DATASET: 1.5 BILLION TOKENS

EPOCHS

Isoloss curves

Y

1

87M

X

MODEL SIZE
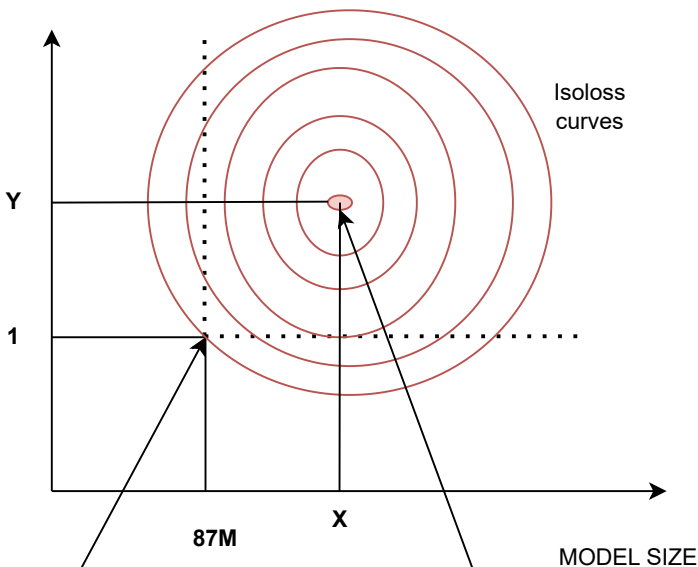
Chinchilla Optimal

Best Model possible given 1.5 Billion Tokens
by training a X parameter model for Y Epochs
(assuming infinite compute)