  |  |  |  | **√** |
| NSGA2\_random30p |  |  |  |  |  |  |  | **√** |

* 1. **多目标-mlp5**

注：mlp5模型，相比MLP模型，考虑bias的优化，且三层感知机模型的隐藏层值被设为5。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 优化目标\算法 | FPA, AAE | FPA, NNZ | FPA, AAE, NNZ | FPA, MSE | FPA, MSE, NNZ | FPA, L1 | FPA, L1, AAE | FPA, L1, MSE |
| NSGA2 |  |  |  |  |  |  |  |  |
| NSGA2-toZero |  |  |  |  |  |  |  |  |
| NSGA2-10p-toZero |  |  |  |  |  |  |  |  |
| NSGA2-20p-toZero |  |  |  |  |  |  |  |  |
| NSGA2-30p-toZero |  |  |  |  |  |  |  |  |
| NSGA2\_random10p |  |  |  |  |  |  |  | **√** |
| NSGA2\_random20p |  |  |  |  |  |  |  | **√** |
| NSGA2\_random30p |  |  |  |  |  |  |  | **√** |

* 1. **单目标-优化FPA**

|  |  |  |  |
| --- | --- | --- | --- |
| 算法\模型 | Linear | mlp3 | mlp5 |
| CoDE | **√** | **√** |  |
| CoDE\_toZero | **√** | **√** |  |
| CoDE\_10p\_toZero | **√** | **√** |  |
| CoDE\_20p\_toZero | **√** | **√** |  |
| CoDE\_10p\_lr\_toZero | **√** | **√** |  |
| CoDE\_20p\_lr\_toZero | **√** | **√** |  |
| CoDE\_random10p\_toZero | **√** | **√** |  |
| CoDE\_random20p\_toZero | **√** | **√** |  |
| CoDE\_random30p\_toZero | **√** | **√** |  |
| DE\_rand\_1\_bin | **√** |  |  |

* 1. **传统机器学习算法**

|  |  |  |
| --- | --- | --- |
| 算法\设置运行次数 | 1 | 10 |
| Linear regression | **√** |  |
| Ridge regression（默认参数） | **√** |  |
| RidgeCV(调参) | **√** | **√** |
| LassoLarsCV（调参） | **√** | **√** |
| RandomForestRegressor | **√** |  |
| MLPRegressor | **√** |  |

1. **实验对比(附****lasso regression，random forest，mlp对比)**

【table/result】记录所有跑出来的算法的——单目标算法测试结果；多目标算法测试结果中FPA最高的解；传统机器学习算法;

|  |  |  |
| --- | --- | --- |
| 算法 | train | test |
| RidgeCV | 0.757989555 | 0.688552667 |
| LassoLarsCV | 0.730934226 | 0.698958096 |
| MLPRegressor | 0.704706235 | 0.662779361 |
| RandomForestRegressor | 0.880300706 | 0.708868051 |

* 1. **单目标-linear结果对比**

【图片见plot/single\_linear】（train文件夹存放：对每个数据集，在图上绘制出各个算法的训练结果；test文件夹存放：对每个数据集，在图上绘制出各个算法的测试结果；图片仅绘制了测试结果前三的算法。）

以下表格为训练结果：

|  |  |  |
| --- | --- | --- |
| 算法 | train | test |
| linear/CoDE | 0.781313 | 0.699055 |
| linear/CoDE\_toZero | 0.781571 | 0.69462 |
| linear/CoDE\_10p\_toZero | 0.781571 | 0.692735 |
| linear/CoDE\_20p\_toZero | 0.780626 | 0.695438 |
| linear/CoDE\_10p\_lr\_toZero | 0.781029 | 0.693202 |
| linear/CoDE\_20p\_lr\_toZero | 0.780683 | 0.691953 |
| linear/CoDE\_random10p\_toZero | 0.778431 | 0.695293 |
| linear/CoDE\_random20p\_toZero | 0.776606 | 0.697349 |
| linear/CoDE\_random30p\_toZero | 0.775617 | 0.701456 |

总结：引入random因素的CoDE算法可以得到更优结果。

* 1. **单目标-mlp3结果对比**

【图片见plot/single\_mlp3】（train文件夹存放：对每个数据集，在图上绘制出各个算法的训练结果；test文件夹存放：对每个数据集，在图上绘制出各个算法的测试结果）

