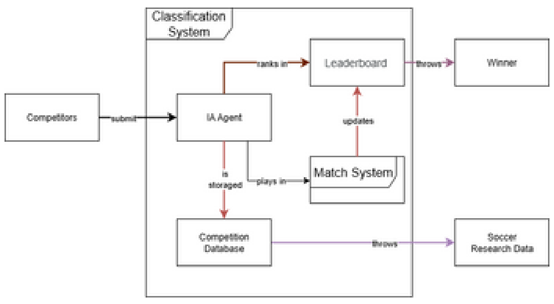
USE OF SYSTEMATIC THINKING IN THE DESIGN AND IMPLEMENTATION OF AN AI AGENT THAT PLAYS FOOTBALL

JUAN DAVID ESCALLON GUZMÁN, JUAN DIEGO LOZANO LUNA, JORGE EDUARDO MUÑOZ GOMEZ

SYSTEMS ANALYSIS
2025-1

INTRODUCTION

The AI Football Competition, organized by Manchester City F.C. and Google Research, challenges participants to develop AI agents capable of playing football in a simulated environment. While the simulation offers structured observations (e.g., player positions, ball dynamics), it lacks real-world unpredictability (e.g., weather, referee errors). Key limitations include perfect simulation conditions and restricted action sets, which hinder the transferability of strategies to real matches. Previous solutions focused on reinforcement learning (RL) in controlled environments.

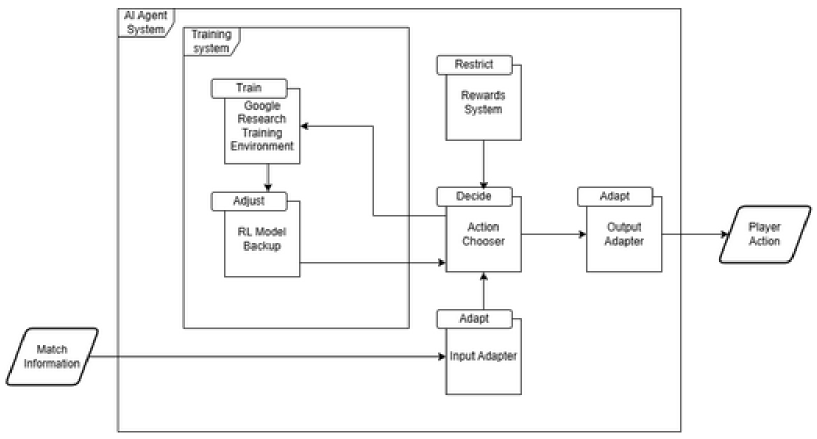


METHODOLOGY & PROPOSED SOLUTION

Architecture:

1. Input Adapter: Standardizes match data (raw).
2. Action Chooser: RL model (PPO algorithm) selects actions from 19 options.
3. Reward System: Customized reward system based on prior analysis..
4. Output Adapter: Formats actions for the Google Research Environment.
5. Backup System: Saves RL models to prevent regression due to chaotic learning.

Technical Stack: Python, TensorFlow (TF-Agents), VirtualBox , GitHub.



An abstraction of the system was made in order to understand it in the simplest possible way, identifying the goal-scoring action as our main focus and how it connects with the entire system. Based on this, a diagram was created to design the initial customized reward system.

GOAL

1. How can an AI agent be designed to perform optimally in a simulated football competition while addressing sensitivity to input formats, chaotic gameplay, and how a specialized reward system affects learning?
2. Expected Product: A scalable RL-based agent with adaptive reward systems, robust to environmental noise and capable of dynamic decision-making.

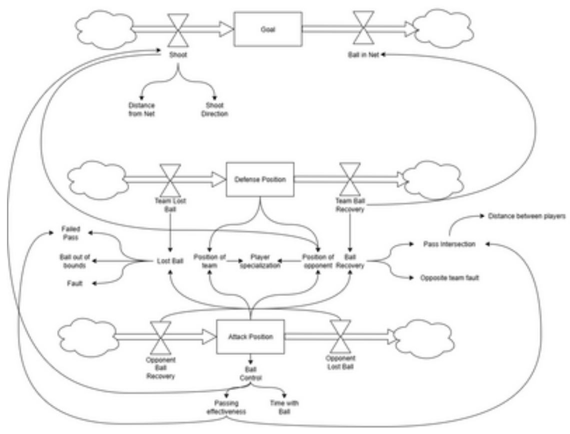
REWARD SYSTEM DESIGN

.Positive Incentives:
- Reward for scoring, shooting on target, short passes, pressuring opponents.

Negative Incentives:
- Penalties for long-distance shots, lost possession, defensive errors.

Contextual Logic:
- Defense: penalize fouls/losses; reward pressure/recovery.
- Attack: penalize long possession, reward short effective passes.

Position-Specific Design:
- Rewards are adapted based on player roles (defender, midfielder, attacker).



RESULTS

The agent demonstrated effective behavior in specific tasks such as passing, high pressing, and scoring. However, its performance lacked overall balance and generalization. In many cases, agents exploited weaknesses in the reward structure, like repeating long passes or remaining in fixed zones. Despite some promising decisions, behavior often became inconsistent in similar situations. Additionally, the expected behavioral differences based on player roles were not significantly observed.

CONCLUSIÓN

The experiment confirmed that the reward system plays a crucial role in shaping agent behavior. While specific improvements were achieved, the agent still lacks adaptability and consistent decision-making across varied scenarios. A promising improvement would be to divide the agent into two specialized networks — one for attack and one for defense — allowing for more context-aware and effective learning.