

项目经验

2017.01 – 2017.06 脑电波信号控制机器人

- 利用 MATLAB 分析脑电波数据，通过机器学习线性判别分析方法处理与训练数据
- 验证机器学习结果，编写 Python 程序实时操控机器人

2017.09 – 2017.10 [森林类别分析](#)

- 利用 Python 可视化工具 Matplotlib, Seaborn 对森林的各种特征进行提取与分析
- 用 Sklearn 机器学习包的算法随机森林, Extra Tree 对数据进行训练和预测
- 分析误差矩阵来评估和优化机器学习算法

2017.10 – 2017.11 [Kaggle 房价预测竞赛](#)（前 16%）

- 用 Python Pandas 对数据进行清理，构造特征向量，结合线性回归及 Xgboost 机器学习方法提高预测房价的精确度

2017.11 – 2017.12 [Kaggle Mercari 电商价格预测竞赛](#)（前 20%）

- 用 Python NLTK, 正则表达式, 和 Sklearn 清理及向量化文字数据并利用线性回归预测商品价格

知识技能

1. 掌握 Python 语法，熟练使用 Pandas, Numpy, Scipy 进行特征提取及分析
2. 掌握 SQL 语言，在 MySQL 环境下对数据进行增删查改
3. 了解机器学习基本算法，如 SVM、决策树、随机森林、线性回归、K 近邻、朴素贝叶斯
4. 掌握 Linux 基本操作
5. 能利用 Python Matplotlib, Seaborn, 及 Power BI 可视化数据
6. 掌握 Java 基本语法，熟练使用 Eclipse