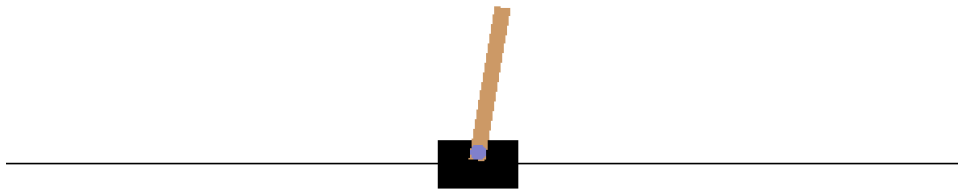Assignment 6.1:

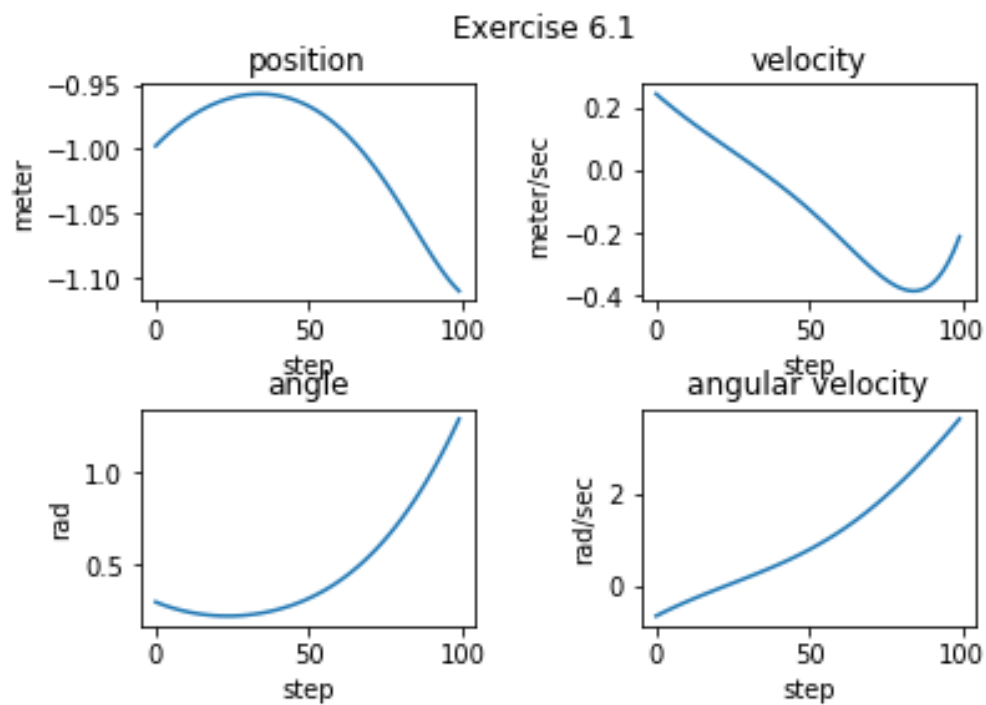Fig – 1: Cart-Pole System Time Simulation

Fig – 2: Change in state of Cart – Pole with time
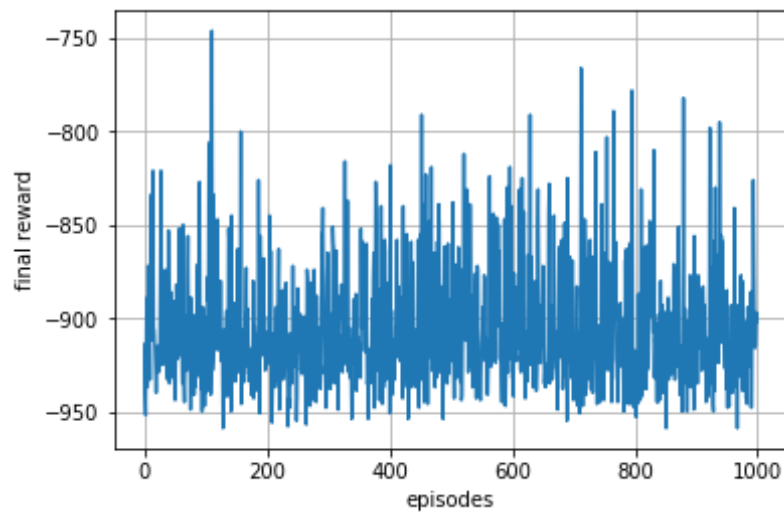
Assignment 6.2

Fig – 3: Accumulated rewards per episodes using state-feedback policy

Note: The optimisation of the state and rewards depend on the choice of K value. The K value were chosen randomly and hence we did not get an optimised result.
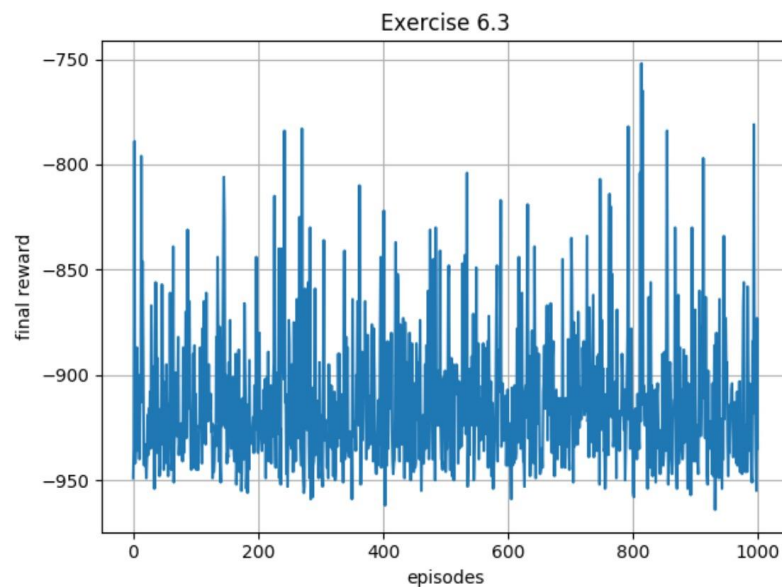
Assignment 6.3

Fig – 4 Accumulated rewards per episodes using policy gradient method

Note: We have used a neural Network as a function Approximator