Thank you all very much for carefully reviewing our paper. Since the common feedback was to add ablation studies, in this document, we first give ablation study results. We will add these to the paper. Then we explain the advantage of TGAN against TVAE. And finally, we address individual questions and comments.

Ablation Study

Table 1: Ablation study results. Numbers are absolute performance change on real datasets.

Experiment	EXP1			EXP2		EXP3		
Model	`GMM5`	`GMM10`	`MinMax`	`w/o R.`	`w/o C.`	`GAN`	`WGANGP`	`GAN+PacGAN`
Performance	-4.1%	-8.6%	-25.7%	-17.8%	-36.5%	-6.5%	+1.75%	-5.2%

EXP1. Mode-specific normalization: In TGAN, we use variational Gaussian mixture model (VGM) to normalize continuous columns. We compare it with (1) `GMM5`: Gaussian mixture model with 5 modes, (2) `GMM10`: Gaussian mixture model with 10 modes, and (3) `MinMax`: min-max normalization to $[-1, 1]$. Using GMM slightly decreases the performance while min-max normalization gives the worst performance.

EXP2. Resampling and condition vector: We successively remove these two components. (1) `w/o R.`: we first disable resampling in training, but the generator still gets a condition vector and its loss function still has the cross-entropy term. The condition vector is sampled from training data frequency instead of log frequency. (2) `w/o C.`: We further remove the condition vector in the generator. These ablation results show that both resampling and condition vector are important for imbalanced datasets. Especially on highly imbalanced dataset such as `credit`, removing resampling results in 0% on F1 metric.

EXP3. WGANGP and PacGAN: In the paper, we use WGANGP+PacGAN. Here we compare it with three alternatives, WGANGP only, vanilla GAN loss only, and vanilla GAN + PacGAN. We observe that WGANGP is more suitable for synthetic data task than vanilla GAN, while PacGAN is helpful for vanilla GAN loss but not as important for WGANGP.

Why we put TGAN as the main method in our paper.

TGAN has a few advantages over TVAE, namely (1) since the generator in GAN is not directly optimized by mean square error, it's easier to make it differentially private using existing frameworks like DPGAN and PATE-GAN. Empirically, we compute the distance between synthetic data and nearest neighbor in training data. We observe TGAN gets 13% larger distance than TVAE, while achieving the same accuracy or F1 score on the real data. We will add this to the paper. (2) TGAN is more flexible in the sense that they are capable of capturing interactions amongst variables through their architecture, though TVAE is not intrinsically capable of doing so. To this end, in scenarios where strong complex underlying structures are involved, TGAN shall outperform TVAE.

Individual Comments and Questions

To Reviewer #1: (1) We agree with you and per your advice, we will remove unnecessary equations and use figures for NN architectures. We will move figure about data transformation process to the main paper. (2) Regarding GANs evaluation, given that the data employed here are not images, visual fidelity (which is the most common metric in image generation tasks) could not be applied. Moreover, Fréchet distance inception (Heusel et. al, 2017) could not be employed to all synthetic datasets given that it is applied on Gaussian distributed data. Hence, we were restricted to employ metrics that could be applied and are widely used in mixed data scenarios (Theis et. al, 2015). (3) Regarding TGAN convergence, theoretical guarantees for GANs to convergence to a Nash equilibrium are hard to derive in the case of continuous data, and harder for mixed data scenarios. Nevertheless, we have empirically checked that our algorithm does converge for a fixed set of hyperparameters regardless of random seeds. We recognize a recent work by Daskalakis et. al, 2017 and would extend it to provide theoretical results in the future.

To Reviewer #2: (1) We conducted a set of ablation studies as proposed in the review (results presented above). We will add these to paper. (2) All our code, real and benchmark datasets are publicly available. We developed our code as an open benchmarking framework. We had our anonymous repository hosted on `anonymous.4open.science`. The URL is at line 2 in the supplementary material. We are really sorry that the service was down recently and it's back online now.

To Reviewer #3: (1) We tried several hyperparameter sets for TVAE and TGAN. We will add this detail to the paper. In future there is scope for improvement with hyperparameter tuning. (2) In our revision, we will add figures to show generated synthetic data. We will add more references to recent advances in GAN. We have an additional page to add these.