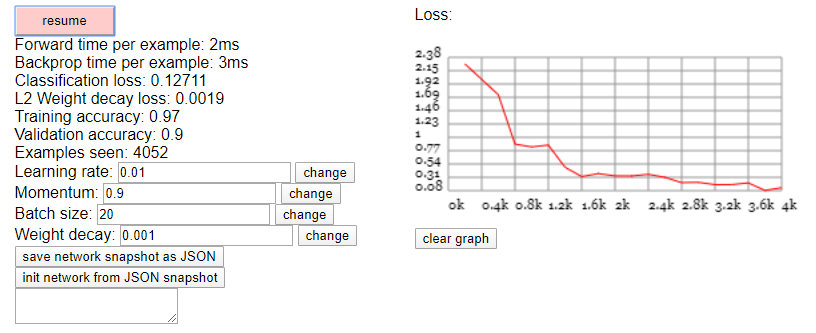
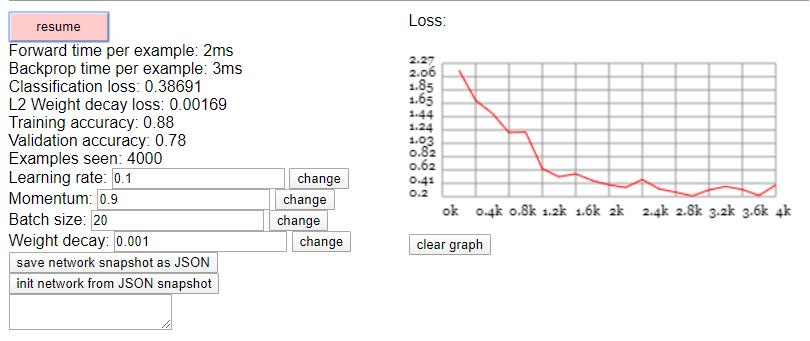
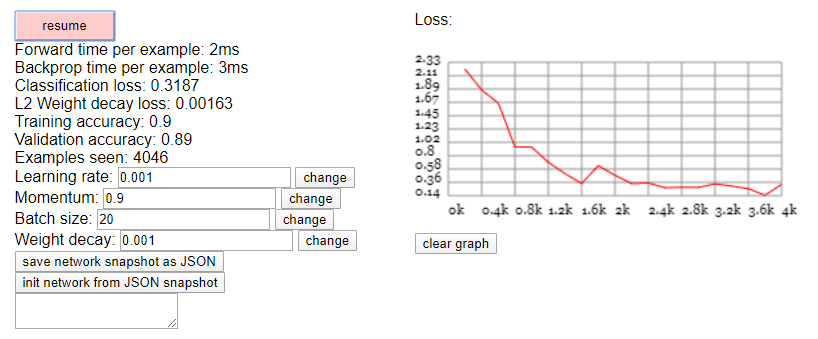
Lucas Fonseca 221150154

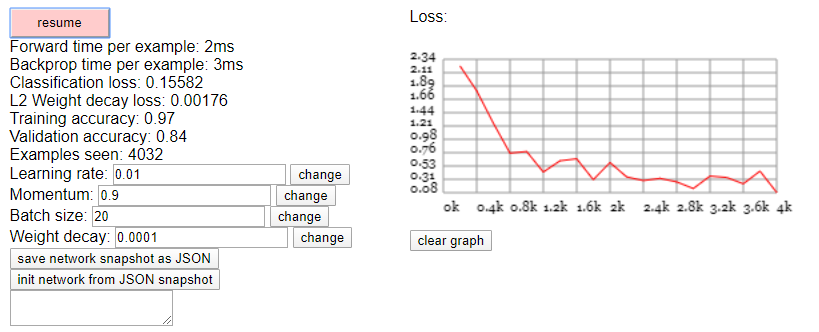
Digitos:

