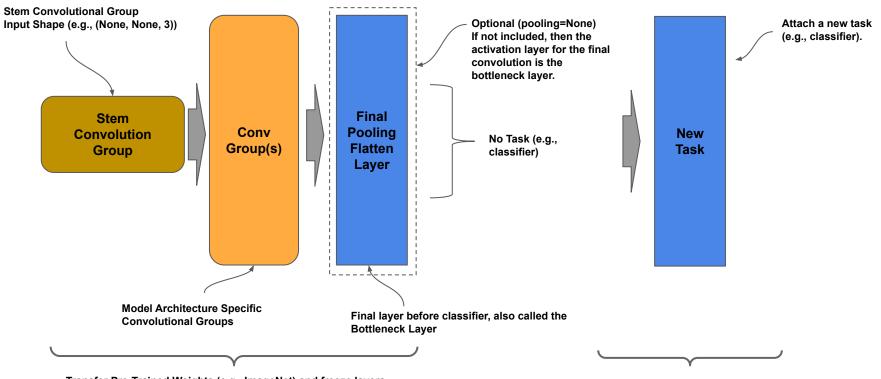




Google Cloud Al Developer Relations Automatic Learning

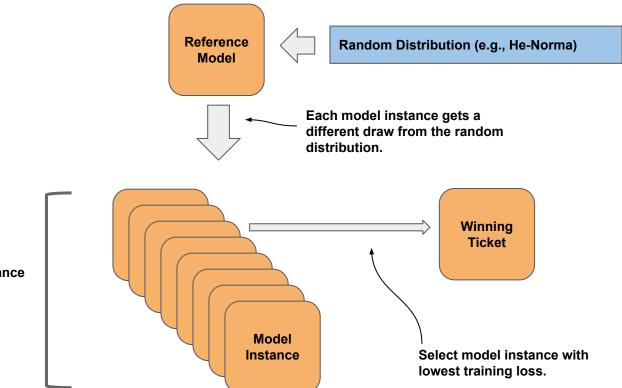
Transfer Learning



Transfer Pre-Trained Weights (e.g., ImageNet) and freeze layers.

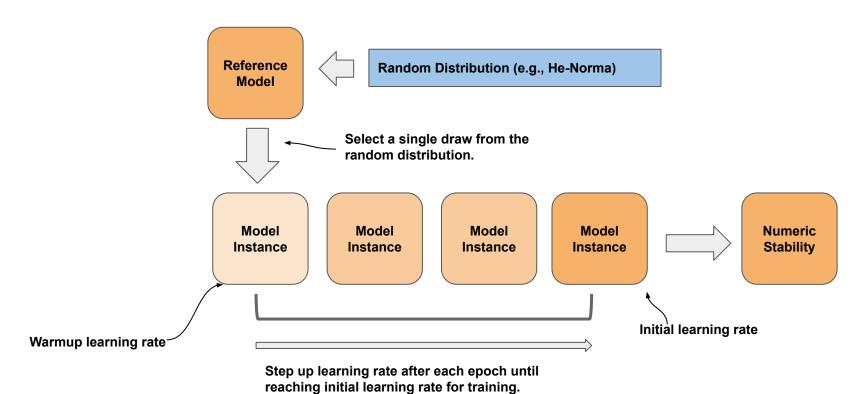
Train the layer(s) in the new task.

Weight Initialization (Lottery)

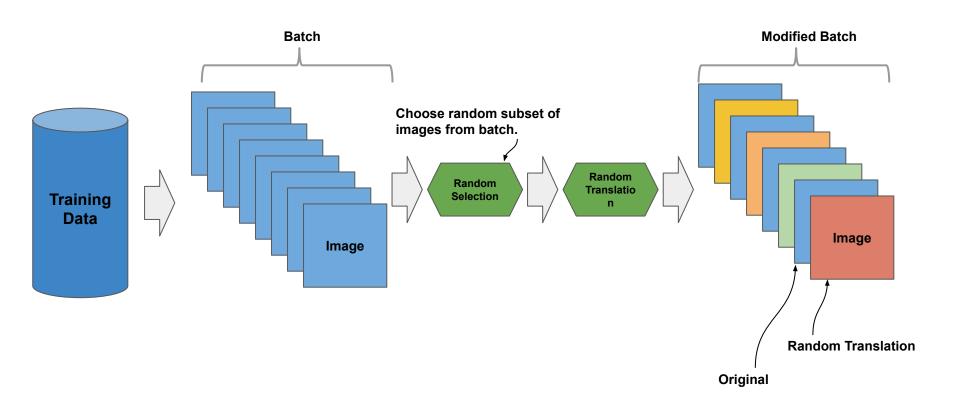


Pre-train each model instance with tiny learning rate.

