

FTTTransformerModel (48 T params, 180 epochs)

