

GANDALFModel (302 T params, 110 epochs)

