

FTTransformerModel (48 T params, 118 epochs)

