

FTTransformerModel (130 T params, 121 epochs)