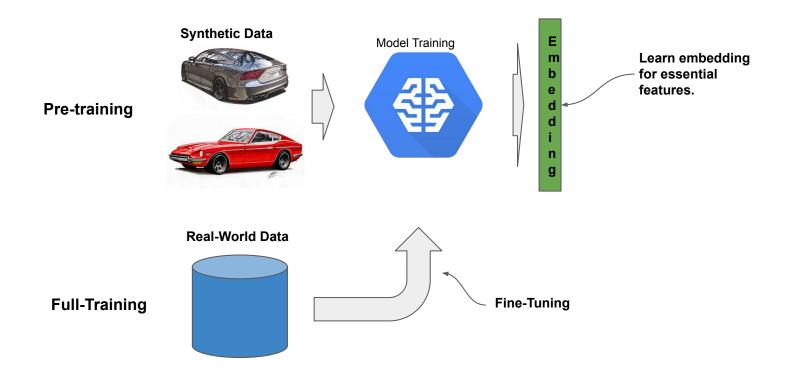
Weight Initialization (Warmup)



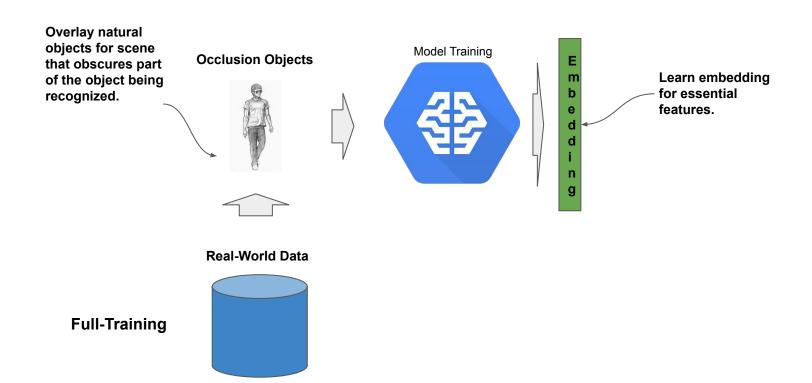
Translational/Scale Invariance



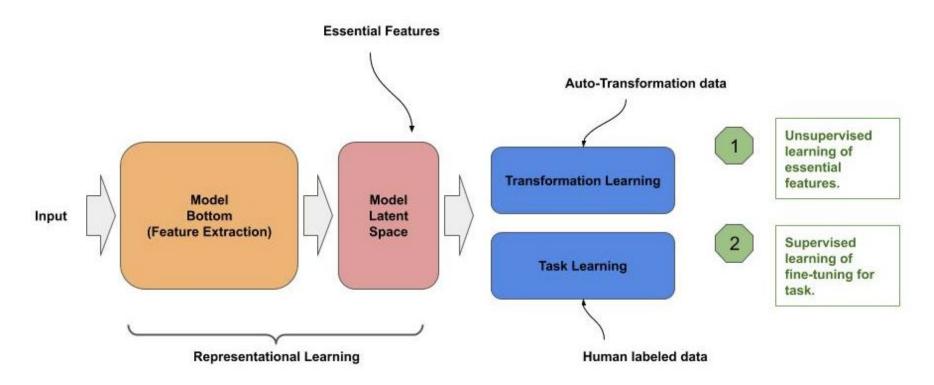
Essential Feature Learning (Synthetic)



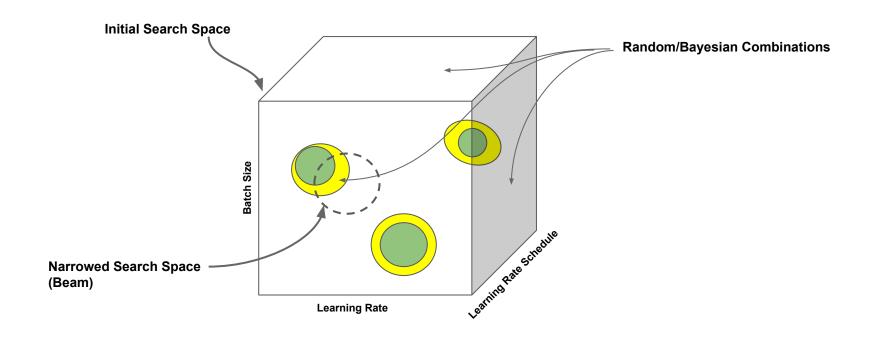
Essential Feature Learning (Occlusion)



