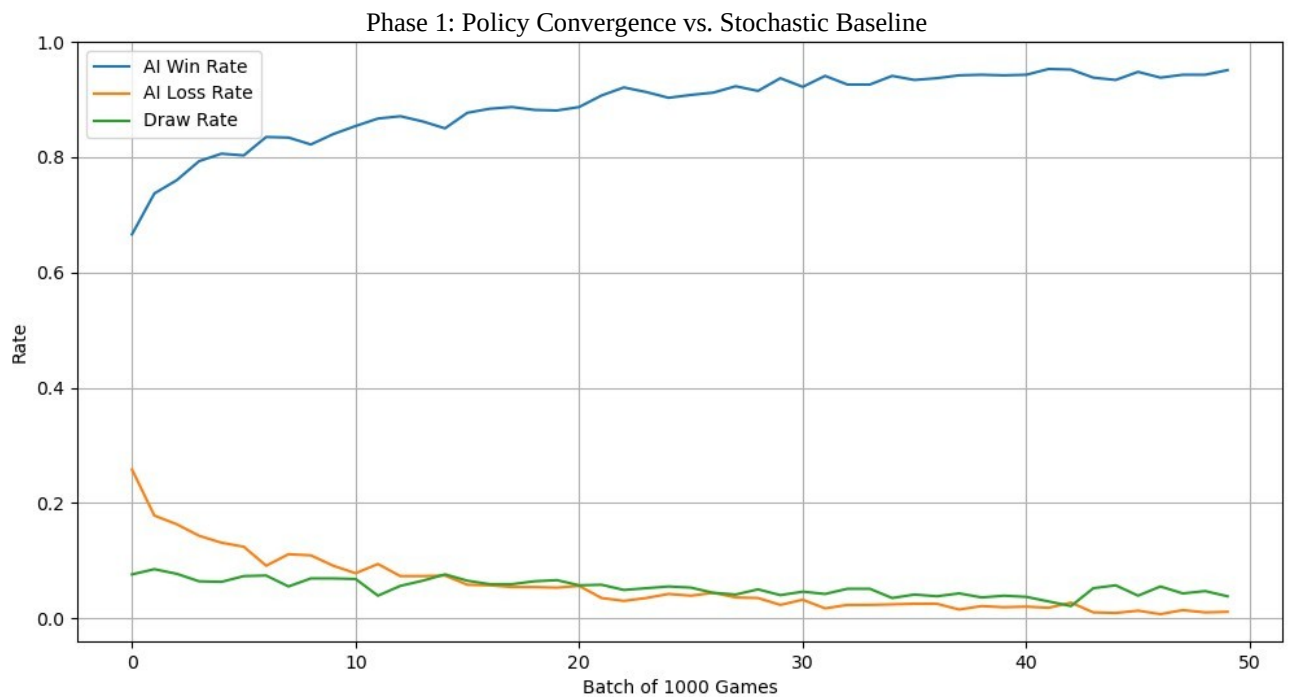
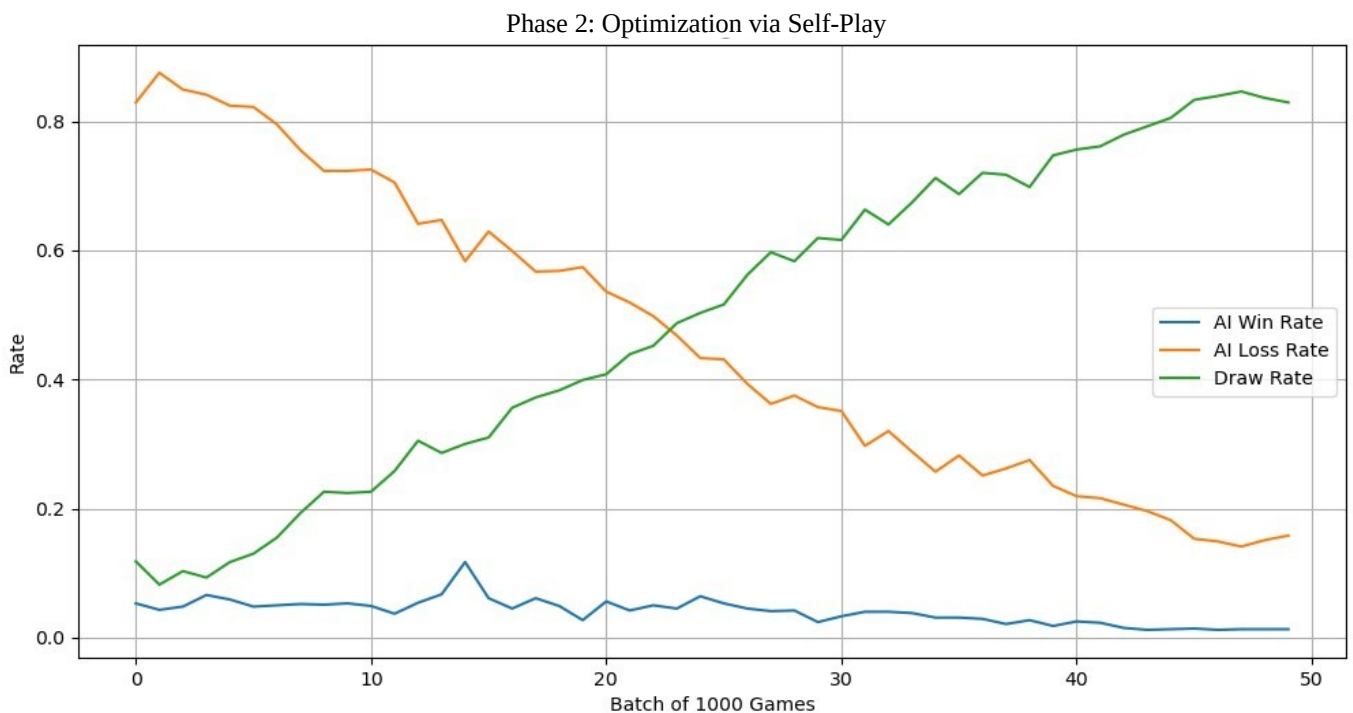


