Assignment 2

Tic Tac Toe is a two-player game where the goal is to get three of your symbols (X or O) in a row. The game becomes more interesting when humans play against a computer using algorithms like **MiniMax** and **Reinforcement Learning**.

MiniMax is a decision-making algorithm for two-player games. It works by:

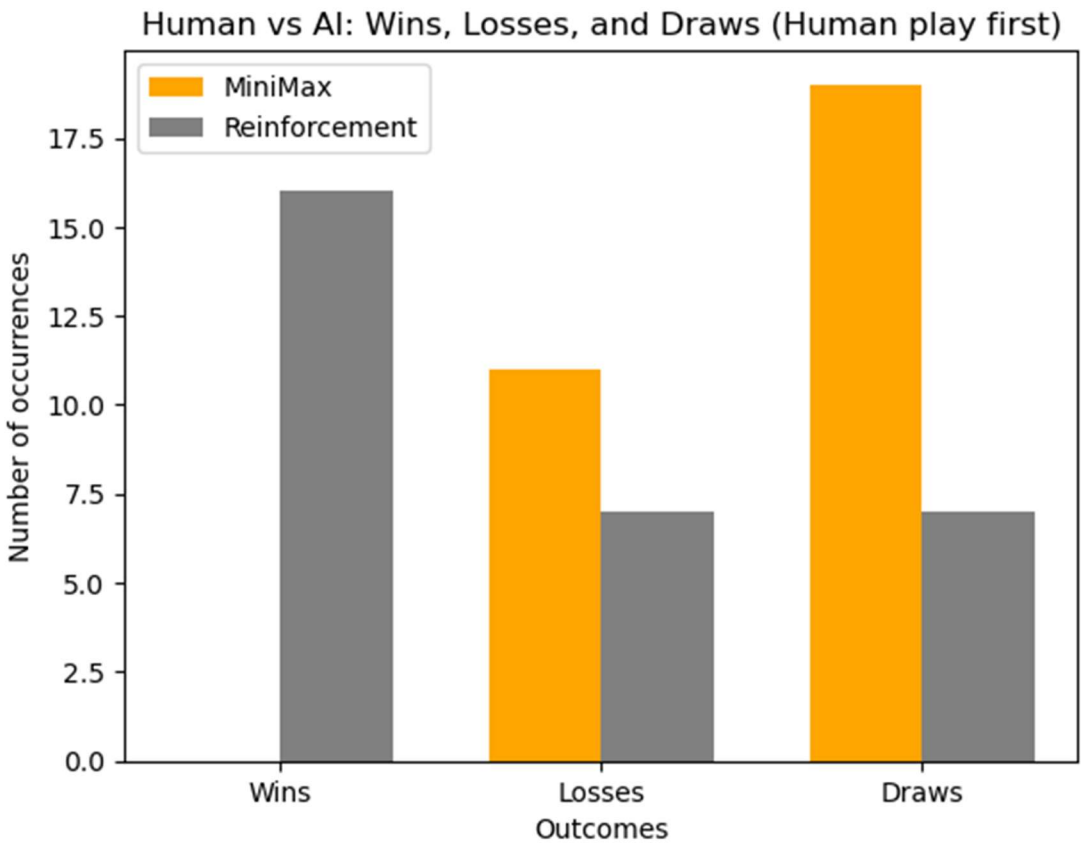
- Considering all possible moves.
- Recursively analyzing the opponent's possible moves.
- Assigning a value to each outcome (+1 for win, -1 for loss, 0 for draw).
- Choosing the move that maximizes the chance of winning or minimizes the chance of losing.

Reinforcement Learning is a method for training AI agents. It works by:

- Setting a goal for the AI agent.
- Allowing the agent to interact with the environment and make decisions.
- Providing feedback in the form of rewards for good decisions and penalties for bad ones.
- Over time, the agent learns to make better decisions to maximize rewards.

