

FTTransformerModel (162 T params, 78 epochs)

