



Policy Function could be modelled as a Neural Network $\pi_\theta(a_t|o_t)$

State Evolving Function $p(s_{t+1}|s_t, a_t)$ and Cost Function $c_t = c(s_t, a_t)$