

---

# MindWare

发布 *beta*

Thomas (Yang) Li

2022 年 01 月 26 日



1 概览 3

2 安装指南 5

2.1 1 系统要求 . . . . . 5

2.2 2 预先准备 . . . . . 5

2.3 3 安装 MindWare . . . . . 6

2.4 4 进阶功能安装（可选） . . . . . 6

2.5 5 疑难解答 . . . . . 7

3 快速入门 9

3.1 数据准备 . . . . . 9

3.2 优化 . . . . . 10

4 使用实例 11

4.1 Single-Objective Black-box Optimization . . . . . 11

5 研究成果 13

5.1 会议论文 . . . . . 13

6 更新历史 15



Coming soon!

---



# CHAPTER 1

---

概览

---

To be filled





## 2.1 1 系统要求

安装要求：

- Python  $\geq$  3.6 （推荐版本为 Python 3.7）

支持系统：

- Linux (Ubuntu, ...)
- macOS
- Windows

## 2.2 2 预先准备

我们**强烈建议**您为 MindWare 创建一个单独的 Python 环境，例如通过 [Anaconda](#):

```
conda create -n mindware python=3.7
conda activate mindware
```

我们建议在安装 OpenBox 之前通过以下命令更新 pip 和 setuptools:

```
pip install pip setuptools --upgrade
```

## 2.3 3 安装 MindWare

### 2.3.1 3.1 使用 PyPI 安装

只需运行以下命令：

```
pip install mindware
```

### 2.3.2 3.2 从源代码手动安装

使用以下命令通过 Github 源码安装 MindWare：

```
git clone https://github.com/thomas-young-2013/mindware.git && cd mindware
cat requirements/main.txt | xargs -n 1 -L 1 pip install
python setup.py install
```

### 2.3.3 3.3 安装测试 (TODO)

运行以下代码以测试您安装是否成功：

```
from mindware import run_test

if __name__ == '__main__':
    run_test()
```

如果成功，将输出以下信息：

```
===== Congratulations! All trials succeeded. =====
```

如果您在安装过程中遇到任何问题，请参考 [疑难解答](#)。

## 2.4 4 进阶功能安装（可选）

如果您想使用更高级的功能，比如在超参数优化过程中使用 `pyrfr`（概率随机森林）作为代理模型，请参考 [Pyrfr 安装教程](#) 并安装 `pyrfr`。

## 2.5 5 疑难解答

如果以下未能解决您的安装问题, 请在 [Github](#) 上提交 [issue](#) 或发送邮件至 [liyang.cs@pku.edu.cn](mailto:liyang.cs@pku.edu.cn).

如果您未能正确安装 `openbox`, 您也可以参考 [OpenBox](#) 安装指南.

### 2.5.1 Windows

- ‘Error: [WinError 5] 拒绝访问’. 请使用管理员权限运行命令行, 或在命令后添加 `--user`.
- ‘ERROR: Failed building wheel for ConfigSpace’. 请参考 [提示](#).

### 2.5.2 macOS

- 对于 macOS 用户, 如果您在安装 `pyrfr` 时遇到了困难, 请参考 [提示](#).
- 对于 macOS 用户, 如果您在编译 `scikit-learn` 时遇到了困难. [该教程](#) 或许对你有帮助.



本教程将指导您运行第一个 **MindWare** 程序。

## 3.1 数据准备

首先，为端到端自动化机器学习系统 **准备数据**。这里我们用 sklearn 中的 iris 数据集。

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split

iris = load_iris()
X, y = iris.data, iris.target
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33, random_
↪state=1, stratify=y)
```

在导入并切分数据集后，使用 **DataManager** 包装数据。

```
from mindware.utils.data_manager import DataManager

dm = DataManager(X_train, y_train)
train_data = dm.get_data_node(X_train, y_train)
test_data = dm.get_data_node(X_test, y_test)
```

## 3.2 优化

**MindWare** 提供了一种便捷的、端到端的方式，来完成机器学习任务。在这个例子中，我们使用 **Classifier** 来解决分类任务。请指定优化的时间约束 **time\_limit**，然后只需调用 **Classifier.fit**，系统就会自动执行特征工程、模型选择、超参数优化和模型集成过程。对于大搜索空间，系统还会通过搜索空间分解来加速优化过程。

```
from mindware.estimators import Classifier

clf = Classifier(time_limit=3600)
clf.fit(train_data)
```

优化结束后，调用 **Classifier.predict** 来获取自动化机器学习系统对于测试数据集的预测结果。

```
pred = clf.predict(test_data)
```

(todo: result and visualization)

#### 4.1 Single-Objective Black-box Optimization

To be filled





#### 5.1 会议论文

1. **VolcanoML: Speeding up End-to-End AutoML via Scalable Search Space Decomposition** Yang Li, Yu Shen, Wentao Zhang, Jiawei Jiang, Bolin Ding, Yaliang Li, Jingren Zhou, Zhi Yang, Wentao Wu, Ce Zhang and Bin Cui International Conference on Very Large Data Bases (VLDB 2021).
2. **Efficient Automatic CASH via Rising Bandits** Yang Li, Jiawei Jiang, Jinyang Gao, Yingxia Shao, Ce Zhang and Bin Cui Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2020). <https://ojs.aaai.org/index.php/AAAI/article/view/5910>
3. **MFES-HB: Efficient Hyperband with Multi-Fidelity Quality Measurements** Yang Li, Yu Shen, Jiawei Jiang, Jinyang Gao, Ce Zhang and Bin Cui Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2021). <https://arxiv.org/abs/2012.03011>
4. **MindWare: A Generalized Black-box Optimization Service** Yang Li, Yu Shen, Wentao Zhang, Yuanwei Chen, Huaijun Jiang, Mingchao Liu, Jiawei Jiang, Jinyang Gao, Wentao Wu, Zhi Yang, Ce Zhang and Bin Cui ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD 2021). <https://arxiv.org/abs/2106.00421>



# CHAPTER 6

---

更新历史

---

Coming Soon!