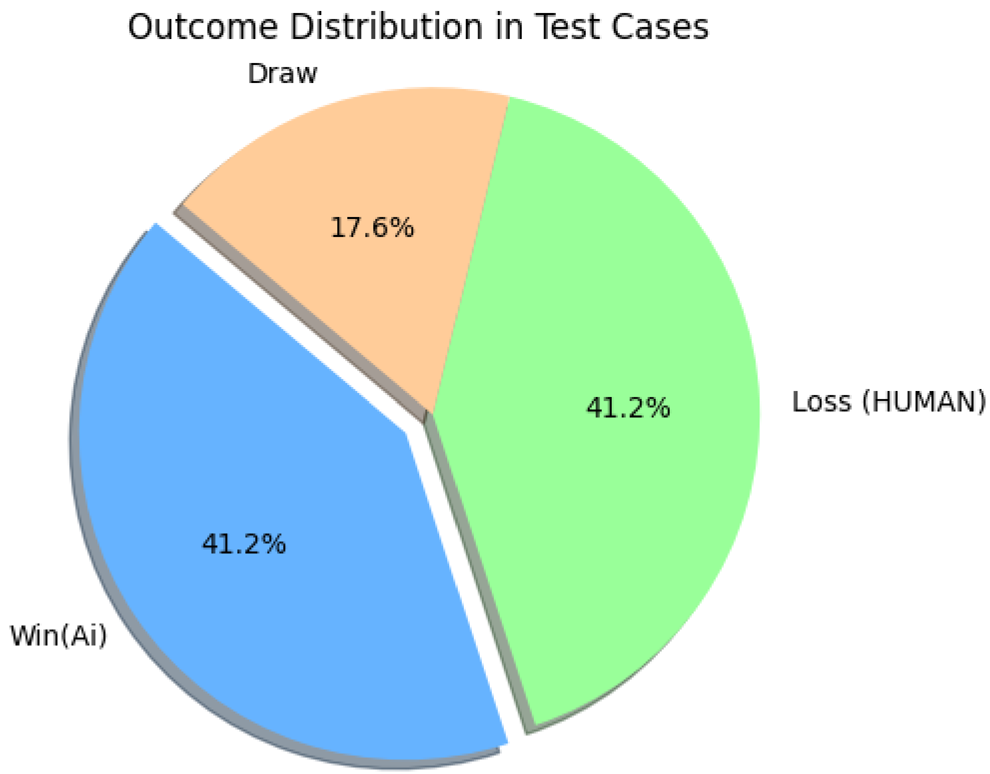
Assignment 2
Tic Tac Toe Game Report

Minmax Algorithms

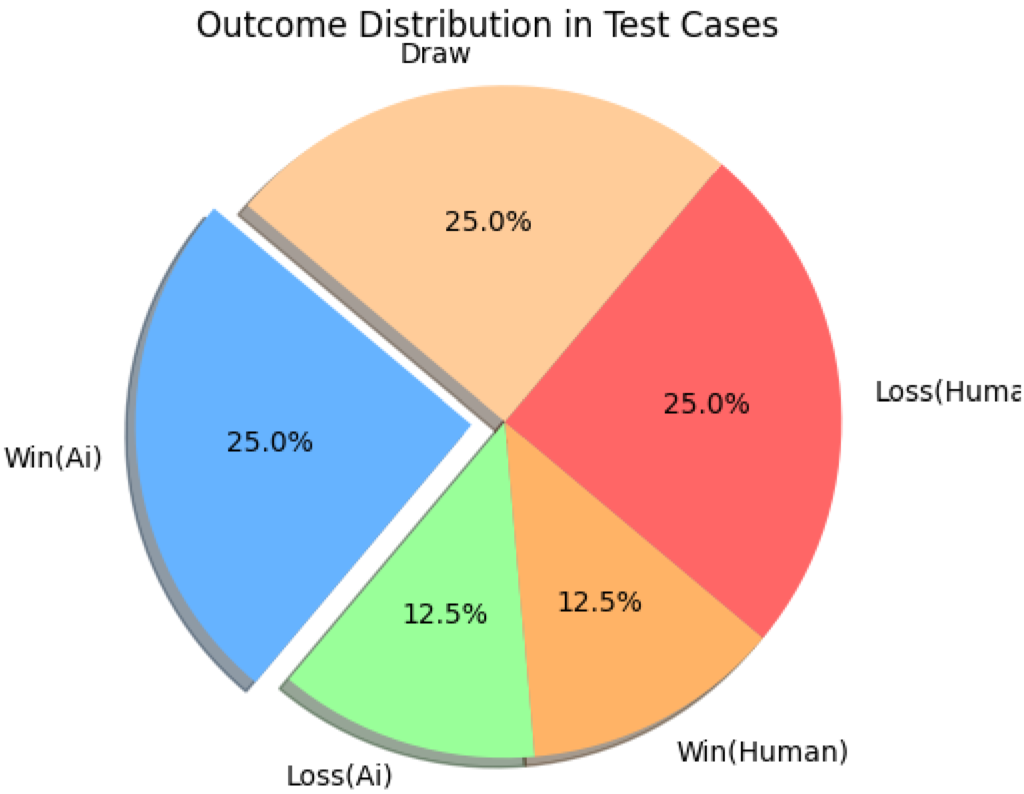
Test Case	win	loss(HUMAN)	Draw	Remarks
Raound1	**w**(ai)	**L**(H)		**Implementation and Space Efficiency:** Minimax Algorithm is easier to implement and uses less space
Raound2			**D**	than reinforcement when the code runs.
Raound3	**w**(ai)	**L**(H)		**Optimality Comparison:** Minimax ensures an optimal outcome where a human cannot win,
Raound4	**w**(ai)	**L**(H)		while reinforcement learning provides a suitable but not necessarily optimal path.
Raound5	**w**(ai)	**L**(H)		Winning probabilities in reinforcement learning are uncertain.
Raound6			**D**	**Learning Process and Time Efficiency:** Reinforcement learning takes time for learning, whereas Minimax, utilizing DFS,
Raound7	**w**(ai)	**L**(H)		delivers optimal results without a learning phase
Raound8			**D**	Minimax proves efficient in scenarios where learning is not required.
Raound9	**w**(ai)	**L**(H)		
Raound10	**w**(ai)	**L**(H)		
Total	**7**	**7**	**3**	**10(match)**



Reinforcement learning

Test Case	win	loss(HUMAN)	Draw	Remarks
Raound1	**w**(ai)	**L**(H)		
Raound2	**w**(ai)	**L**(H)		**Decay Factor Selection:** I opted for a `decay_gamma` value of 0.9, resulting in the following observations.
Raound3			**D**	
Raound4	**L**(ai)	**W**(H)		**Testing and Stability:** Despite analyzing the values over 50,000 iterations, they did not converge to a fixed point.