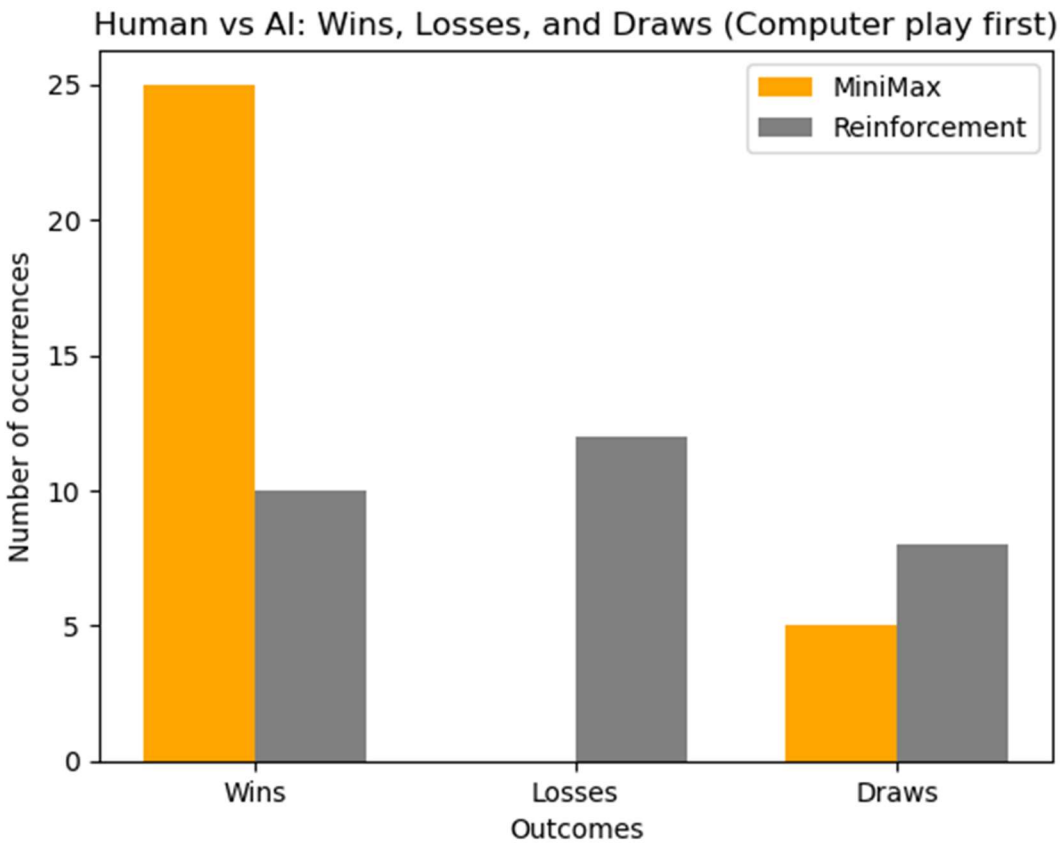
Case I: When the AI plays first

- In this scenario, the MiniMax algorithm, playing optimally, always wins or draws the game, with **25 wins and 5 draws**, and no losses. This is because it always chooses the best possible moves. On the other hand, the Reinforcement Learning algorithm, which is willing to take risks, has a more varied outcome with **8 wins, 10 draws, and 12 losses**.



Case II: When the Human Plays First

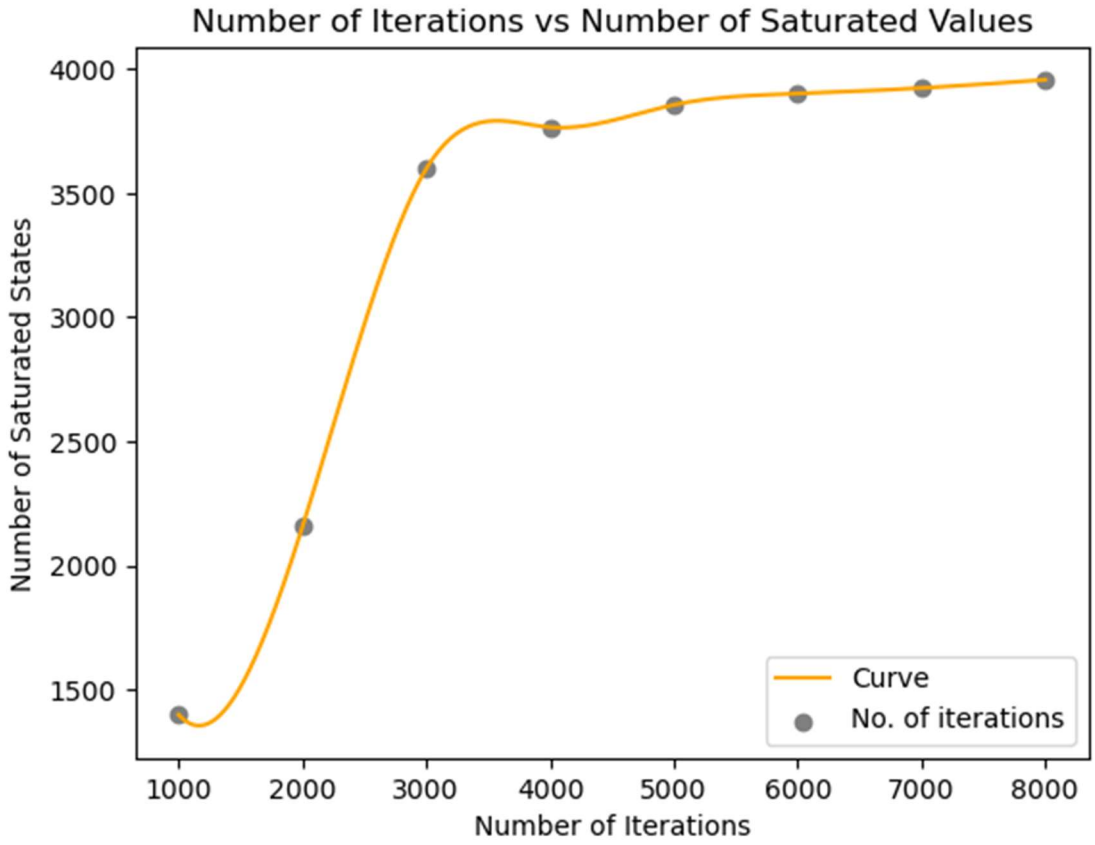
- When the human player makes the first move, the MiniMax algorithm doesn't win any games but manages to draw 12 games and loses 18 games. This is because the first move advantage is with the human player and MiniMax plays defensively. The Reinforcement Learning algorithm, on the other hand, wins 16 games, draws 7, and loses 7. This shows that the Reinforcement Learning algorithm, while still taking risks, can adapt better to the human player's first-move advantage.