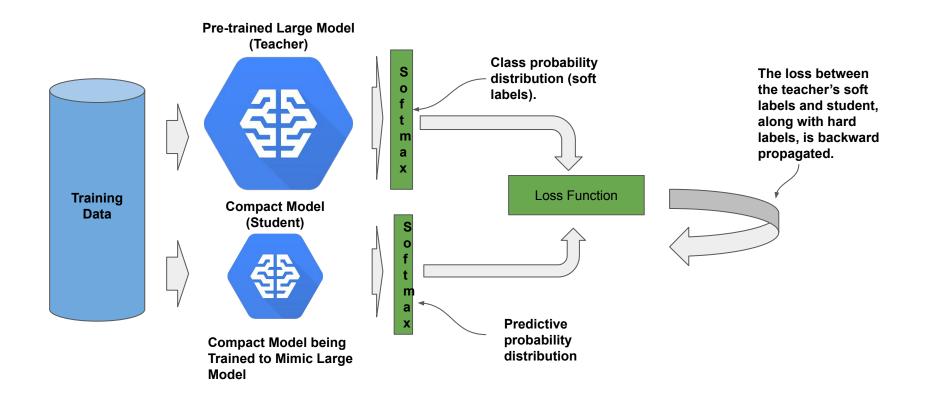
Unsupervised Pre-Text Tasks



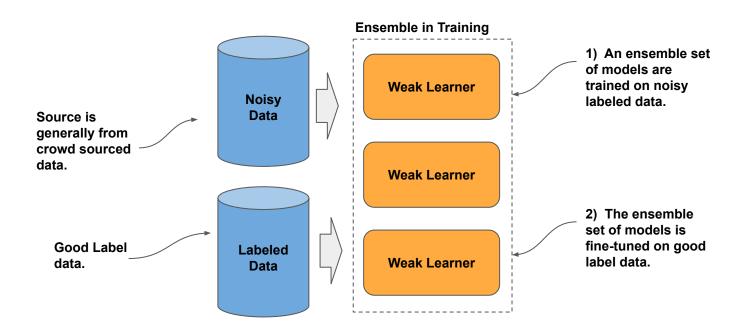
Automatic (Guided) Hyperparameter Tuning



