

UDs



computing capacity



data embeddings



parameter features

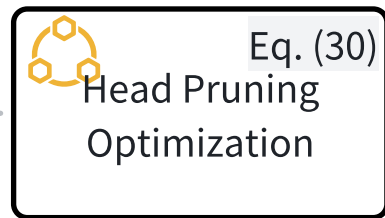
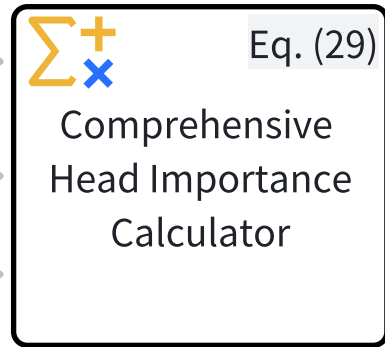
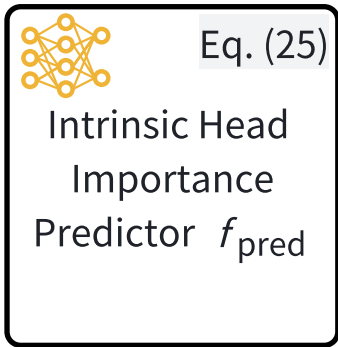


historical gradients

final $z^{l,h}$

pruning decision

ES

 y_m S_{base} $z^{l,h}$

iterative

 S_{comp} $z^{l,h}$