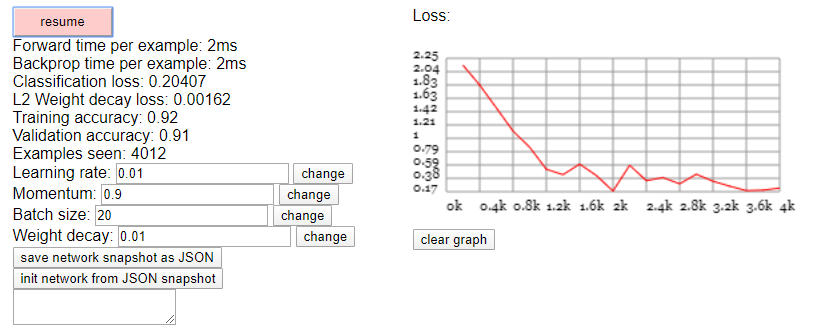




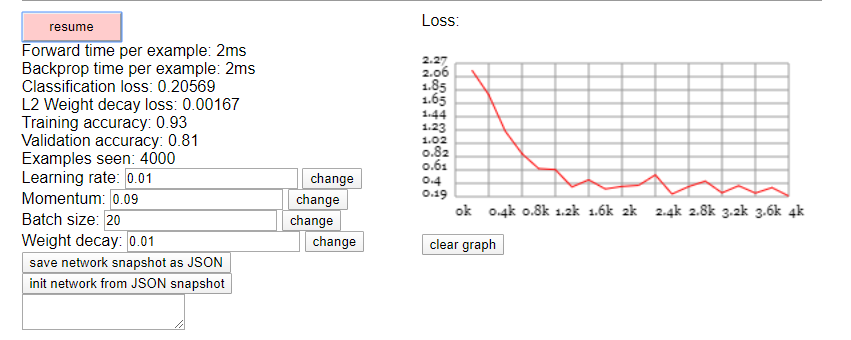


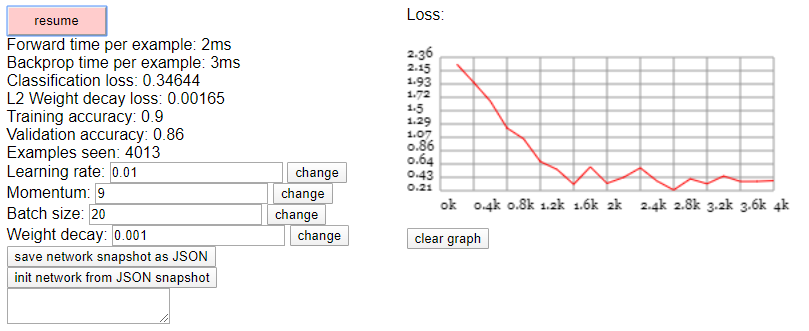
Depois de mudar o learning rate de 0.01 para 0.1 e 0.001 observa-se que a melhor rede é a que possui um learning rate de 0.01, e então pegamos esta rede e mudamos o weight decay de 0.001 para 0.0001 e 0.01



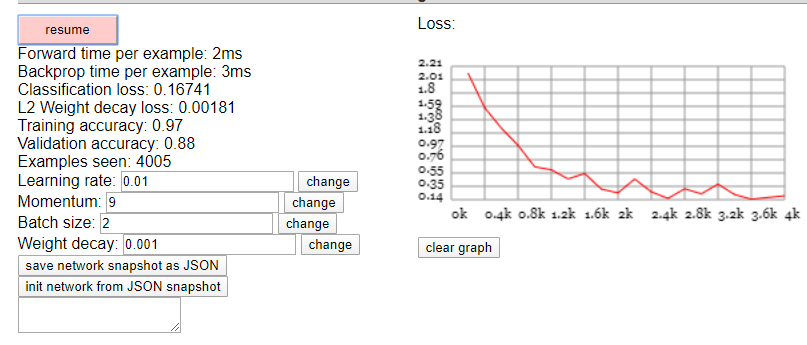


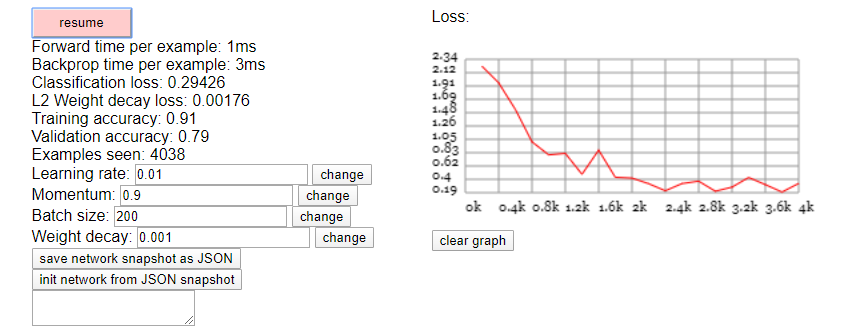
Mudando o momentum para a rede que possui learning rate de 0.01 e weight decay de 0.01





