Q3. Written:

How do you think this compares with your manual policy?
This policy is much worse than my manual policy. I do not however feel accomplished by this.

What are some reasons for the difference in performance?
- I have a predefined understanding of the rules to win the game while random choice does not.
- Random choice does not "learn" thus it does not achieve better performance while playing. I can get better at the game by discovering faster paths and choosing to take those in the future.

Q4 Written:

What strategy does each policy use, and why does it lead to generally worse/better performance?

(Better) We can see that the winning square of the gameboard lies to the right of the starting square. Thus, if we alter the random probability of 0.25 for R/L/U/D to skew towards the next action being right, we should theoretically obtain better performance since we are more likely to move towards the direction of the winning square. I accomplish this by altering the probabilities of each move from 0.25 for all to 0.4 for right and 0.2 for others.

(Worse) In the same thought process, we can see that the winning square of the gameboard gets farther away the more we move to the left. Thus, if we alter the random probability of 0.25 for R/L/U/D to skew left, we should theoretically obtain worse performance since we are more likely to move away from the winning square. I accomplish this by altering the probabilities of each move from 0.25 for all to 0.4 for left and 0.2 for others.