

FTTTransformerModel (549 T params, 132 epochs)

