Episodes' rewards using DQN agent with heuristic reward