**Prediction and Control with Function Approximation: Learning Objectives**

**Module 00: Welcome to the Course**
Understand the prerequisites, goals and roadmap for the course.

**Module 01: On-policy prediction with approximation**

**Lesson 1: Estimating Value Functions as Supervised Learning**
Understand how we can use parameterized functions to approximate value functions
Explain the meaning of linear value function approximation
Recognize that the tabular case is a special case of linear value function approximation.
Understand that there are many ways to parameterize an approximate value function
Understand what is meant by generalization and discrimination
Understand how generalization can be beneficial
Explain why we want both generalization and discrimination from our function approximation
Understand how value estimation can be framed as a supervised learning problem
Recognize not all function approximation methods are well suited for reinforcement learning

**Lesson 2: The Objective for On-policy Prediction**
Understand the mean-squared value error objective for policy evaluation
Explain the role of the state distribution in the objective
Understand the idea behind gradient descent and stochastic gradient descent
Outline the gradient Monte Carlo algorithm for value estimation
Understand how state aggregation can be used to approximate the value function
Apply Gradient Monte-Carlo with state aggregation

**Lesson 3: The Objective for TD**
Understand the TD-update for function approximation
Highlight the advantages of TD compared to Monte-Carlo
Outline the Semi-gradient TD(0) algorithm for value estimation
Understand that TD converges to a biased value estimate
Understand that TD converges much faster than Gradient Monte Carlo

**Lesson 4: Linear TD**
Derive the TD-update with linear function approximation
Understand that tabular TD(0) is a special case of linear semi-gradient TD(0)
Highlight the advantages of linear value function approximation over nonlinear
Understand the fixed point of linear TD learning
Describe a theoretical guarantee on the mean squared value error at the TD fixed point

## Module 02: Constructing Features for Prediction

### Lesson 1: Feature Construction for Linear Methods

Describe the difference between coarse coding and tabular representations
Explain the trade-off when designing representations between discrimination and generalization
Understand how different coarse coding schemes affect the functions that can be represented
Explain how tile coding is a (computationally?) convenient case of coarse coding
Describe how designing the tilings affects the resultant representation
Understand that tile coding is a computationally efficient implementation of coarse coding

### Lesson 2: Neural Networks

Define a neural network.
Define activation functions.
Define a feedforward architecture
Understand how neural networks are doing feature construction
Understand how neural networks are a non-linear function of state
Understand how deep networks are a composition of layers.
Understand the tradeoff between learning capacity and challenges presented by deeper networks

### Lesson 3: Training Neural Networks

Compute the gradient of a single hidden layer neural network
Understand how to compute the gradient for arbitrarily deep networks
Understand the importance of initialization for neural networks
Describe strategies for initializing neural networks
Describe optimization techniques for training neural networks

## Module 03: Control with Approximation

### Lesson 1: Episodic Sarsa with Function Approximation

Explain the update for Episodic Sarsa with function approximation
Introduce the feature choices, including passing actions to features or stacking state features
Visualize value function and learning curves
Discuss how this extends to Q-learning easily, since its a subset of Expected Sarsa

### Lesson 2: Exploration under Function Approximation

Understanding optimistically initializing your value function as a form of exploration

**Lesson 3: Average Reward**
Describe the average reward setting
Explain when average reward optimal policies are different from discounted solutions
Understand how differential value functions are different from discounted value functions

**Module 04: Policy Gradient**

**Lesson 1: Learning Parameterized Policies**
Understand how to define policies as parameterized functions
Define one class of parameterized policies based on the softmax function
Understand the advantages of using parameterized policies over action-value based methods

**Lesson 2: Policy Gradient for Continuing Tasks**
Describe the objective for policy gradient algorithms
Describe the results of the policy gradient theorem
Understand the importance of the policy gradient theorem

**Lesson 3: Actor-Critic for Continuing Tasks**
Derive a sample-based estimate for the gradient of the average reward objective
Describe the actor-critic algorithm for control with function approximation, for continuing tasks

**Lesson 4: Policy Parameterizations**
Derive the actor-critic update for a softmax policy with linear action preferences
Implement this algorithm
Design concrete function approximators for an average reward actor-critic algorithm
Analyze the performance of an average reward agent
Derive the actor-critic update for a gaussian policy
Apply average reward actor-critic with a gaussian policy to a particular task with continuous actions