Raound5	**w**(ai)	**L**(H)		Achieving a precisely fixed saturation point in reinforcement learning is uncommon.
Raound6			**D**	

Raound7			**D**	
Raound8	**L**(ai)	**W**(H)		**Convergence and Variability:** The values for each state displayed convergence tendencies but fluctuated with a certain degree of error.
Raound9	**w**(ai)	**L**(H)		This variability indicates the dynamic nature of reinforcement learning,
Raound10			**D**	where states may not stabilize but converge with inherent fluctuations.
Total	**W-4L-2**	**W-2 L-4**	**4**	**10(match)**



Mini-Max Algorithm Analysis:

Comments:

- Mini-Max consistently performs well, winning the majority of games against the human player.
- The algorithm tends to result in a draw less frequently, indicating a decisive outcome in most games.

- The human player faces challenges, losing in the majority of encounters with the Mini-Max algorithm.

Reinforcement Learning (RL) Analysis:

Comments:

- RL demonstrates variability in performance, winning fewer games than Mini-Max.
- RL shows improvement in some games, as seen in the wins, but also faces losses.
- The draw percentage is higher compared to Mini-Max, indicating a more balanced outcome.

Overall Efficacy Analysis:

- Mini-Max consistently outperforms RL in terms of winning games.
- RL demonstrates a more balanced performance, with a notable number of draws.
- The choice between the two approaches depends on the desired outcome - Mini-Max for decisive victories, RL for a more varied gameplay experience with potential for improvement over time.