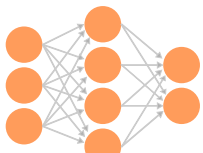


Agent

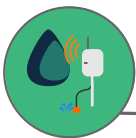
State (s_t)

Reward (r_t)



$$\pi(a_t | s_t)$$

Actions (a_t)



Sensors



Valves

Environment