Reinforcement Learning: Solving Tic-Tac-Toe Game with Q-Learning.

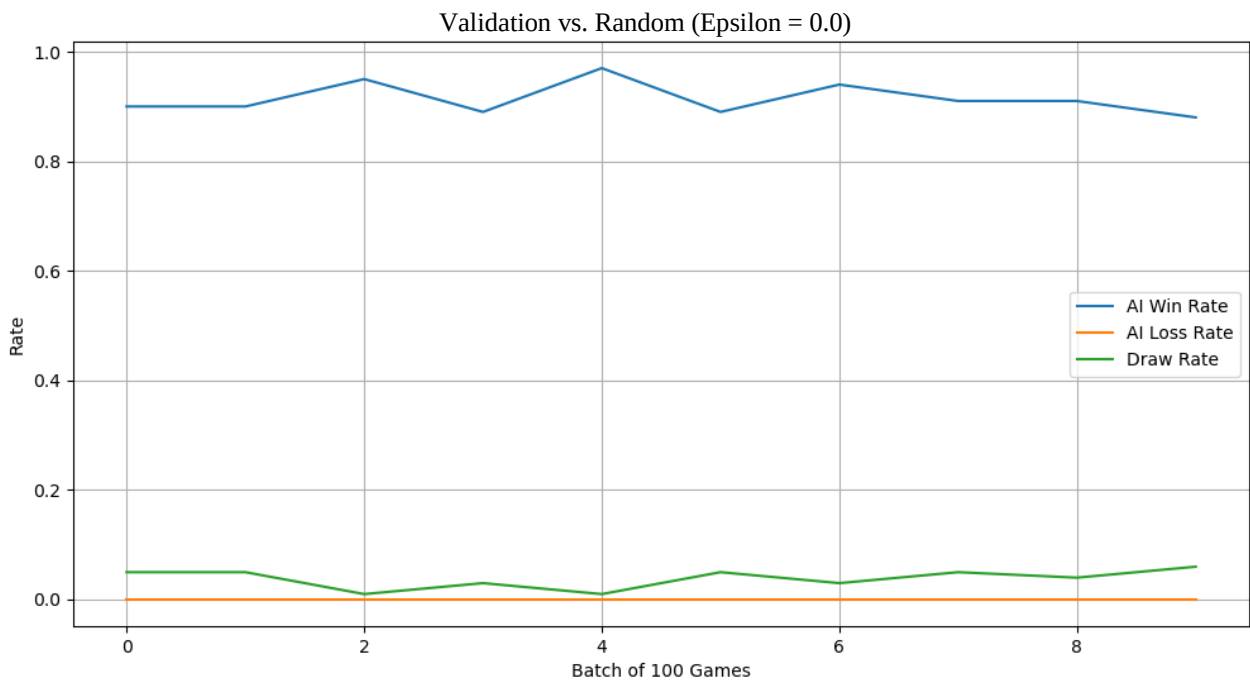
1. Initial Training
2. Convergence to Nash Equilibrium (from untrained Agent_2's perspective)
3. Validation
4. Validation Agent vs. Agent (Nash Equilibrium)



