

FTTTransformerModel (474 T params, 85 epochs)

