**《项目进展汇报》**

# **一、第 [2] 周汇报（日期：[2025.6.30]-[2025.7.4]）**

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**

* 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。学习棋盘状态的张量表示等，掌握双头输出（策略头/价值头）的损失函数设计。
* 理解算法原理与代码实现的对应关系。理解四阶段（选择、扩展、模拟、回传）的数学意义

**项目进展**

* 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搜索过程中的选择阶段的实现
* 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。
* 构建CNN架构并尝试实现损失函数组合。

1. **下周计划**

* 蒙特卡洛树搜索（MCTS）的完善与优化，完成MCTS的扩展和回传阶段代码实现。优化选择策略的参数调优，平衡探索与利用。测试MCTS在简单棋局中的搜索效率，确保节点扩展和回溯逻辑正确。
* 将当前CNN模型与MCTS结合，实现策略-价值网络的初步训练，将MCTS自我对弈生成的棋局数据用于CNN训练。
* 构建自玩训练循环，使用当前CNN指导MCTS生成对弈数据。定期从经验池采样数据更新CNN参数。