

GANDALFModel (60 T params, 76 epochs)

