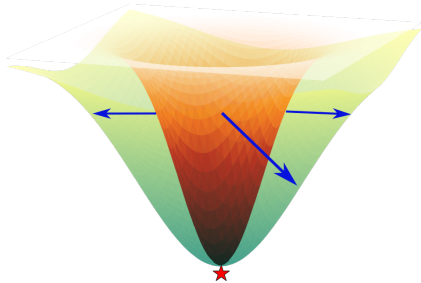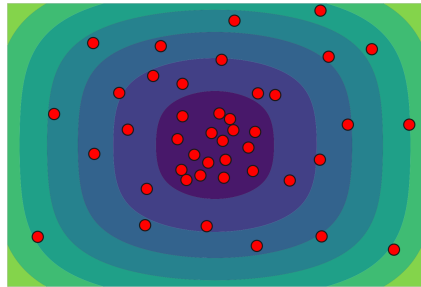(a) Learning a prior   (b) Prior rescaling   (c) Bayesian inference

Bayesian transfer learning pipeline. We (a) learn a prior by fitting a probability distribution over feature extractor parameters to a pre-training posterior mode, (b) rescale the prior for a downstream task, and (c) use the prior for Bayesian inference on downstream tasks.