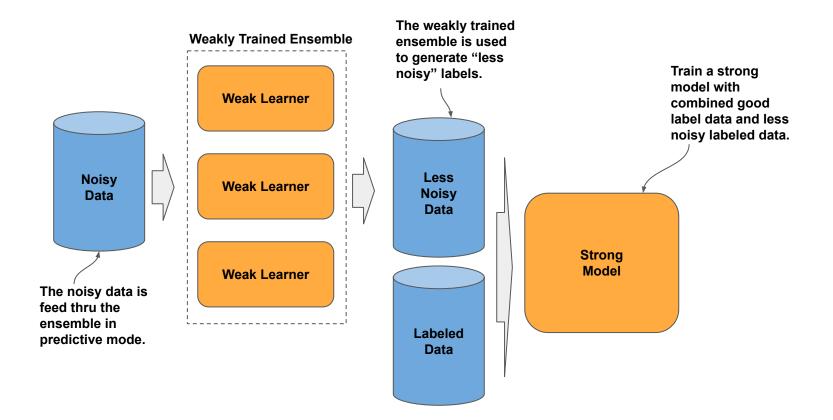
Knowledge Distillation



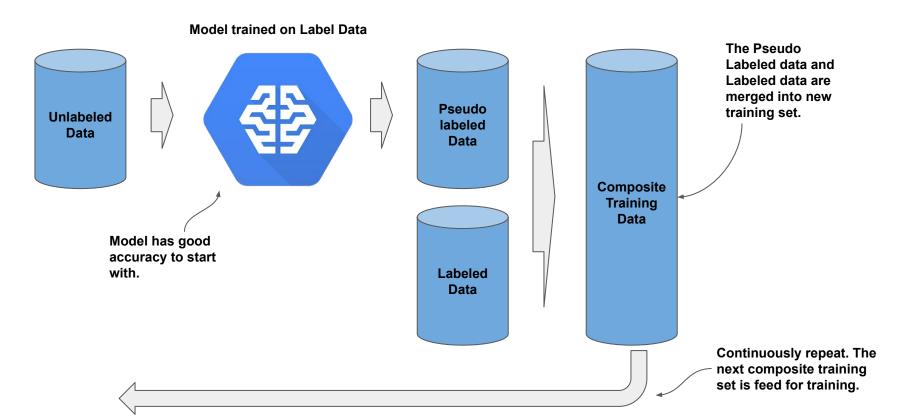
Data Denoising (Weakly Supervised)



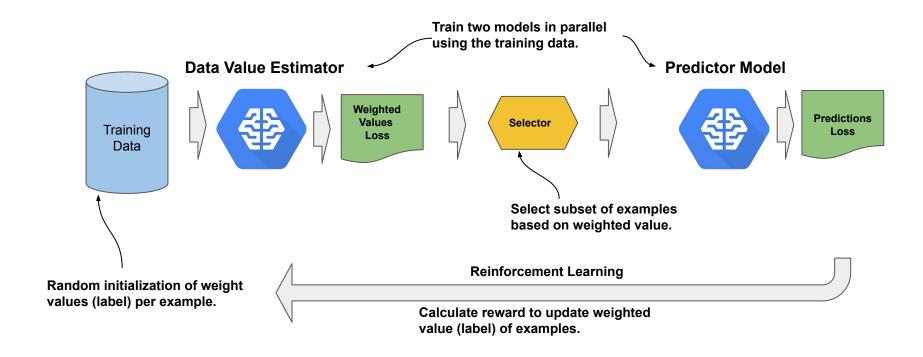
Data Denoising (Weakly Supervised)



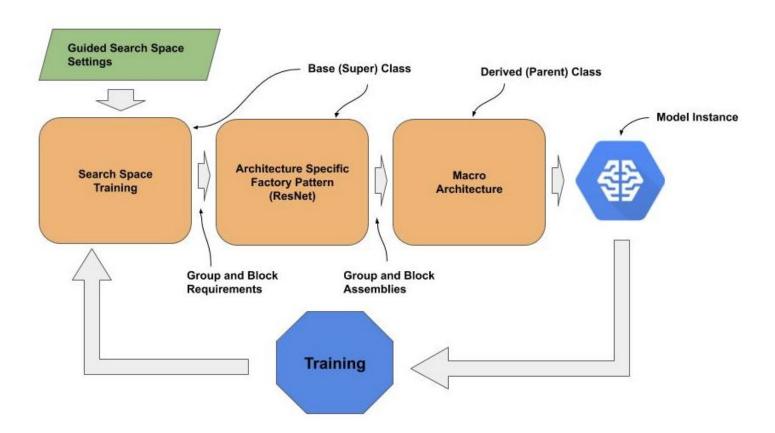
Data Labeling (Self-Supervised)



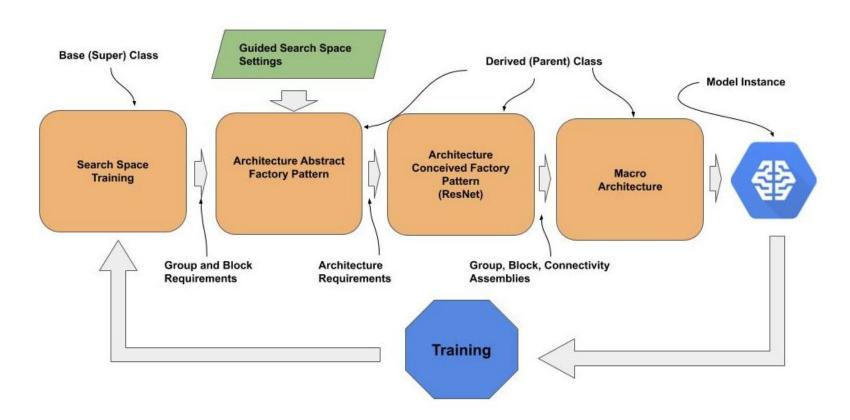
Data Validation (RL)



