# Reinforcement Learning Training

1. Day 1 session 1

    a. **Reinforcement learning**
        i. Basic terminologies and conventions
        ii. Optimality criteria
        iii. The value function for optimality
        iv. The policy model for optimality
        v. The Q-learning approach to reinforcement learning
        vi. Asynchronous advantage actor-critic
        vii. Introduction to TensorFlow and OpenAI Gym
        viii. Basic computations in TensorFlow
        ix. An introduction to OpenAI Gym
        x. The pioneers and breakthroughs in reinforcement learning
        xi. David Silver
        xii. Pieter Abbeel
        xiii. Google DeepMind
        xiv. The AlphaGo program
        xv. Libratus

    b. **Environments**
        i. Training Reinforcement Learning Agents Using OpenAI Gym
        ii. The OpenAI Gym
        iii. Understanding an OpenAI Gym environment
        iv. Programming an agent using an OpenAI Gym environment
        v. Q-Learning
        vi. The Epsilon-Greedy approach
        vii. Using the Q-Network for real-world applications

2. Day 1 session 2

    a. **Markov Decision Process**
        i. The Markov property
        ii. The S state set
        iii. Actions
        iv. Transition model
        v. Rewards
        vi. Policy
        vii. The sequence of rewards - assumptions
        viii. The infinite horizons
        ix. Utility of sequences
        x. The Bellman equations
        xi. Solving the Bellman equation to find policies
        xii. An example of value iteration using the Bellman equation
        xiii. Policy iteration
        xiv. Partially observable Markov decision processes
        xv. State estimation

xvi. Value iteration in POMDPs
xvii. Training the FrozenLake-v0 environment using MDP

3. Day 2 session 1

    a. **Policy Gradients**
        i. The policy optimization method
        ii. Why policy optimization methods?
        iii. Why stochastic policy
        iv. Example 1 - rock, paper, scissors
        v. Example 2 - state aliased grid-world
        vi. Policy objective functions
        vii. Policy Gradient Theorem
        viii. Temporal difference rule
        ix. TD(1) rule
        x. TD(0) rule
        xi. TD() rule

4. Day 2 session 2

    a. **Policy Gradients**
        i. The Monte Carlo policy gradient
        ii. Actor-critic algorithms
        iii. Using a baseline to reduce variance
        iv. Vanilla policy gradient
        v. Agent learning pong using policy gradients

    b. **Why reinforcement learning?**
        i. Model based learning and model free learning
        ii. Monte Carlo learning
        iii. Temporal difference learning
        iv. On-policy and off-policy learning

5. Day 3 Session 1

    a. **Q-learning**
        i. The exploration exploitation dilemma
        ii. Q-learning for the mountain car problem in OpenAI gym

    b. **Deep Q-networks**
        i. Using a convolution neural network instead of a single layer neural network
        ii. Use of experience replay
        iii. Separate target network to compute the target Q-values
        iv. Advancements in deep Q-networks and beyond

v. Double DQN
vi. Duelling DQN
vii. Deep Q-network for mountain car problem in OpenAI gym
viii. Deep Q-network for Cartpole problem in OpenAI gym
ix. Deep Q-network for Atari Breakout in OpenAI gym

6. Day 3 Session 2

   a. **Reinforcement Learning Algorithms**
      i. The Monte Carlo tree search algorithm
      ii. Minimax and game trees
      iii. The Monte Carlo Tree Search
      iv. The SARSA algorithm
      v. SARSA algorithm for mountain car problem in OpenAI gym

   b. **Asynchronous Methods**
      i. Why asynchronous methods?
      ii. Asynchronous one-step Q-learning
      iii. Asynchronous one-step SARSA
      iv. Asynchronous n-step Q-learning
      v. Asynchronous advantage actor critic

7. Day 4 Session 1

   a. **Case Study**
      i. Real Strategy Gaming
      ii. Reinforcement Learning in Autonomous Driving
      iii. Financial Portfolio Management
      iv. Reinforcement Learning in Robotics

8. Day 4 Session 2

   a. **Case Study**
      i. Deep Reinforcement Learning in Ad Tech
      ii. Reinforcement Learning in Image Processing
      iii. Deep Reinforcement Learning in NLP