Each of you have to come up with a Gridworld environment based on the rules assigned to you in the [Gridworld rules sheet](https://docs.google.com/spreadsheets/d/1d0hZ2jLyAPmU3igEa8JO1gg856R0CbiWkBf_A-3WQlo/edit?usp=drive_link), refer to the [notebook](https://colab.research.google.com/drive/12MPicuzwRHibcBxQu51hQTGMPpru62_8?usp=sharing) discussed in class to modify the transition probabilities and/or reward functions and come up with:

* Optimal state values through value iteration.
* What are the average total rewards that the agent gets, with and without training.
* Create a heatmap for the optimal state values.

NOTE: To understand Jump states, you can refer to Example 3.8 from Sutton Textbook (<https://drive.google.com/file/d/1ThdA8GgOgNbK4i2UU-XIQkxRMpp85LVP/view?usp=drive_link>). Please reach out if you have doubts or want to discuss how the transition probabilities can be modified for Jump states.

Example of jump state,

***{***

***"jump\_state":***

***{***

***"from": [3, 4],***

***"to": [3, 3],***

***"one-step-reward": 10***

***}  
}***

The condition to be satisfied here is that, the agent always transitions from [3, 4] state to [3, 3]

irrespective of the action it took, gets a one step reward of +10 for the transition.