SIT796 Reinforcement Learning

Distinction Task 7.1D: Function approximation implementation

Overview

During week 7, you have learnt about function approximation methods. You have also been provided with an example of the implementation of gradient descent in the workshop for the week.

In this task you need to implement the following methods

- Semi-Gradient Sarsa(0)
- Semi-Gradient TD(λ)

and test them on the environment you scoped in Tasks 1.1P. In your report, discuss the differences and similarities between these and how these affect the result. In your discussion, refer to your code. For your results, compare the results for both methods plotting the average rewards over 100 episodes as a function of the necessary number of steps.

To complete this assignment, you need to refer back to Week 7 lecture and practical material.

Submission Details

For this task you need to provide the code and a report on the implementation of the methods above that allows you to achieve the goal. The implementation of the methods has to be from scratch, with the environment itself being that which you scoped in Task 1.1P.

The implementation should be done using a Jupiter notebook. The report and the Jupiter notebook need to be submitted to OnTrack. To do so, Convert the Jupyter Notebook to a **PDF** and append it to your report so as to submit that document. Also include a URL link to a GitHub repo containing the Notebook file and the environment implementation.

You may have to install pandoc to convert a Jupyter Notebook to a PDF document:

http://pandoc.org/installing.html

Constraints

The code used for the environment should be attributable only to the author of the report. The report should have a **high-quality writing style** not exceeding **600 words** in length.