

```

    k-armed bandits Upper confidence bound
bandit:''' UpperConfidenceBoundBandit
init(self,k,c,ifers,mu='random'):Numberofarmsself.k=kExplorationparameterself.c=cNumberofiterationsself.ifers=ifersStepcountself.n=1Stepcount
module==np.name;User-definedaveragesself.mu=np.array(mu)elifmu=='random':Drawmeansfromprobabilitydistributionself.mu=np.random.normal(mu,1/self.c)
reward+self.c*np.sqrt((np.log(self.n))/self.k_n))
n[a]+=1
reward=self.mean_reward+(reward-self.mean_reward)/self.n
self.k_reward[a]=self.k_reward[a]+(reward-self.k_reward[a])/self.k_n[a]
reward
n=np.ones(self.k)self.mean_reward=0self.reward=np.zeros(ifers)self.k_reward=np.zeros(self.k)ifmu=='random':Drawmeansfromprobabilitydistributionself.mu=np.random.normal(mu,1/self.c)
rewards=np.zeros(ifers)Initializebanditsuch=ucb_bandit(k,2,ifers)episodes=1000Runexperimentsforiinrange(episodes):
rewards=ucb_rewards+(ucb.reward-ucb_rewards)/(i+1)
rewards,label="UCB")plt.legend(bbox_inches=(1.2,0.5))plt.xlabel("Iterations")plt.ylabel("AverageReward")

```