

GANDALFModel (483 T params, 71 epochs)

