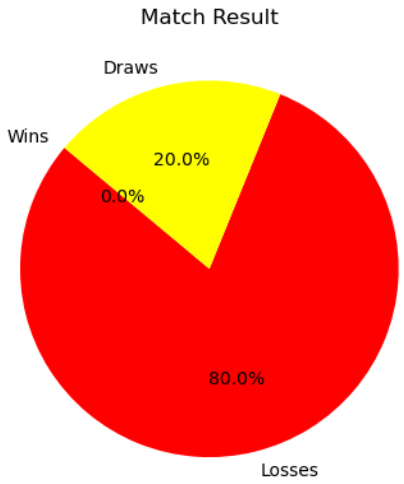
Assignment 2 Report

About assignment:- We need to observe the behavior, pattern and outcomes of a Game known as Tic-Tac-Toe. This game is implemented in two different ways, one using MinMax Algorithm and other one is Reinforcement learning.

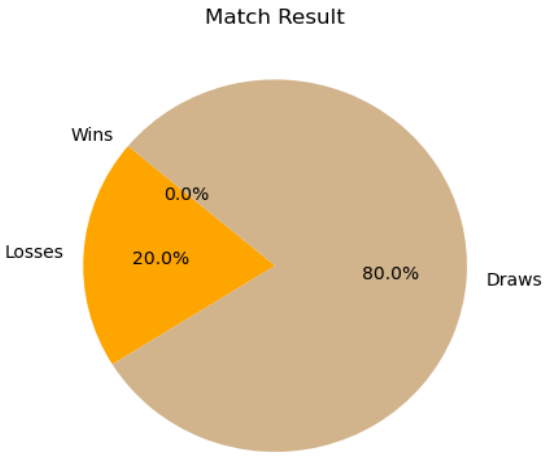
Analysis of the outcomes:-

Below charts are on the basis of human results.

MinMax

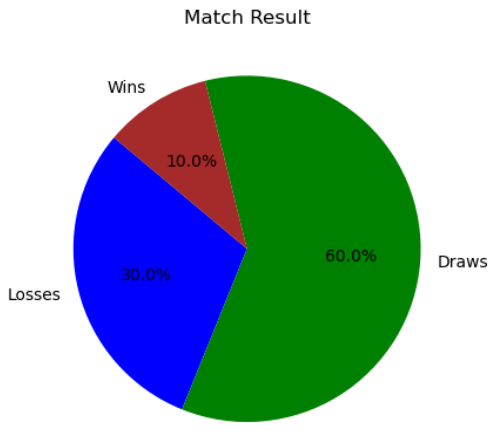


Computer play first

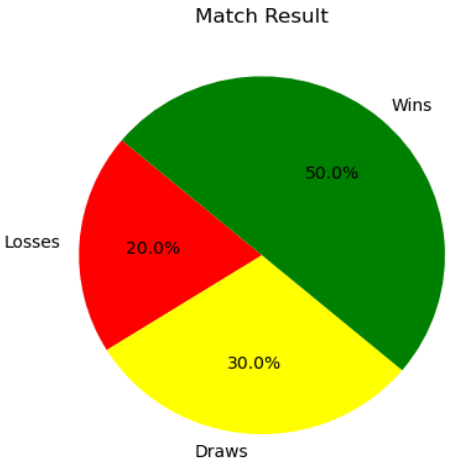


Human play first

Reinforcement Learning



Computer First



Human First

Observations while playing the game multiple times:

MinMax:-

1. Computer play first:-
    a. Most of the time the computer's first move is at a fixed position or last position.
    b. Who is going to win the game can be decided after the two moves of the computer.
    c. Human is not able to win the game.
    d. Game results in computer’s wins.
2. Human play first:-
    a. Most of the time the computer's first move is at a fixed position.
    b. Who is going to win the game can be decided after the two moves of the human.
    c. Most of the time the game results in draws.

Reinforcement:-

1. Computer play first:-
    a. Most of the time games result in computer’s wins or draws and very few human’s win.
    b. In some cases when a computer can win the game, it results in draws because of the computer's move.
2. Human play first:-
    a. Human is able to win the game more than any other way.
    b. In some cases when a computer can win the game, it results in draws because of the computer's move.

Observation While training:-

MinMax:-

- Around 81% percent the value of state in the game is updated.
- Most of the time the positive value is updated.

Reinforcement:-

- I have taken alpha as 0.3104 as I found this value as best for training as per my requirement of the code and saturation point.
- We have to update the value till the saturation point does not occur. Saturation point means when the value of state is not changing or starts converging to one value.
- I have done training in two ways
    - Running the algorithm on the terminal list multiple times.

    - Running the algorithm to a fixed number and randomly selecting a terminal.
- Since I didn't use Exploration in training, learning of the game was not too good.

By which way of training, I would like to play games?

After playing the game multiple times I found that training by Reinforcement learning will be good for me and as a human we always try to win the game and we are able to win the game in this case. But when I was playing the game trained by the MinMax algorithm, I was not able to win the game at any condition. Few times games result in draws.