以下表格为训练结果：

|  |  |  |
| --- | --- | --- |
| 算法 | train | test |
| mlp3/CoDE | 0.7537 | 0.661167 |
| mlp3/CoDE\_toZero | 0.754124 | 0.661236 |
| mlp3/CoDE\_10p\_toZero | 0.753634 | 0.663287 |
| mlp3/CoDE\_20p\_toZero | 0.757774 | 0.666017 |
| mlp3/CoDE\_10p\_lr\_toZero | 0.755336 | 0.671893 |
| mlp3/CoDE\_20p\_lr\_toZero | 0.759007 | 0.664515 |
| mlp3/CoDE\_random10p\_toZero | 0.754948 | 0.663571 |
| mlp3/CoDE\_random20p\_toZero | 0.755474 | 0.670945 |
| mlp3/CoDE\_random30p\_toZero | 0.760608 | 0.673672 |

总结：以mlp3为模型的单目标优化算法效果不佳。

* 1. **多目标-linear结果对比**

【图片见plot/multi\_linear】（仅绘制优化FPA+MSE+L1的结果：/train文件夹存放：对每个数据集，在图上绘制出各个算法的训练结果；/test文件夹存放：对每个数据集，在图上绘制出各个算法的训练结果）

|  |  |  |
| --- | --- | --- |
| 算法 | train | test |
| linear/nsga2/FPA\_AAE | 0.784844 | 0.702313 |
| linear/nsga2/FPA\_NNZ | 0.775208 | 0.685318 |
| linear/nsga2/FPA\_AAE\_NNZ | 0.783888 | 0.702926 |
| linear/nsga2/FPA\_MSE | 0.78134 | 0.701537 |
| linear/nsga2/FPA\_NNZ\_MSE | 0.783054 | 0.709035 |
| linear/nsga2/FPA\_L1 | 0.775335 | 0.724109 |
| linear/nsga2/FPA\_AAE\_L1 | 0.780866 | 0.731516 |
| linear/nsga2/FPA\_L1\_MSE | 0.780607 | 0.73178 |
| linear/nsga2\_10p\_toZero/FPA\_AAE | 0.782715 | 0.711981 |
| linear/nsga2\_10p\_toZero/FPA\_NNZ | 0.772965 | 0.710297 |
| linear/nsga2\_10p\_toZero/FPA\_AAE\_NNZ | 0.779599 | 0.728557 |
| linear/nsga2\_10p\_toZero/FPA\_MSE | 0.782957 | 0.707554 |
| linear/nsga2\_10p\_toZero/FPA\_NNZ\_MSE | 0.780779 | 0.731268 |
| linear/nsga2\_10p\_toZero/FPA\_L1 | 0.775699 | 0.724323 |
| linear/nsga2\_10p\_toZero/FPA\_AAE\_L1 | 0.780246 | 0.732254 |
| linear/nsga2\_10p\_toZero/FPA\_L1\_MSE | 0.780301 | 0.731335 |
| linear/nsga2\_20p\_toZero/FPA\_AAE | 0.780088 | 0.709393 |
| linear/nsga2\_20p\_toZero/FPA\_NNZ | 0.769466 | 0.714252 |
| linear/nsga2\_20p\_toZero/FPA\_AAE\_NNZ | 0.778201 | 0.723624 |
| linear/nsga2\_20p\_toZero/FPA\_MSE | 0.780327 | 0.709141 |
| linear/nsga2\_20p\_toZero/FPA\_NNZ\_MSE | 0.777368 | 0.72602 |
| linear/nsga2\_20p\_toZero/FPA\_L1 | 0.773931 | 0.719056 |
| linear/nsga2\_20p\_toZero/FPA\_AAE\_L1 | 0.777772 | 0.727257 |
| linear/nsga2\_20p\_toZero/FPA\_L1\_MSE | 0.778538 | 0.730681 |
| linear/nsga2\_30p\_toZero/FPA\_AAE | 0.779493 | 0.714065 |
| linear/nsga2\_30p\_toZero/FPA\_NNZ | 0.768666 | 0.718313 |
| linear/nsga2\_30p\_toZero/FPA\_AAE\_NNZ | 0.774647 | 0.723453 |
| linear/nsga2\_30p\_toZero/FPA\_MSE | 0.778262 | 0.717974 |
| linear/nsga2\_30p\_toZero/FPA\_NNZ\_MSE | 0.775879 | 0.729116 |
| linear/nsga2\_30p\_toZero/FPA\_L1 | 0.770819 | 0.724483 |
| linear/nsga2\_30p\_toZero/FPA\_AAE\_L1 | 0.776635 | 0.728081 |
| linear/nsga2\_30p\_toZero/FPA\_L1\_MSE | 0.77575 | 0.729932 |
| linear/nsga2\_random10p\_toZero/FPA\_MSE | 0.7841 | 0.717175 |
| linear/nsga2\_random10p\_toZero/FPA\_NNZ\_MSE | 0.780234 | 0.725537 |
| linear/nsga2\_random10p\_toZero/FPA\_L1\_MSE | 0.782483 | 0.733901 |
| linear/nsga2\_random20p\_toZero/FPA\_MSE | 0.78063 | 0.710933 |
| linear/nsga2\_random20p\_toZero/FPA\_NNZ\_MSE | 0.777499 | 0.727436 |
| linear/nsga2\_random20p\_toZero/FPA\_L1\_MSE | 0.777198 | 0.740582 |
| linear/nsga2\_random30p\_toZero/FPA\_MSE | 0.776794 | 0.712907 |
| linear/nsga2\_random30p\_toZero/FPA\_NNZ\_MSE | 0.772861 | 0.720324 |
| linear/nsga2\_random30p\_toZero/FPA\_L1\_MSE | 0.773681 | 0.734325 |
| linear/nsga2\_toZero/FPA\_NNZ | 0.776045 | 0.711963 |
| linear/nsga2\_toZero/FPA\_MSE | 0.78261 | 0.706886 |
| linear/nsga2\_toZero/FPA\_NNZ\_MSE | 0.781963 | 0.718236 |
| linear/nsga2\_toZero/FPA\_L1 | 0.775175 | 0.718999 |
| linear/nsga2\_toZero/FPA\_L1\_MSE | 0.781518 | 0.733622 |

