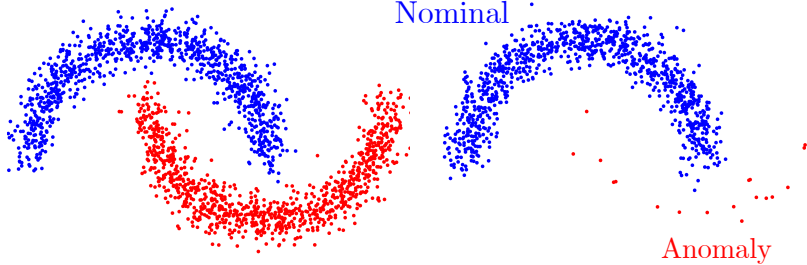
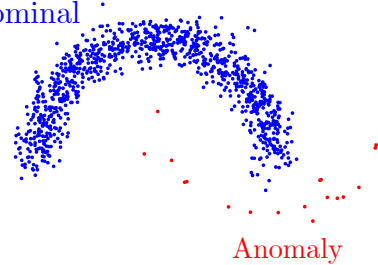


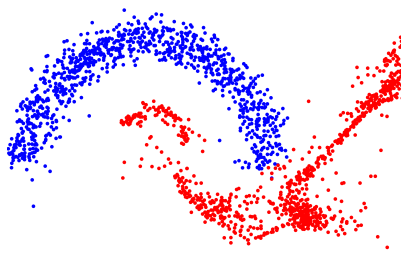
Nominal



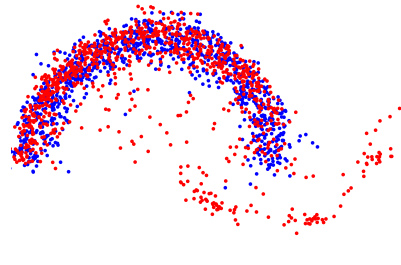
(a) GT



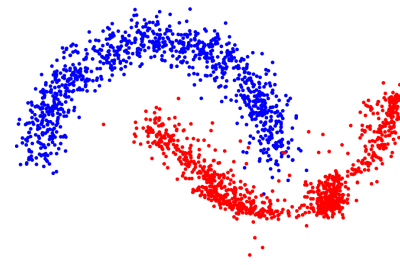
(b) Imbalanced dataset



(c) KL regularized, $\beta=0.01$



(d) KL regularized, $\beta=1.0$



(e) Calibrated (ours)