

Figure 1: The results on Colored-MNIST dataset, considering isotropic Gaussian prior. (First Row): $d_z = 8$; (Second Row): $d_z = 64$; (Third Row): estimated information leakage $I(\mathbf{S}; \mathbf{Z})$ using MINE; (Fourth Row): estimated useful information $I(\mathbf{U}; \mathbf{Z})$ using MINE. (First Column): utility task is digit recognition ($|\mathcal{U}| = 10$), while the adversary's goal is the digit color ($|\mathcal{S}| = 3$), setting $P_S(\text{Red}) = P_S(\text{Green}) = P_S(\text{Blue}) = \frac{1}{3}$; (Second Column): utility task is digit recognition ($|\mathcal{U}| = 10$), while the adversary's goal is the digit color, setting $P_S(\text{Red}) = \frac{1}{2}$, $P_S(\text{Green}) = \frac{1}{6}$, $P_S(\text{Blue}) = \frac{1}{3}$; (Third Column): utility task is digit color recognition ($|\mathcal{U}| = 3$), while the adversary's interest is the digit number ($|\mathcal{S}| = 10$).

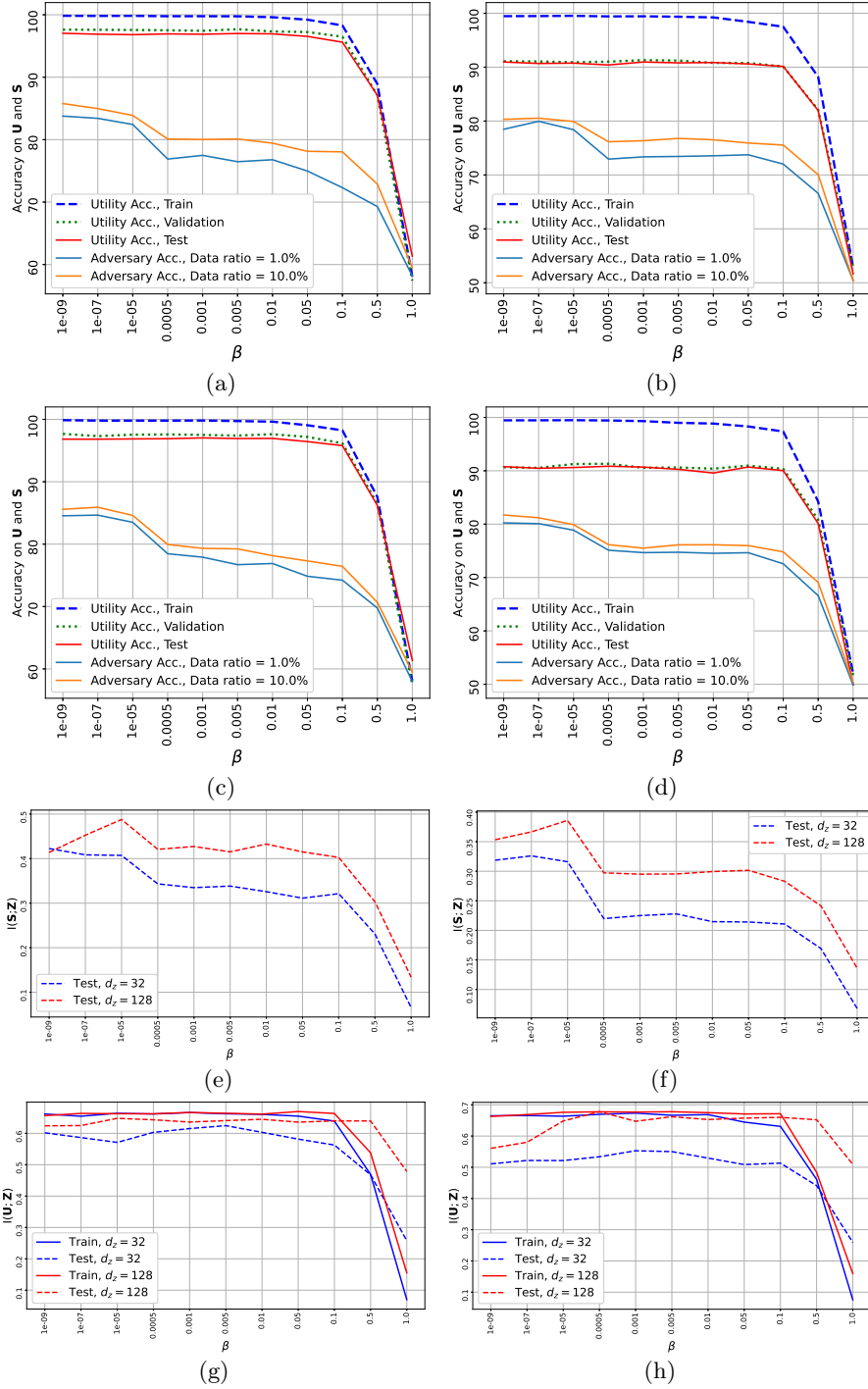


Figure 2: The results on CelebA dataset, considering isotropic Gaussian prior. (First Row): $d_z = 64$; (Second Row): $d_z = 128$; (Third Row): estimated information leakage $I(\mathbf{S}; \mathbf{Z})$ using MINE; (Fourth Row): estimated useful information $I(\mathbf{U}; \mathbf{Z})$ using MINE. (First Column): utility task is gender recognition ($|\mathcal{U}|=2$), while the adversary’s interest is Heavy Makeup ($|\mathcal{S}|=2$); (Second Column): utility task is emotion (smiling) recognition ($|\mathcal{U}|=2$), while the adversary’s interest is mouth slightly open ($|\mathcal{S}|=2$).