

Legend:

B = batch size
L = sequence length

ChimeraLM Architecture