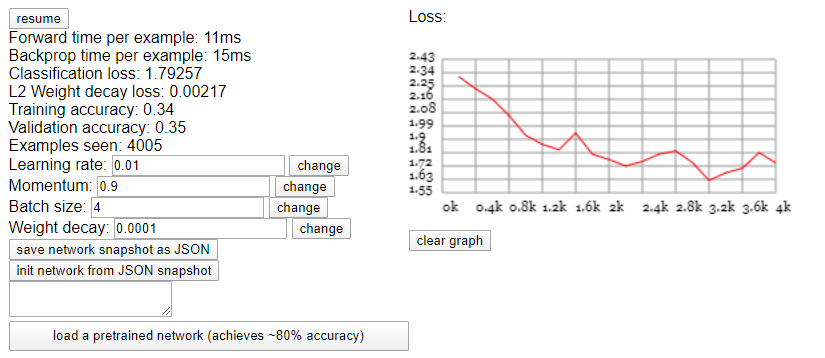
Mudando o batch size para a rede que possui learning rate de 0.01 momentum de 9 e weight decay de 0.001

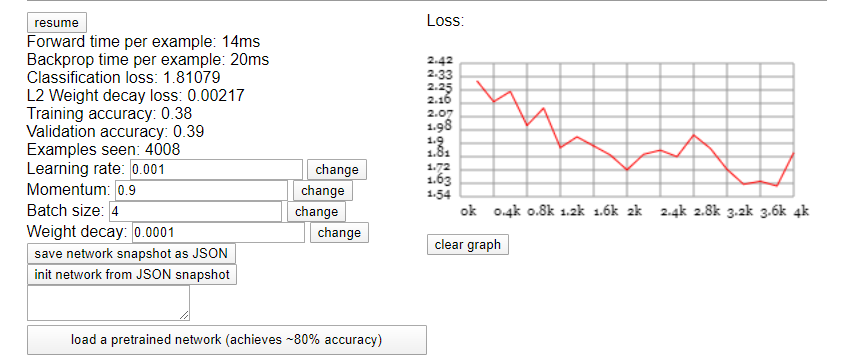


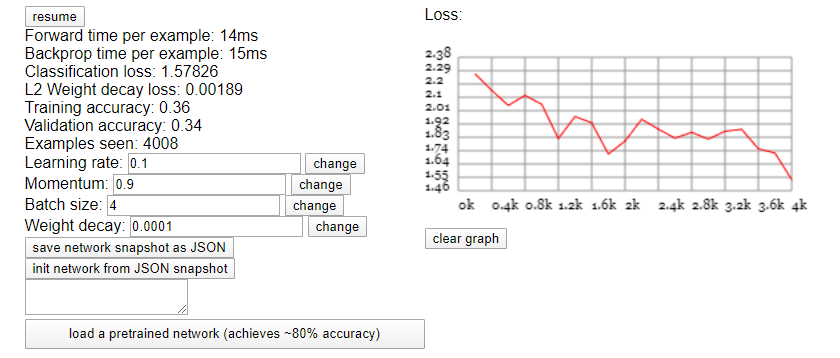


A melhor rede é a que possui learning rate 0.01, momentum 0.9, batch size20 e weight decay 0.01

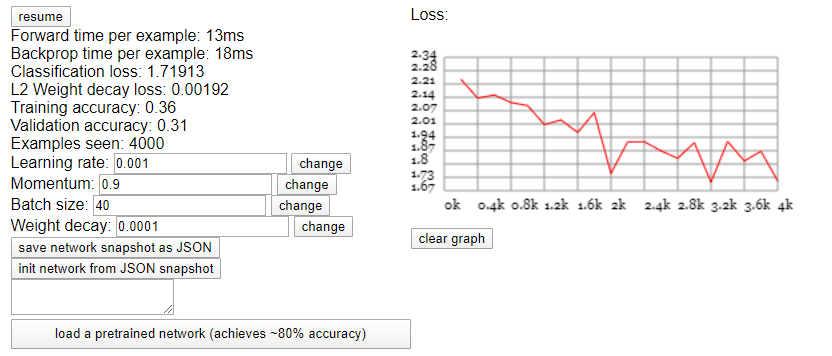
Imagem:

