具有知识约束的机器学习在露天钙钛矿太阳能电池制造过程优化中的应用

# 介绍

钙钛矿是什么。扩大工艺是关键。尽管旋涂效率超过25%，但无法用到生产线。有人提出rspp技术。Rspp有优点。

但Rspp需要调参，高维优化仅靠实验无法解决。

Bo成为有效策略，bo存在成功案例，所以选择bo。

经典bo有两个缺点，1.无法整合相关信息。2.无法引入人员反馈。

所以，领域知识作为概率约束引入，dfe计算也作为概率约束，加速优化。

而且，人员反馈具有意义，定义概率约束是个好方法。

本工作开发概率约束bo，优化钙钛矿pce。流程如图所示，进行介绍。

为了证明本算法能力，目标超过目前最佳pce。1.证明可以提高效率，效率达到18.5%。2.描述概率约束方法。3.分析输入输出之间关系。4.和传统方法比较，虚拟数据中进行优化模拟。

# 结果和讨论

## 提高效率的顺序学习过程优化

## 将知识约束纳入贝叶斯优化

## 样本分布和参数相关性的迭代可视化

## 利用虚拟优化设计实验方法进行基准测试

# 总结

为自动驾驶实验室的数据驱动实验计划配备语义记忆：化学反应优化中的迁移学习案例研究

# 介绍

数据驱动方法能探索复杂参数空间，引起人们兴趣。Bo能对黑盒函数进行有效优化。Bo包括两部分，代理模型和获取函数。

Bo最近代理自动科学平台而受欢迎，如sdlabs。Bo驱动科学有不少案例。

尽管bo成功，但sdlabs针对孤立优化。没有先验信息下，bo初始只能随机采样。人类专家靠经验迁移，迁移优点。对迁移过程模拟是机器学习一大挑战。

本工作提出semopt框架，实现知识迁移。Semopt介绍。

工作流程如下：第二节是问题设置，第三节是相关研究，第四节是本方法介绍。第五节在基准函数上和其他方法对比。最后，化学反应数据上的优化对比。

# 问题设置

通用优化框架介绍

# 相关工作

1. 人类专家知识迁移，存在两个缺点。
2. 代理活动知识迁移，代理活动不太准确，本质与高保真度贝叶斯优化有关。
3. 先前活动知识迁移，先前活动可以从网上收集，也可以自己制作。

后两者有本质区别。

ML社区已有几种先前活动知识迁移的策略。1.MTBO，源活动数据丰富便无法实际使用。**提出解决方法，每个源活动单独设置gp，加入动态加权协议，通过i），更新模型。ii）更新获取函数，实现信息整合**。2.热启动方法。3.神经网络参数共享。

上述方法用于超参数优化，而我们要用于化学实验中。

# 方法

# 基准分析

# 钯催化反应模拟

# 讨论和总结

zombi工作思路

# 介绍

首先介绍了两种优化问题，平滑优化和“针”类优化，针类优化问题具有普遍性，而普通GP不能解决针类优化。

对于针类优化存在三个挑战，前人的方法不能同时应对这三种挑战，因此，作者提出了新算法，并将该算法和标准bo比较，在三种应用场景中和另外三种算法比较。

# 方法

该论文主要有两个贡献，zombi算法和自适应获取函数，以用来解决三个问题，并和其他算法在几个例子上进行比较。然后又测试zombi的适用性，通过144个数据集测试成功条件，在另外30个数据集补充条件后再评估

* 1. zombi算法
  2. 自适应获取函数

# Zombi机制演示

* 1. 范围缩放
  2. 内存修剪
  3. 防鸽子洞

# 实验

4.1，4.2，4.3，4.4介绍了四种问题，将zombi和其他算法比较。

# 总结

# 利用大型材料数据集中的冗余度，以较少的数据实现高效的机器学习

# 介绍

数据对机器学习很重要。通常认为数据稀缺，但大数据时代正在兴起，数据量会加速增长。

与收集数据相比，数据丰富性没人注意。但这是重要的，因为数据冗余可能会降低采样效率。

消除冗余可提高模型效率。一方面大数据会花费算力和时间，另一方面在图像工作中，小数据表现与整个数据集相当。材料学中还没有数据冗余研究，揭示数据冗余有很多优点。

数据冗余在主动学习算法中也很重要。

本工作首先基于神经网络性能退化对大型数据集冗余做了研究。其次提出修剪算法，最后对不确定性主动学习策略和修剪算法进行比较。