eval accuracy: 0.6869565