Deep Q Learning

Policy Network

Target Network

Replay Buffer

Environment