Dropout(p=0.5)
[forward\_compute\_time=0.057,backward\_compute\_time=0.101,activation\_size=1048576.0,parameter\_size=0.0]

Linear(in\_features=4096, out\_features=1000, bias=True) [forward\_compute\_time=0.209,backward\_compute\_time=0.949,activation\_size=256000.0,parameter\_size=16388000.0]