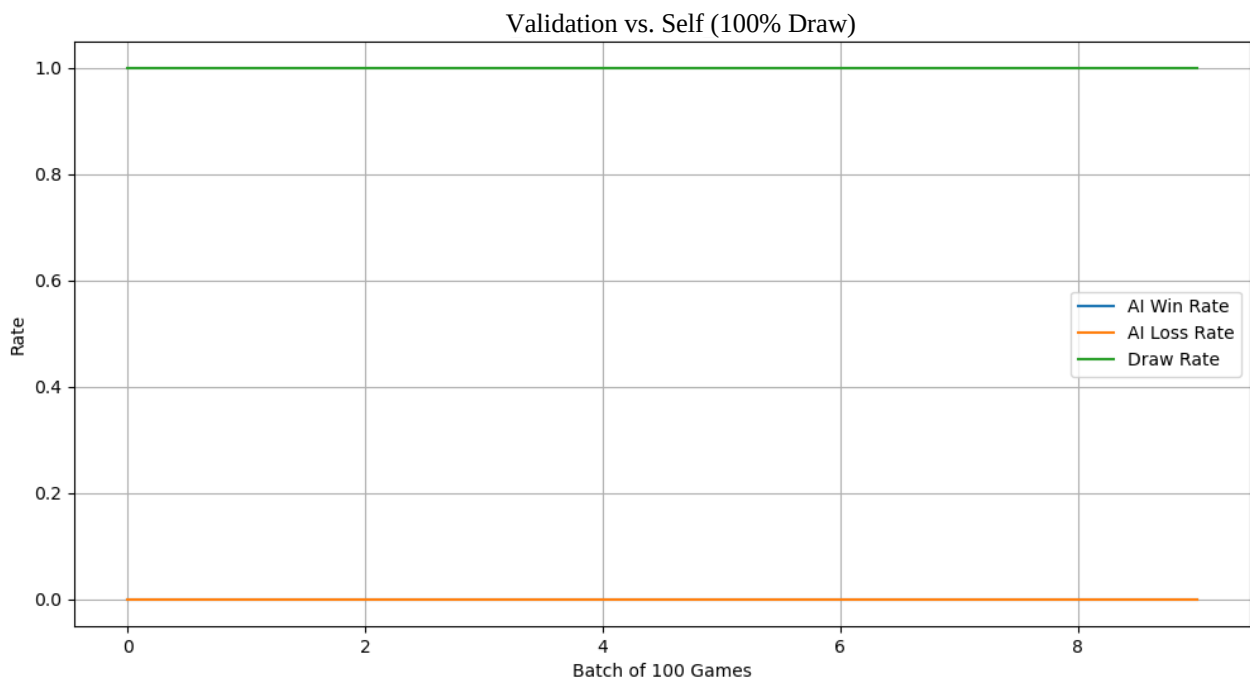
1- Initial Training: Agent quickly exploits a random opponent, reducing loss rate from 40% to <1%.



2- Convergence to Nash Equilibrium (from untrained Agent_2's perspective): As the Agent_2 plays against Agent_1 which has already played against a random bot, the Loss Rate drops and Draw Rate approaches 100% which indicates convergence toward an optimal policy..



3-Validation: The frozen model (Epsilon=0.0) achieves 0 losses in 1000 Games against random bot and forcing a draw in all contested scenarios.



4-Validation Agent vs. Agent (Nash Equilibrium) : Both frozen models (Epsilon=0.0) have perfected the game and are playing the best available moves which is shown through the 100% draw rate from 1000 Games.