

FTTransformerModel (162 T params, 90 epochs)

