ARTIFICIAL INTELLIGENCE
CSL-333
ASSIGNMENT - 4 (Blackjack)

The state is determined by the hand value. The dealer's states vary from 2 to 21 valued hand (ace counted as 1 if it exists in hand) and A1 to A9 valued hand meaning hand values from 12 to 21 with ace being counted as 11.

For the dealer, we compute the one step transition probability matrix that would store probabilities of dealer going from one state to another in a single step. Let us call this matrix D, then Dn would store the state transition probabilities for n steps. In our case, we have computed D20 as after 20 chances dealer must have either busted or had taken a stand.

Now we calculate expected reward if the player is at a particular state using the above matrix (using winning and losing probabilities of dealer from a starting state). Thus, we compute an expected reward matrix that would store the expected winnings if player stands corresponding to every player and dealer state. Let's call this M.
We compute expected winnings by hitting once and then standing using a matrix which is product of D and M. We made another matrix V which is max(M,DM) meaning max value for each entry. For 15 iterations, we did V = max(V,MV).
If value of V exceeds value in M, we hit else we stand.

Doubling –

2DM would store reward obtained from doubling the amount and making a stand after exactly one step. Thus if value in 2DM is greater than V and M, we would go for doubling.

Splitting –

On splitting, a player gets 2 new hands each having a card already in the hand and a new card. Then the expected winning from each of the new hands formed would be sum of (probability of getting hand k * Expected reward if player is holding hand k provided k doesn't form a pair) + (probability of getting a pair * twice of expected winning from the pair). Thus, the above equation is solved to compute expectation of winnings per split. The special case of AA split has been considered independently.