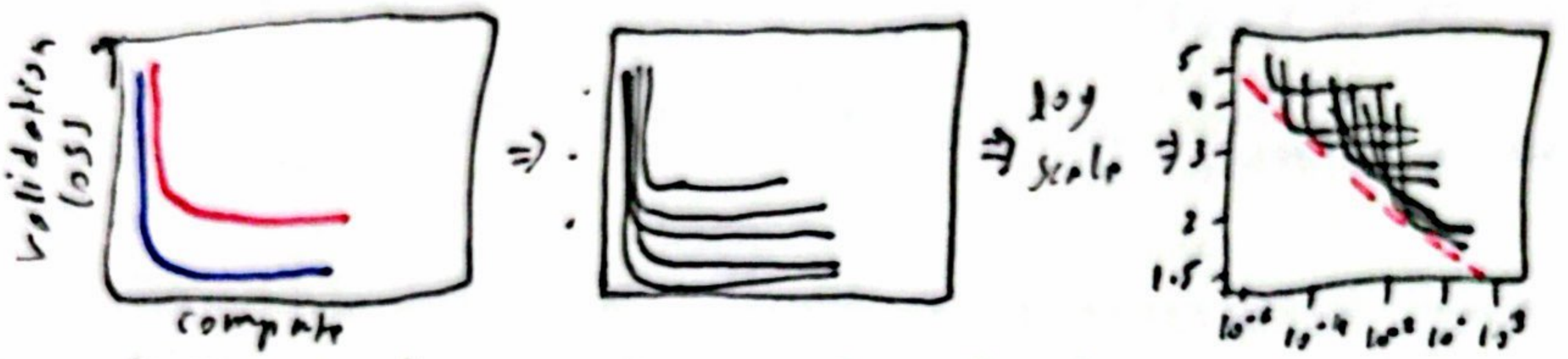
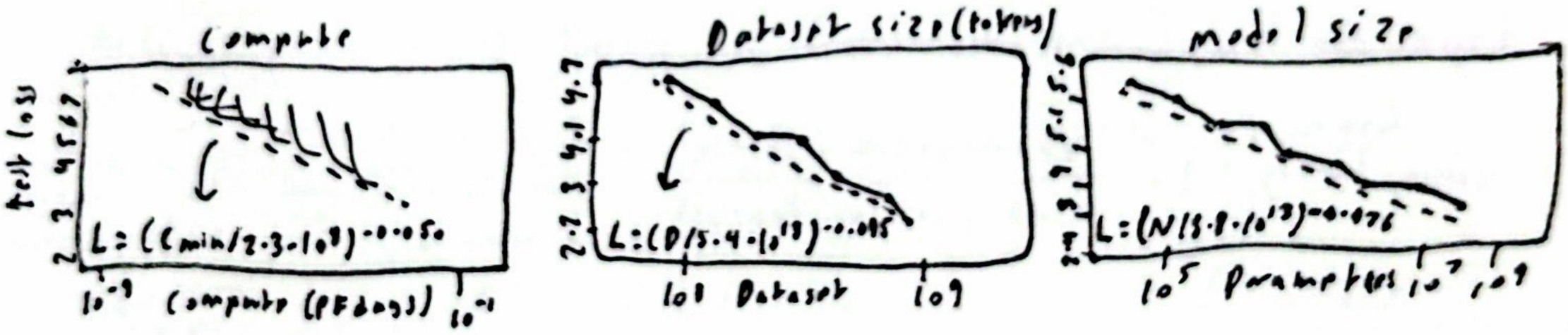
Nural scaling Laws

- as we train a AI model its error rate generally drops off Quickly and then levels off if we train a larger model it will achive a lower error rate but requires more compute Scalling to larger and larger models we end up with a family of curves switching to log scale a trend emerges. where do model can cross this line AKA the compute Efficency frontier



- This trend is one of three nural scalling laws that have been observed error rate scales in a simmilar way with compute. Model size and Data set size. and does not depend much on model architecture or other Agorithmic details.



- on log plots these show straight lines and the slope of each line is equal to the exponent of fit equation (L) Larger exponents make for steeper lines and more rapid preformance improvment.