

GANDALFModel (664 T params, 98 epochs)

