Week 8 Assignment – Reinforcement Learning

By: Zach Adair

Regis University

MSDS 680 – Machine Learning

# Introduction

For this week’s assignment I will be seeking out a RStudio example of using Reinforcement Learning and describing the reinforcement components of the application. The example I found is for a game of tic-tac-toe and the code will be worked out in the R script attached in the turn in section.

# RStudio Packages

Devtools, ReinforcementLearning, MDPtoolbox

# Objective

The objective our this Reinforcement Learning agent will be to learn to play a game of tic-tac-toe. It will learn how to play based on several steps which are goal oriented with the aim to learn sequences of actions which will lead to its goal or objective function. For this assignment the goal is to learn how to play tic-tac-toe.

# Policy

Policy dictates the correct path to reach the final state. Policy is the mapping of steps to achieve the desired result. In this case it will be mapping the steps to winning a game of tic-tac-toe.

# Reward Function

The reward function is associated with the transition and the goal of the RL agent is to collect as much reward as possible. The agent can pick any action as a function of its history and prior knowledge from previous iterations of the process. Reward in this example will be when the agent wins a game of tic-tac-toe.

# State Transition

The state transition deals with the current situation of the agent, after every iteration that state can change based off what it had just learned and if so the state has now transitioned.

# Exploitation

The exploitation is where the RL agent makes the best decision given the current information it possesses. This can change after every run but this is what it uses to make it’s decision on its next step.

# Exploration

Exploration is where the RL agent gathers more information which will hopefully lead to better decisions in its future. So with every iteration the agent learns more about the game and how to fulfill its reward more often.

# Example Reference

dataaspirant.com/2018/02/05/reinforcement-learning-r/