NAME: SHUBHAM SHARMA
ROLL NO. : 18I190002
MSc-PhD(OR)

Design Decisions/Assumptions:

a) Tolerance is taken to be 1e-12 for Value Iteration
b) For ending states(s) in episodic tasks, $r(s, a, s') = 0 \ \forall s' \in S, a \in A$
c) For ending states(s) in episodic tasks, $T(s, a, s') = 0 \ \forall s' \in S \backslash \{s\}, , a \in A \ and \ T(s, a, s) = 0$ where s is an ending state.
d) V(s) = 0 in Harword policy iteration for s∈Ending states. Linear equations are only solved non terminating states, as it is a singular matrix otherwise.

Observations:

a) Value iteration is faster than hpi and lp.
b) vi is taking less than 3 minutes to run MazeVerifyOutput.py, while lp and hpi are taking more time. So, it is advisable to run vi for MazeVerifyOutput.py.
c) lp algorithm is less precise than hpi and vi.

Maze problem formulation:

a) Let X be the array given in the grid
b) A loop in run over X for each and every element checking for $X[i, j] \neq 1$ and making that a state.
c) $S = \{(ij), \ X[i, j] \neq 1\}$
d) The set of action is A = {N, W, S, E}.
e) The transition probabilities are kept to be like this that if it is possible to go from state s to s', then $T(s, a, s') = 1$, otherwise, if it is not possible to go from s to s', then $T(s, a, s') = 0$.
f) If there is a wall or 1 in left(W), right(E), up (N) or down(S), then $T(s, a, s) = 1$ for that state.
g) It is possible to go from one state to other if and only f the other state is zero and nearby, i.e., touching that particular state(as per mentioned in the question statement)
h) The reward functions are kept like this:
    1) $r(s, a, s') = -1,$ if we are at 0 and the other state s' is also 0.
    2) $r(s, a, s) = -5 \ \forall \ s \in S \backslash \{t\},$ where let say t is the terminal state. -5 will prevent the self loops to happen.
    3) $r(s, a, t) = 100$, where t is the terminal state and s be a state from where we can go to t ,i.e. $T(s, a, t) = 1$. This id done to make this a perfect sink.
    4) $r(t, a, s) = \ 0$, where t is the terminal state and s be a state from where we can go to t ,i.e. $T(s, a, t) = 1$
i) For all other cases, $T(s, a, s') = 0 \ and \ r(s, a, s') = 0$