

KL divergence for checke1; darker = larger training sets

