

GANDALFModel (423 T params, 89 epochs)

