# DQN

Iniziamo con:

REWARDS = {

    "win": 1.0,           # vittoria

    "lose": -1.0,         # sconfitta

    "draw": 0.5,          # pareggio

    "valid\_move": 0.0,    # mossa valida ma non vincente

    "invalid": -0.5,      # mossa non valida (colonna piena)

}

In fase training:

*learning\_rate*=1e-3,

*buffer\_size*=50000,

*batch\_size*=64,

*exploration\_initial\_eps*=1.0,

*exploration\_final\_eps*=0.05,

*exploration\_fraction*=0.3,

*gamma*=0.99,

*target\_update\_interval*=500,

Il trainign si compone di 3 fasi (circa 28 minuti):

**fase 1 -> DQN vs Random [150\_000 timesteps]**

**fase 2 -> DQN vs RuleBasedL1 [200\_000 timesteps]**

**fase 3 -> DQN vs RuleBasedL2 [200\_000 timesteps]**

Dopo il training attraverso file agent\_evaluator:

200 partite DQN vs Random

Immagine che contiene testo, linea, Diagramma, diagramma

Il contenuto generato dall'IA potrebbe non essere corretto.

Immagine che contiene testo, diagramma, Diagramma, linea

Il contenuto generato dall'IA potrebbe non essere corretto.

200 partite DQN vs L1

Immagine che contiene testo, diagramma, linea, Diagramma

Il contenuto generato dall'IA potrebbe non essere corretto.

Immagine che contiene testo, Diagramma, linea, diagramma

Il contenuto generato dall'IA potrebbe non essere corretto.

200 partite DQN vs L2

Immagine che contiene testo, linea, diagramma, Diagramma

Il contenuto generato dall'IA potrebbe non essere corretto.

Immagine che contiene testo, diagramma, Diagramma, linea

Il contenuto generato dall'IA potrebbe non essere corretto.