

FTTransformerModel (89 T params, 110 epochs)

