Implemented Double DQN network to avoid overestimating the q values

Prioritized replay buffer

Deuling DQN

Architecture of Deuling DQN

* + input size = 37
  + output size = 4
  + 2 hidden layers
    - each hidden layer has 128 hidden units and is followed by a ReLU activation layer
  + one value stream network with a single hidden layer (128 hidden units) and is followed by a ReLU activation layer. The output layer of the value stream is followed by a linear activation unit with 1 outputs
  + one advantage stream network with a single hidden layer (128 hidden units) and is followed by a ReLU activation layer. The output layer of the advantage stream is followed by a linear activation unit with 4 outputs
  + addition of state value from value stream to the advantage values from the advantage stream to compute the q values