Project Proposal: *"Could multi-agent environments foster the emergence of empathy-like behaviors if agents are trained to care about each other's rewards?"*

Context Overview
In multi-agent settings, artificial systems begin to mirror aspects of human interaction.
This project explores whether introducing empathy-like objectives can promote cooperation and higher collective welfare among learning agents.

Research Focus (Problem to be Addressed)
In multi-agent reinforcement learning, agents are typically designed to maximize only their own rewards.
This raises the question: can introducing empathy-like reward sharing encourage cooperative behaviors and improve overall system welfare?

Proposed Solution
Develop a multi-agent reinforcement learning simulation where agents are trained under two regimes:

1. Selfish agents: optimize only personal rewards.
2. Empathic agents: partially value others' rewards.

- Compare agents' performance and aggregate welfare outcomes in cooperative board-game-like settings (e.g. prisoner's dilemma, stag hunt, resource sharing).

Math / ML Topics Involved

- Reinforcement Learning: define reward functions and then apply Q-learning (a reinforcement learning algorithm which means it teaches an agent how to act in an environment by trial and error, aiming to maximize long-term rewards). + policy gradient (optional ~ more advanced)
- Game Theory Strategy:
    - Start with payoff matrices → build environment
    - Continue with Nash equilibrium → explain how standard self-interest leads to a stable but inefficient outcome
    - Discussion on social welfare → show that empathy increases overall outcomes beyond the Nash prediction
- Simulation: Multi-agent training dynamics & exploration vs exploitation.
    - Multi-agent training dynamics: studying how multiple agents learn and adapt simultaneously.
        - Simulate multiple agents (say, two to four).
        - Observe how their strategies evolve during training:
            - Do they converge to cooperation or competition?
            - How does empathy (reward-sharing) change convergence speed or stability?
        - Visualize this with plots
    - Exploration vs exploitation

- Possibly ε-greedy policy
  - Each agent would use an ε-greedy policy to choose moves.
  - Early on → high ε (agents try random cooperative/competitive actions).
  - Later → lower ε (agents stick to what yields higher combined welfare).

Concrete Output
Report for publication accompanied by a GitHub repository