总结：在以线性模型为基础的多目标优化算法中，发现引入random的算法可以得到更优效果，且优化目标为FPA+L1+MSE的组合时性能更佳。

* 1. **多目标-mlp3结果对比**

【图片见plot/multi\_mlp3】（仅绘制优化FPA+MSE+L1的结果：/train文件夹存放：对每个数据集，在图上绘制出各个算法的训练结果；/test文件夹存放：对每个数据集，在图上绘制出各个算法的测试结果）

|  |  |  |
| --- | --- | --- |
| 算法 | train | test |
| mlp3/nsga2\_random10p\_toZero/FPA\_L1\_MSE | 0.784156 | 0.746064 |
| mlp3/nsga2\_random20p\_toZero/FPA\_L1\_MSE | 0.779157 | 0.73962 |
| mlp3/nsga2\_random30p\_toZero/FPA\_L1\_MSE | 0.772148 | 0.741749 |
| RidgeCV | 0.757989555 | 0.688552667 |
| LassoLarsCV | 0.730934226 | 0.698958096 |
| MLPRegressor | 0.704706235 | 0.662779361 |
| RandomForestRegressor | 0.880300706 | 0.708868051 |

总结：上述表格内为算法训练后在所有数据集上的平均值，nsga2\_random10p\_toZero算法总体来说可以得到更优结果,，且相对于lasso regression，random forest，mlp，ridge regression可以得到更高的FPA值。

* 1. **多目标-mlp5结果对比**

【图片见plot/multi\_mlp5】（仅绘制优化FPA+MSE+L1的结果：/train文件夹存放：对每个数据集，在图上绘制出各个算法的训练结果；/test文件夹存放：对每个数据集，在图上绘制出各个算法的测试结果）

|  |  |  |
| --- | --- | --- |
| 算法 | train | test |
| mlp5/nsga2\_random10p\_toZero/FPA\_L1\_MSE | 0.784585 | 0.739426005 |
| mlp5/nsga2\_random20p\_toZero/FPA\_L1\_MSE | 0.777993 | 0.739569736 |
| mlp5/nsga2\_random30p\_toZero/FPA\_L1\_MSE | 0.77278 | 0.740476391 |
| RidgeCV | 0.757989555 | 0.688552667 |
| LassoLarsCV | 0.730934226 | 0.698958096 |
| MLPRegressor | 0.704706235 | 0.662779361 |
| RandomForestRegressor | 0.880300706 | 0.708868051 |

