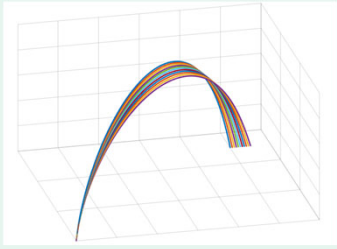
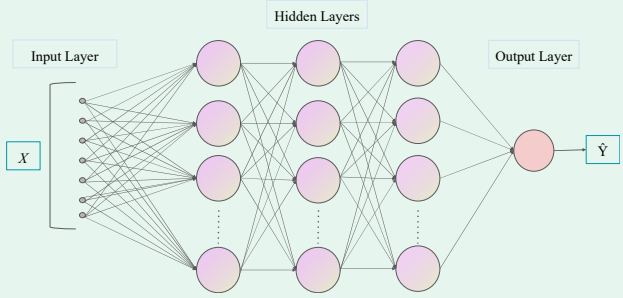
(a) Behavior cloning approach



Labeled trajectory dataset



Train the neural network

(b) Strategy exploration approach

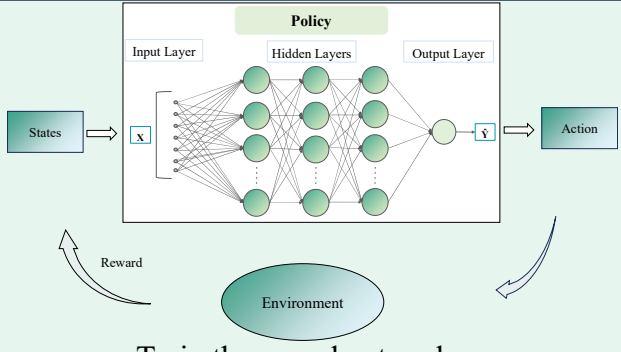
$$r_1 = \begin{cases} a, if \dots \\ b, if \dots \end{cases} \quad \text{(sparse Reward)}$$

$$r_2 = e^{-\frac{c}{\sigma^2}} \quad \text{(intensive Reward)}$$

$$r_n = \cdots$$

$$Reward = r_1 + r_2 + \cdots + r_n$$

Design of reward functions



Train the neural network