

常见机器学习算法：

- 1. 线性回归
- 2. 逻辑回归
- 3. SVM
- 4. 决策树
- 5. k-NN
- 6. k-means
- 7. 朴素/贝叶斯
- 8. PCA
- 9. ensemble/bagging/随机森林
- 10. ensemble/boosting/Adaboost
- 11. ensemble/boosting/GBDT
- 12. ensemble/boosting/XGBoost

可解释性、对异常点是否敏感、有/无监督、可/不可并行、bagging/ensemble、数学假设、过拟合、能/不能处理缺失值、可否处理高维数据

|      | 可解释性 | 对异常点是否敏感-泛化性 | 有 / 无监督 | 可 / 不可并行 | e n s a m b l e | 数学假设  | 过拟合 | 缺失值 | 适合 / 不适合处理高位数据 |
|------|------|--------------|---------|----------|-----------------|---|-----|-----|----------------|
| 线性回归 | 较好   | 敏感           | 有       | 不可       | 不是              | 随机误差独立同分布，服从期望（条件期望）为0的正态分布；解释变量和目标变量近似满足线性 | 容易  | 不能  | -              |

|         |     |     |   |    |          |               |     |    |     |
|---------|-----|-----|---|----|----------|---------------|-----|----|-----|
|         |     |     |   |    |          | 关系；           |     |    |     |
| 逻辑回归    | 较好  | 敏感  | 有 | 不可 | 不是       | 事件发生概率服从二项分布  | 容易  | 不能 | -   |
| SVM     | 较好  | 不敏感 | 有 | 不可 | 不是       | -             | 不容易 | 不能 | 适合  |
| 决策树     | 较好  | 不敏感 | 有 | 不可 | 不是       | -             | 容易  | 不能 | 适合  |
| k-NN    | 较好  | 不敏感 | 有 | 不可 | 不是       | -             | -   | 不能 | 不适合 |
| k-means | 较好  | 敏感  | 无 | 不可 | 不是       | -             | -   | 不能 | 不合适 |
| 朴素贝叶斯   | 较差  | 不敏感 | 有 | 不可 | 不是       | 各个特征变量取值是相互独立 | -   | 能  | -   |
| 随机森林    | 比较差 | 不敏感 | 有 | 可以 | bagging  | -             | 不容易 | 能  | 合适  |
| Ada     | 较   | 敏   |   | 不  | boosting |               | 不   | 看基 | 不   |

|                                 |             |             |   |             |                                      |   |             |             |             |
|---------------------------------|-------------|-------------|---|-------------|--------------------------------------|---|-------------|-------------|-------------|
| b<br>o<br>o<br>s<br>t           | 好           | 感           | 有 | 可           | s<br>t<br>i<br>n<br>g                | - | 容<br>易      | 学<br>习<br>器 | 合<br>适      |
| G<br>B<br>D<br>T                | 比<br>较<br>差 | 敏<br>感      | 有 | 不<br>可<br>以 | b<br>o<br>o<br>s<br>t<br>i<br>n<br>g | - | 不<br>容<br>易 | 能           | 不<br>合<br>适 |
| X<br>G<br>B<br>o<br>o<br>s<br>t | 比<br>较<br>差 | 不<br>敏<br>感 | 有 | 可<br>以      | b<br>o<br>o<br>s<br>t<br>i<br>n<br>g | - | 不<br>容<br>易 | 能           | 合<br>适      |