总结：上述表格内为算法训练后在所有数据集上的平均值，不同算法得到的结果差不多，且相对于lasso regression，random forest，mlp，ridge regression可以得到更高的FPA值。

* 1. **总结**

以下为训练性能前十的算法(完整排序见table\result(sortedtrain))

|  |  |
| --- | --- |
| RandomForestRegressor | 0.880300706 |
| linear/nsga2/FPA\_AAE | 0.784843983 |
| mlp5/nsga2\_random10p\_toZero/FPA\_L1\_MSE | 0.784585348 |
| mlp3/nsga2\_random10p\_toZero/FPA\_L1\_MSE | 0.784155536 |
| linear/nsga2\_random10p\_toZero/FPA\_MSE | 0.784100297 |
| linear/nsga2/FPA\_AAE\_NNZ | 0.783888041 |
| linear/nsga2/FPA\_NNZ\_MSE | 0.783053732 |
| linear/nsga2\_10p\_toZero/FPA\_MSE | 0.782956697 |
| linear/nsga2\_10p\_toZero/FPA\_AAE | 0.782714536 |
| linear/nsga2\_toZero/FPA\_MSE | 0.782609932 |

以下为测试性能为前十的算法(完整排序见table\result(sortedtest))：

|  |  |
| --- | --- |
| mlp3/nsga2\_random10p\_toZero/FPA\_L1\_MSE | 0.746063856 |
| mlp3/nsga2\_random30p\_toZero/FPA\_L1\_MSE | 0.741749155 |
| linear/nsga2\_random20p\_toZero/FPA\_L1\_MSE | 0.740581622 |
| mlp5/nsga2\_random30p\_toZero/FPA\_L1\_MSE | 0.740476391 |
| mlp3/nsga2\_random20p\_toZero/FPA\_L1\_MSE | 0.739619939 |
| mlp5/nsga2\_random20p\_toZero/FPA\_L1\_MSE | 0.739569736 |
| mlp5/nsga2\_random10p\_toZero/FPA\_L1\_MSE | 0.739426005 |
| linear/nsga2\_random30p\_toZero/FPA\_L1\_MSE | 0.734324965 |
| linear/nsga2\_random10p\_toZero/FPA\_L1\_MSE | 0.733900721 |
| linear/nsga2\_toZero/FPA\_L1\_MSE | 0.733621969 |

总体来说：在训练集结果上，randomforest结果最佳；在测试集结果上，多目标优化可以获得更优性能；且以mlp3，mlp5模型为基础的结果优于以线性模型为基础；引入random的多目标优化算法相对于其他置零法表现更好。

1. **非支配集选取较优解**

**3.1新增实验——随机选取数据集的80%为训练集，20%为验证集**

|  |  |  |
| --- | --- | --- |
| **模型** | **优化目标** | **优化算法** |
| MLP | FPA+MSE+L1 | NSGA2\_random10p |
| NSGA2\_random20p |
| NSGA2\_random30p |
| mlp3 | NSGA2\_random10p |
| NSGA2\_random20p |
| NSGA2\_random30p |
| mlp5 | NSGA2\_random10p |
| NSGA2\_random20p |
| NSGA2\_random30p |

**3.2 实验方法**

**(1) 将数据集划分为训练集和验证集，使用训练集进行训练，验证集测试模型的优劣；从非支配集中寻找较优解时，同时参考训练集和验证集中的结果进行选取；随后再在测试集上进行测试。**

**(2) 选取时，同时考虑多个目标的值，即将FPA,L1,MSE的值加权**

**(3) 选取时引入一定得随机性。**

**3.3 目前结果**

**达到接近于randomforest的测试结果。一些实验及相关结果见表格【table\非支配集寻找对比.xlsx】**