**《项目进展汇报》**

# **一、第 [1] 周汇报（日期：[2025.6.23]-[2025.6.27]）**

1. **小组信息**

**姓名** ：唐正阳（组长）

**学号** ：20222241467

**姓名** ：陈嘉炜（组员）

**学号** ：20222241489

**姓名** ：何佩霖（组员）

**学号** ：20222241482

**姓名** ：崔忠延（组员）

**学号** ：20222241018

1. **项目概述**

**项目名称** ：基于AlphaZero 算法的五子棋自我对弈强化学习系统

**项目目标** ：实现融合AlphaZero算法与蒙特卡洛树搜索（MCTS）的五子棋 AI 系统，通过纯自玩训练构建策略-价值网络，使 AI 具备自主学习能力并达到专业水平，同时开发人机交互界面支持实战验证。

1. **本周工作内容**

**理论学习情况**

* 初步了解AlphaZero 自博弈训练闭环机制，学习 "策略网络指导 MCTS 搜索→MCTS 生成高质量数据→数据优化神经网络" 的迭代逻辑。了解AlphaZero算法的具体实现过程。
* 学习蒙特卡洛树搜索（MCTS），自我理解：MCTS是一种基于模拟的搜索方法，通过选择、扩展、模拟和回溯四个阶段迭代优化决策。在AlphaZero中，MCTS结合卷积神经网络指导搜索，替代传统随机模拟，提升效率。初步了解MCTS的核心流程（如节点选择、PUCT算法），对于具体参数的理解以及使用仍然有待提高。
* PyTorch框架，理解：利用PyTorch的动态计算图和自动微分机制便于实现自定义神经网络。初步了解PyTorch框架，学习如何利用框架构建神经网络，学习如何构建自定义神经网络模块与训练循环。

**项目进展**

1. 环境搭建

完成Python环境，PyTorch环境的安装。安装依赖库：numpy, Pygame。

问题：出现了PyTorch等框架和依赖库版本不兼容的情况。解决办法：通过查询相关的博客文章，对依赖库等的版本进行调整。

1. 游戏逻辑功能实现

对于五子棋游戏逻辑进行初步实现，完成人机交互界面的初步设计。

1. 算法实现

初步了解MCTS算法以及AlphaZero的自我博弈过程，了解神经网络构建的相关知识。开始设计策略价值网络的基本结构，计划采用双头CNN架构来分别输出落子概率和局面评估值，对于神经网络的基本结构进行设计。

1. **下周计划**

* 学习卷积神经网络（CNN）在棋盘游戏中的特征提取逻辑，设计棋局状态的张量表示方法（当前玩家、对手棋子、历史落子）等的表示。学习神经网络的模型评估方法。
* 构建CNN架构并实现损失函数组合（策略损失 + 价值损失），实现损失函数等的实现。
* 完成游戏的逻辑功能，实现Game类实现，并对交互界面进行优化设计。
* 实现 MCTS 节点类的实现，包括状态属性、父节点与子节点引用，访问次数、累计价值、先验概率等统计量。
* 实现MCTS搜索过程中的选择阶段的实现。

可能遇到的问题：五子棋规则判定存在漏洞（如斜向连珠未正确检测）。CNN构建出现错误，蒙特卡洛搜索算法实现过程中出现构建失败，对于算法理解程度不够的情况。解决办法：查询相关资料，继续学习相关的算法。