DRL Contents

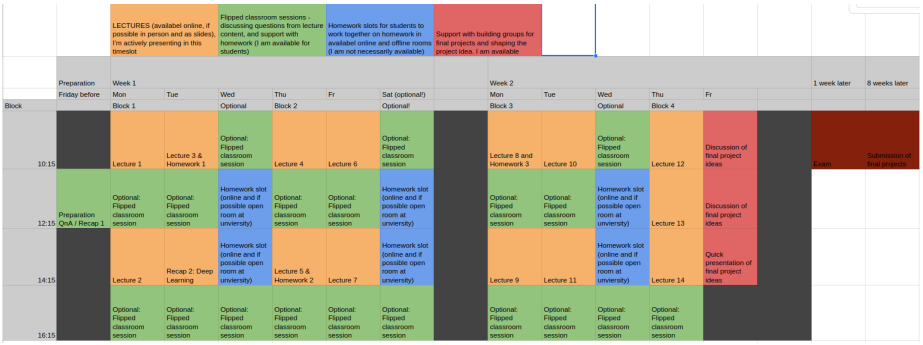
lschmid

July 2020

1 Course Prerequisites

Students participating in the course are expected to be able to implement Deep Learning models on their own (preferably in tensorflow). Furthermore really only high-school level maths is needed. More background in AI and Machine Learning (specifically also RL) is useful, but not necessary. Ability to set up working environment (conda, python, tensorflow) is necessary (but one can also use colab if there are any issues).

2 Content



- Students are free to learn contents from either live lectures (if possible), or courseware (recorded online lectures + blog-like online ressources/available slides)
- For each lecture an inverted-classroom-style session is offered, in which students are invited to discuss the questions and material attached to each of the lectures. Furthermore I'm available for homework related questions in these sessions.
- For each homework a dedicated homework day is available, which features a discussion session in the beginning and end of it. In between students have half a day to actually work on their homework implementations together. Students are not required to take part in these sessions.

- The last day is reserved for developing and discussing final project ideas

2.1 Block 1 - Monday-Tuesday Week 1

2.1.1 Lecture 1

Basics of RL and terminology, starting with MDP and ending with gyms and policy iteration / intuitive version of Q-Learning (without much background) (mostly based on Chapters 1-3 from RL - an introduction)

2.1.2 Lecture 2

Basics of Temporal Difference Learning, Dynamic Programming and Monte Carlo Methods (Chapters 3-6 from RL - an introduction)

2.1.3 Lecture 3

Advanced issues in tabular RL - e.g. Importance sampling, advanced SARSA methods (Tree Backup, TD(λ), Q(σ)), possibly Conservative Q-learning

2.1.4 Homework 1

Implementing tabular offline Watkins Q-Learning and online SARSA(1)

2.1.5 Recap 1

Revisiting basic math and basics of Deep Learning necessary for RL (QnA, students expected to be prepared beforehand with a quick "test yourself sheet" available)

2.1.6 Recap 2

Summary of Deep Learning Methods: Architectures (including advanced Architectures), Basic How-Tos of Training Models (SGD, Minibatch, Adam, Regularization, Normalization, etc.)

2.2 Block 2 - Thursday-Friday Week 1

2.2.1 Lecture 4

DQN - basic: ANNs as function approximators for RL

2.2.2 Lecture 5

DQN - advanced issues and optimizations; Exploring why DQN can't solve continuous action spaces

2.2.3 Lecture 6

Deep Deterministic Policy Gradients and it's basic optmizations (TD3)

2.2.4 Lecture 7

Policy Gradient Theorem and Vanilla Policy Gradients

2.2.5 Homework 2

Implementing DQN

2.3 Block 3 - Monday-Tuesday Week 2 [large course only]

2.3.1 Lecture 8

Advantage Actor Critics and Generalized Advantage Estimation (A2C, A3C, GAEs)

2.3.2 Lecture 9

Conservative Policy / Trust Region Methods (Conservative Q-Leraning, TRPO, PPO) and Exploration and Entropy Regularized Reinforcement Learning (SAC)

2.3.3 Lecture 10

Offline Policy Gradients (ACER, Retrace, V-trace/Impala)

2.3.4 Lecture 11

Optional special lecture with tips and tricks for Implementing DRL (Normalization, Regularization, important frameworks)

2.3.5 Homework 3

Implementation of A2C

2.4 Block 4 - thursday-friday week 2 [large course only]

2.4.1 Lecture 12

Diverse I: Model Based RL (e.g. MCTS), World Models

2.4.2 Lecture 13

Diverse II: Multi Agent RL, Competitive Play, Self-Play, Language Emergence, Cooperative Games

2.4.3 Lecture 14

Diverse III: Leftovers. Includes DRL for Real life applications, Active Learning/Active Exploration and more

2.4.4 Final Project

In small groups students build implementations of slightly advanced papers and try to tackle (very small) research questions / problems (e.g. applying an algorithm to a new environment). Has to contain three parts: Implementation, "Lab report", Short summary video. Is graded. Perfect place to also work on external DRL projects. Students can take their time with it (8 weeks+)

2.4.5 Exam

Exam should happen one or two weeks after the lectures finished. Details will still be discussed, especially regarding the issue of offering it online and offering an alternative (more effort for everyone involved - specifically for students with test anxiety and similar conditions).

3 Remark

Please not that any information in this document is preliminary and might be subject to changes - the final statement about the organizational details will only be available once the course has started (but it's quite unlikely anything will deviate too much from this)