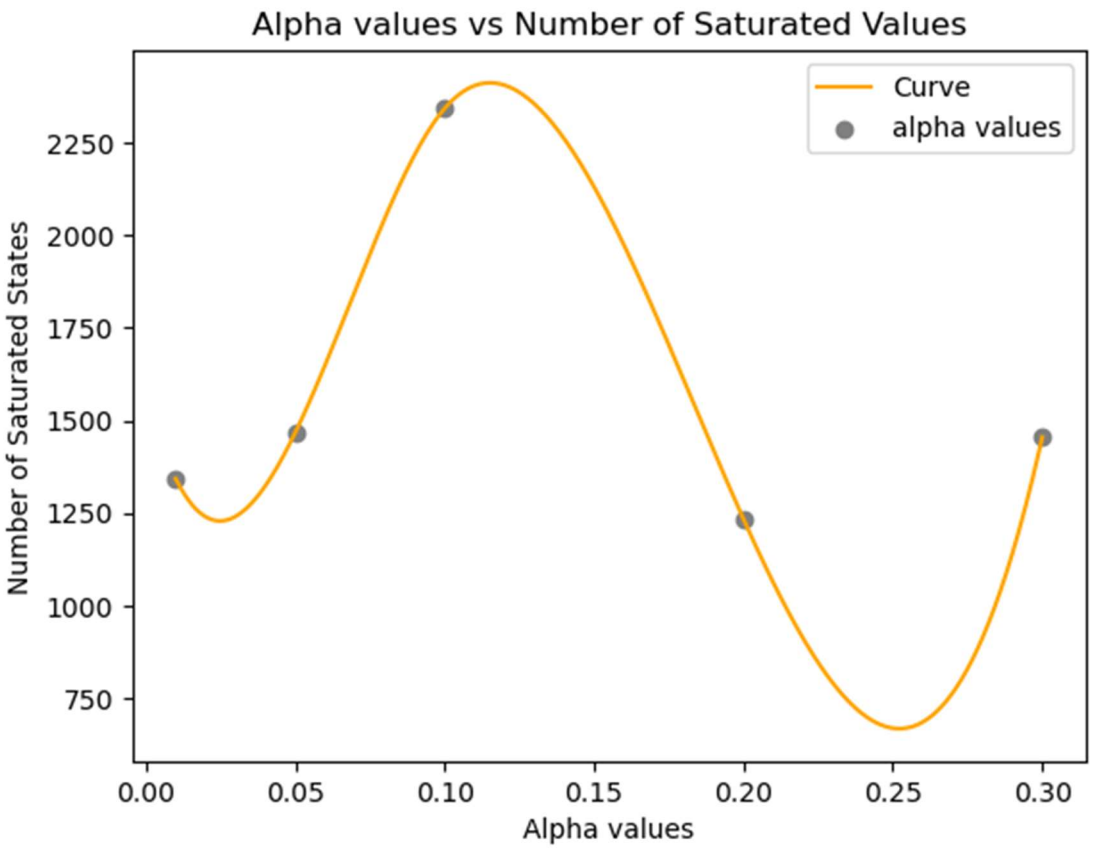


- This analysis shows that while the MiniMax algorithm plays it safe and optimally, the Reinforcement Learning algorithm is more adaptable and can sometimes outperform MiniMax, especially when not having the first move advantage. However, it can also lose more games due to its risk-taking nature. The choice of algorithm can therefore depend on the specific scenario and whether the AI or the human player is making the first move.

Analysis of Saturation Point in Reinforcement Learning:

Fixed-Parameter	Variable Parameter	Values	Observations
Alpha = 0.2	Number of Iterations	1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000	The model shows convergence with increasing iterations, indicating that we are achieving the saturation point in training the Reinforcement Learning model.
Number of Iterations = 2500	Alpha	0.01, 0.05, 0.1, 0.2, 0.3	The best saturation point is achieved at alpha 0.1, suggesting that an alpha of 0.1 is a better choice for reaching the saturation point.




Alpha values vs Number of Saturated Values
Curve
alpha values
Number of Saturated States
2250
2000
1750
1500
1250
1000
750
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Alpha values