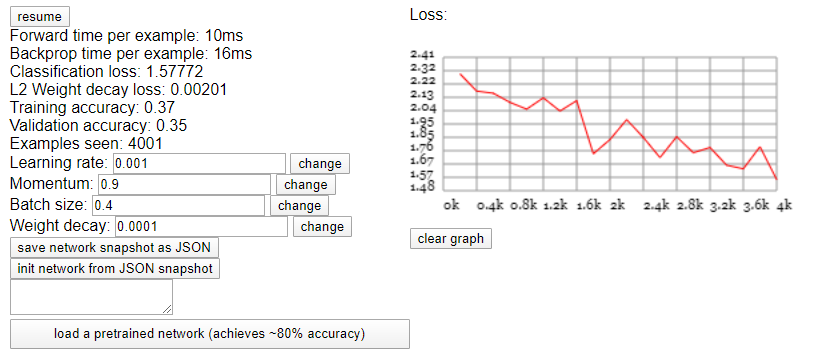




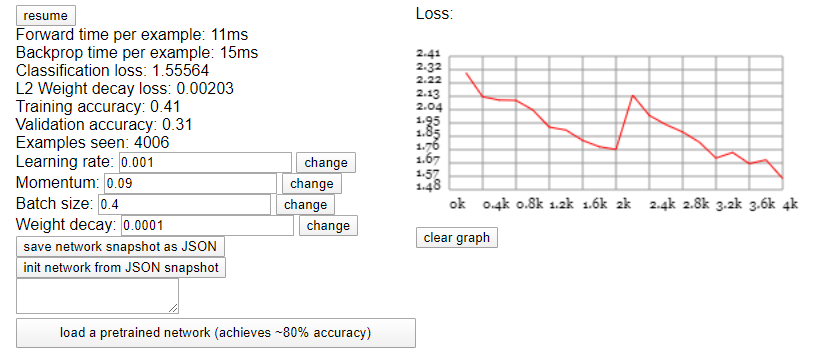


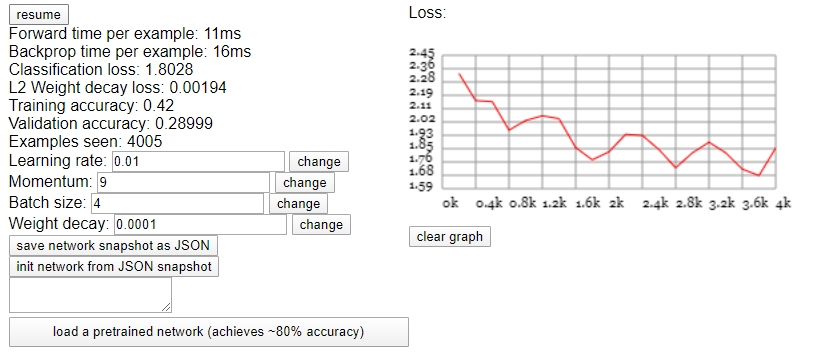
Mudando o learning rate de 0.01 para 0.001 e 0.1 observa-se que o validation accuracy foi melhor para 0.001, então mudamos o batch size de 4 para 40 e 0.4



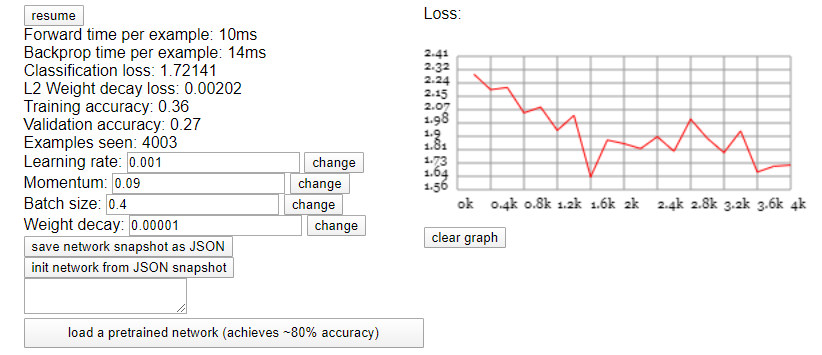


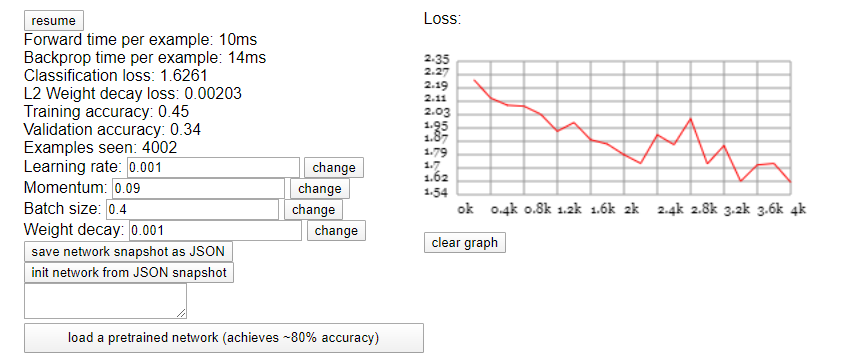
Mudando o momentum para a rede com LR 0.001 e BS 0.4





E então mudando o weight decay para a rede com o LR 0.001, Momentum 0.09 e BS 0.4





A melhor rede é a que possui LR 0.001, Momentum 0.9, BS 4 e WD 0.0001