This paper introduces a novel iterative method for missing data imputation that sequentially reduces the mutual information between data and their corresponding missing mask. Inspired by GAN-based approaches, which train generators to decrease the predictability of missingness patterns, our method explicitly targets the reduction of mutual information.

Specifically, our algorithm iteratively minimizes the KL divergence between the joint distribution of the imputed data and missing mask, and the product of their marginals from the previous iteration. We show that the optimal imputation under this framework corresponds to solving an ODE, whose velocity field minimizes a rectified flow training objective.

We further illustrate that some existing imputation techniques can be interpreted as approximate special cases of our mutual-information-reducing framework. Comprehensive experiments on synthetic and real-world datasets validate the efficacy of our proposed approach, demonstrating superior imputation performance.