

FTTransformerModel (89 T params, 145 epochs)

