**ML Assignment 3**

**Report Part 1**

1. Describe the environment that you defined. Provide a set of actions, states, rewards, main objective, etc.

- The environment defined is a grid world with 3x4 grid. The agent starts at a random location in the grid and its objective is to reach the target location which is fixed at [2, 3].

There are 4 other locations in the grid defined as rewards. The rewards have different values (-5, 2, -2, 2).

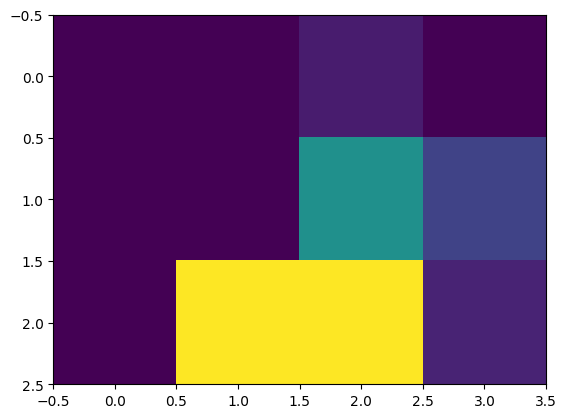
There are four possible actions: up, down, right, left.

The agent’s movements are stochastic meaning that there is a certain probability the agent will move in the intended direction and a certain probability that it will move in the other three directions.

The state space is also discrete and has twelve possible states, which are represented as the flattened version of the grid world. The agent’s current location, the target location and the reward locations are all indicated in the state.

The main objective of the agent is to reach the target location within a fixed number of time steps (14), while avoiding the rewards as much as possible.

2. Provide visualization of your environment.

- 

3. Safety in AI: Write a brief review (∼ 5 sentences) explaining how you ensure the safety of your environment. E.g. how do you ensure that the agent chooses only actions that are allowed, that agent is navigating within defined state-space, etc.

- The agent is moving across a 3x4 grid-based defined state space in the described environment. The only movements the agent can make are up, down, left, and right. These activities are constrained to remain within the grid's constrains.

The environment also has safety measures in place to guarantee that the agent cannot leave the state area or carry out any illegal activities. In order to avoid the agent getting stuck down or requiring an excessive amount of time to finish the task, the agent is also assigned a set number of timesteps.

Last but not the least, the incentives have been created to motivate the agent to hit the mark while avoiding dangerous states, which incentivizes the agent to take safe actions while navigating the environment.