

GANDALFModel (785 T params, 112 epochs)