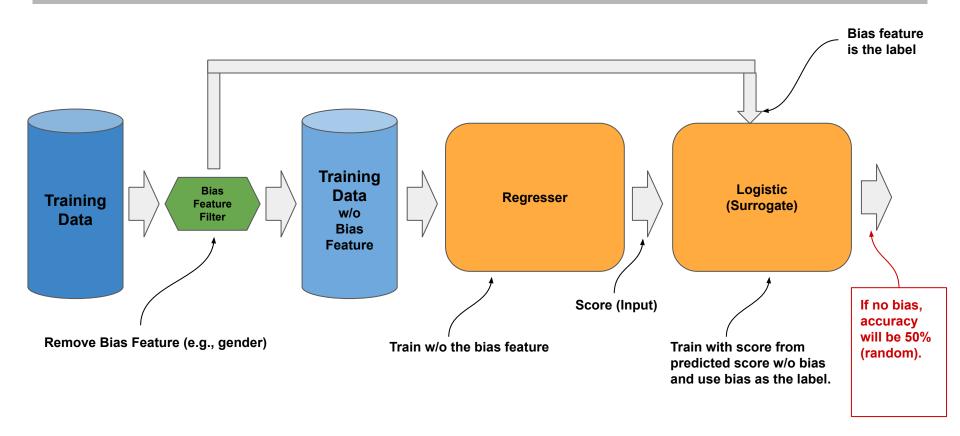
Macro Architecture Search



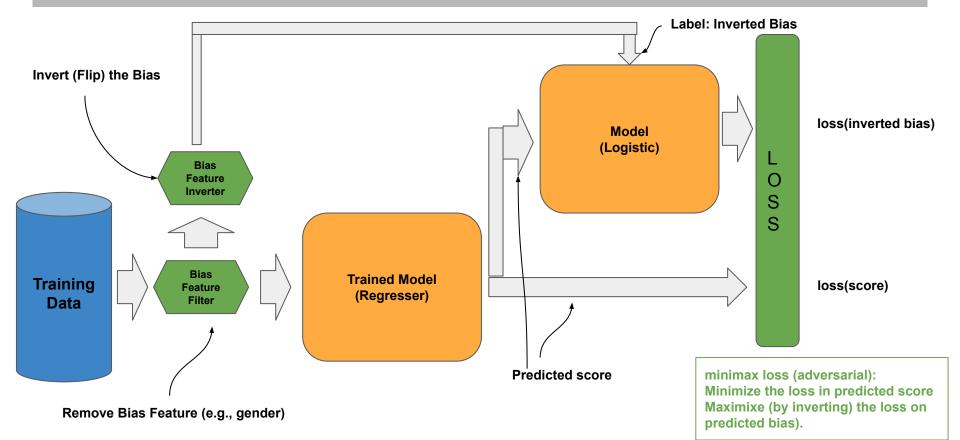
Micro Architecture Search (NAS)



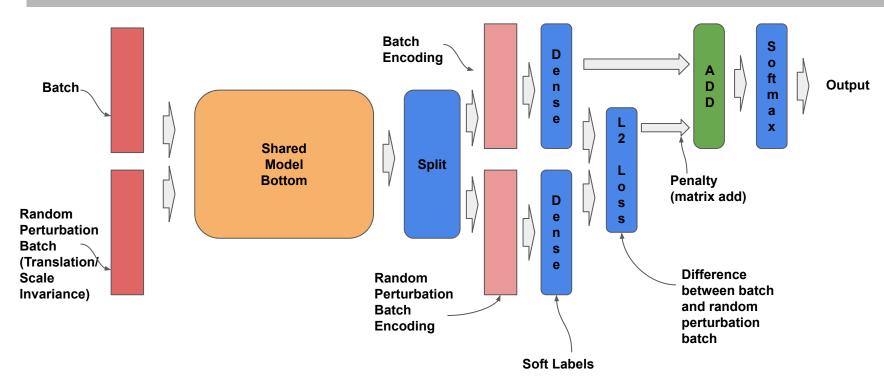
Unseen Covariant (Surrogate)



Unseen Covariant (Adverserial)



Neural Structured Learning



You take a batch. Then take a second batch which is a near mirror of the batch (same labels). If you have true representational learning in the model bottom then the soft label outputs from both should be identical (0 penalty). Otherwise, the difference is the penalty added to the loss (matrix add + softmax) that is backward propagated.

Machine Design Amalgamations (Frontier)

