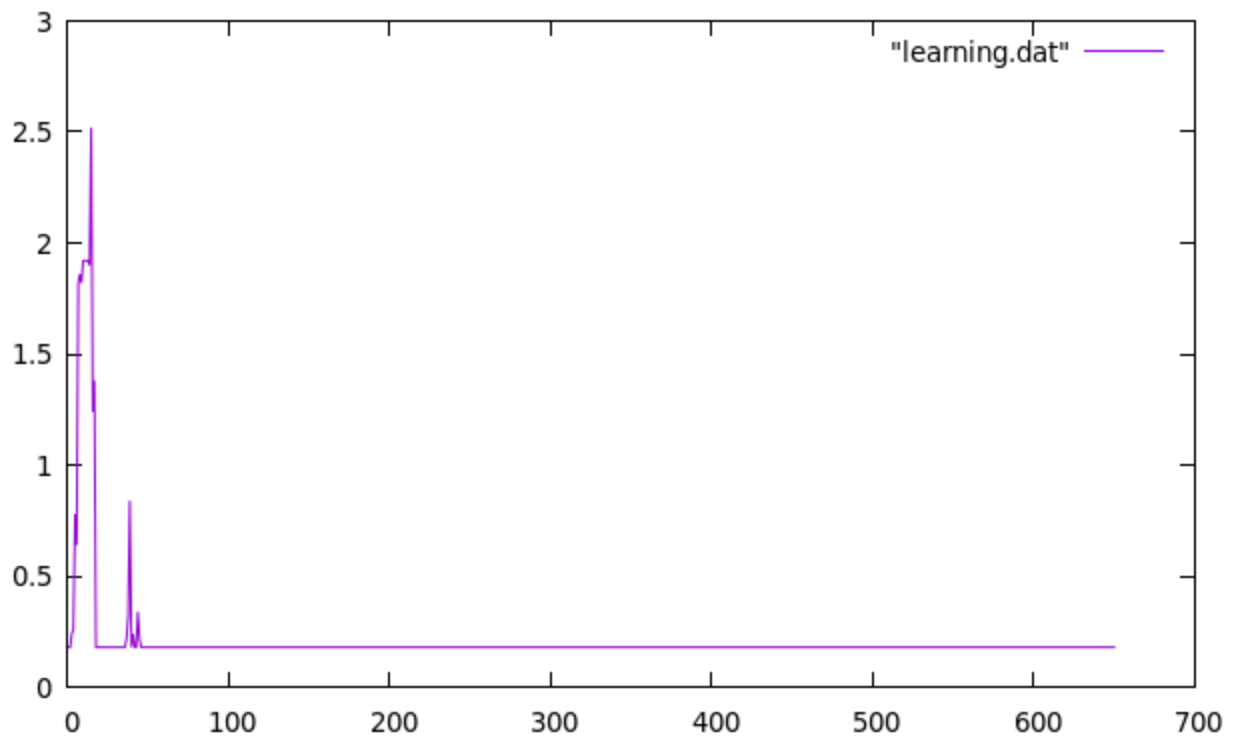**Case 1:**
EPSILON .0001
ALPHA 0.1
GAMMA 0.95(learning1.dat)
The graph converges for these hyper parameters

Case 2:
EPSILON .01
ALPHA 0.5
GAMMA 0.5(learning2.dat)
By increasing epsilon the doesnot learn correctly and doesnot converge