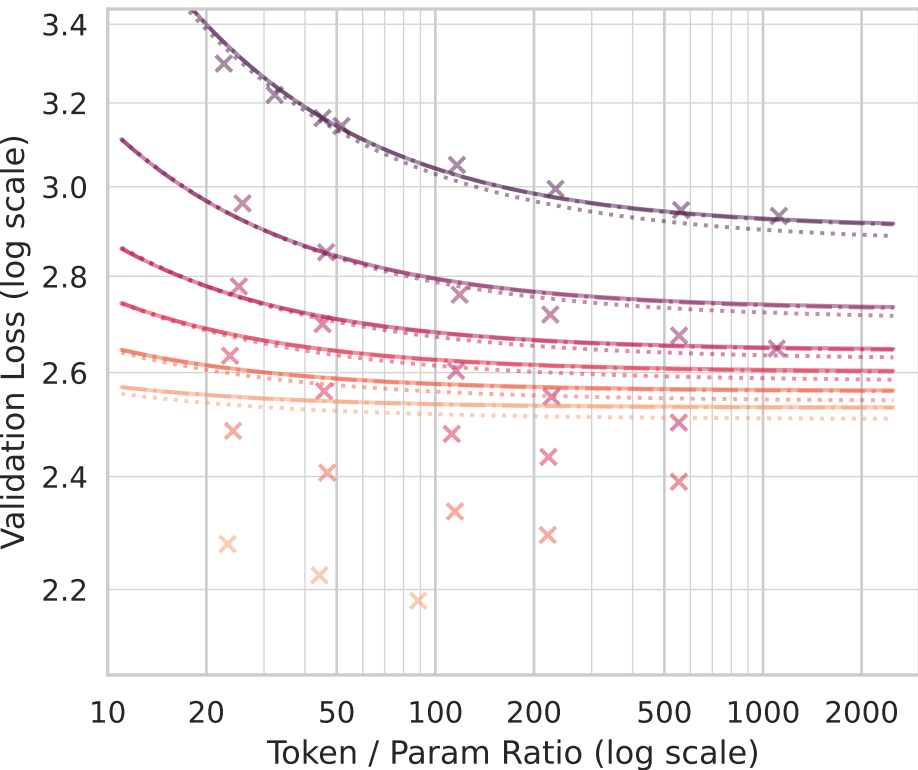
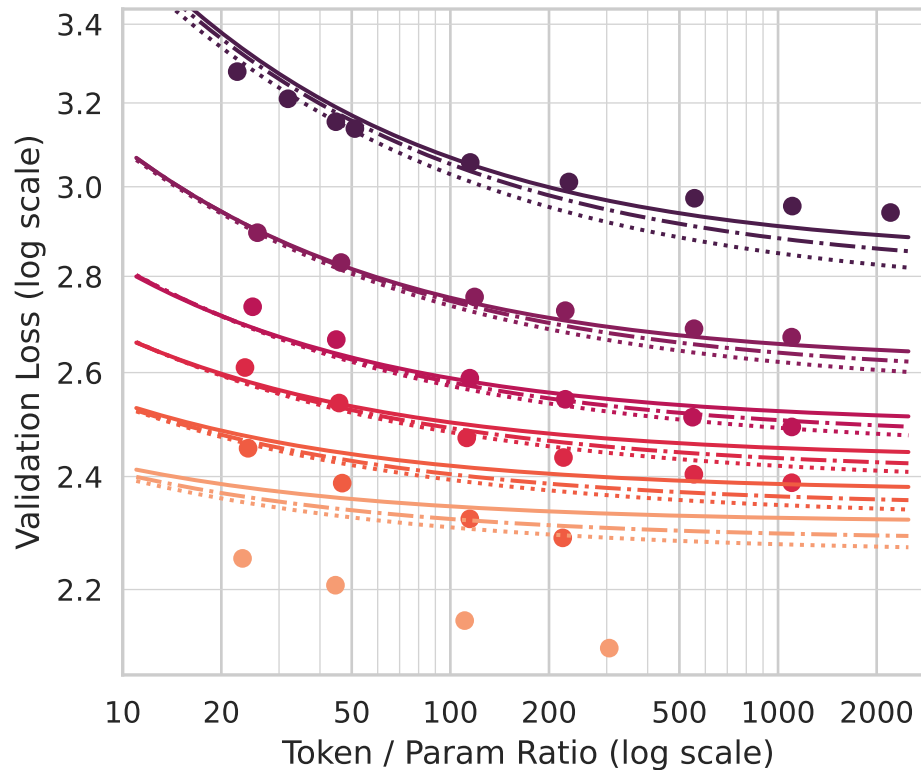


Llama



xLSTM

**Model Sizes**

160M

400M

830M

1.4B

2.7B

7B

**Empirical Data Points**

x Llama Data Points

● xLSTM Data Points

**L(N, D) Token Param Range**.....  $\leq 100$  Token/Param- - -  $\leq 300$  Token/Param

— All Data