Medi-Chal Project

Thomas Gerspacher & Adrien Pavao

April 13$^{\text{th}}$

Ritik presented a recent paper on evaluation metrics for GANs [1].

1 Metrics

Note that the metrics introduced here were tested on images' datasets and we still need to handle mixed-type variables' datasets and see how those metrics can perform.

1.1 Fréchet Inception Distance score

We call the Fréchet distance $d(.,.)$ between the Gaussian with mean (m, C) obtained from $p(.)$ and the Gaussian with mean (m_w, C_w) obtained from $p_w(.)$ the "Fréchet Inception Distance" (FID), which is given by:

$$d^2((m, C), (m_w, C_w)) = ||m - m_w||_2^2 + Tr(C + C_w - 2(CC_w)^{1/2}) \qquad (1)$$

For a generative model, $p(.)$ represents the distribution of model samples and $p_w(.)$ the distributions of the real samples.

The FID is robust to image disturbances.

1.2 1-Nearest Neighbor classifier

The concept of 1-NN is to classify X using the label of the closest neighbor among the training points X'.

1-NN needs a metric to find the nearest neighbor and a possibility could be to use Wasserstein distance.

1.3 Maximum Mean Discrepancy

We compute the mean distance between the distributions[1].

[1] http://alex.smola.org/teaching/iconip2006/iconip_3.pdf

Goal: Estimate $D(p, q, \mathcal{F})$

$$\mathbf{E}_{p,p}k(x, x') - 2\mathbf{E}_{p,q}k(x, y) + \mathbf{E}_{q,q}k(y, y')$$

U-Statistic: Empirical estimate $D(X, Y, \mathcal{F})$

$$\frac{1}{m(m-1)} \sum_{i \neq j} \underbrace{k(x_i, x_j) - k(x_i, y_j) - k(y_i, x_j) + k(y_i, y_j)}_{=:h((x_i, y_i), (x_j, y_j))}$$

With an infinite number of samples, the MMD only converge to 0 if the distributions are equal.

Implementations:
`https://github.com/Diviyan-Kalainathan/CausalDiscoveryToolbox/blob/ca9f3662fc1f22763e4efd2edffa720b8943f3dd/cdt/utils/loss.py`
`https://github.com/dougalsutherland/opt-mmd`
`https://github.com/dougalsutherland/mmd/`
`https://github.com/topics/maximum-mean-discrepancy`

1.4 Minimum Distance Accumulation

This idea was brought by Michèle Sebag. We have two distributions A and B. We allocate to each point from A the distance of its nearest neighbor from B.

Then we compute this graph : a distance θ on x axis and the number of points with a minimum distance smaller than θ on y axis. We can then define a privacy/resemblance trade-off: a threshold distance. The metrics are the areas under the curve on the left and on the right of the threshold.

- For a respect of **privacy** we want the left area the curve (x between 0 and the threshold) to be null. It means that no points from A has an exact match (or really close) from B.
- For **resemblance**, we want the right area of the curve (x from the threshold) to be maximal.

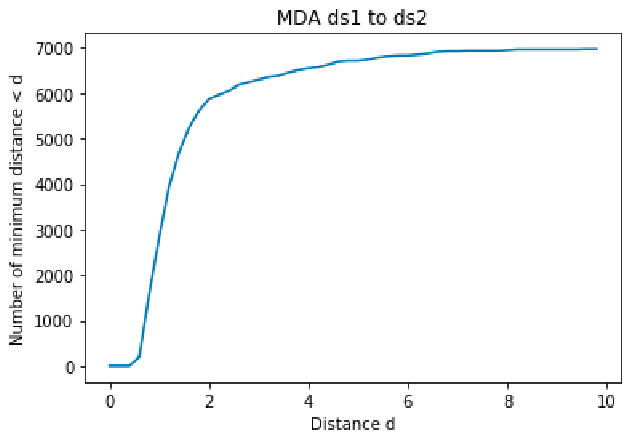

Figure 1: Example of MDA curve

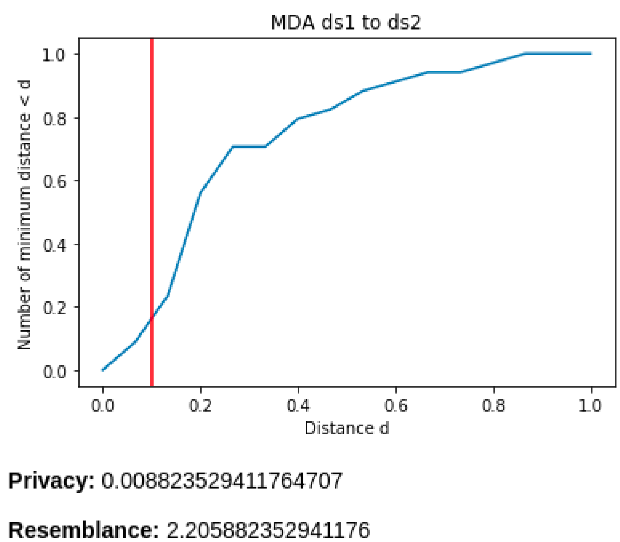

Figure 2: Example of MDA curve with threshold, normalization and areas under curve

2 Workplan

The next step will be to try out different generative models and test out the mentioned metrics to get a first intuition of how they perform. The goal is to fix a metric quickly in order to make further choice on the generative model.

We chose to implement Generative Adversarial Network, Random Forest imputations and Gaussian Copulas.

References

[1] Gao Huang et al. *An empirical study on evaluation metrics of generative adversarial networks*. 2018. URL: https://openreview.net/forum?id=Sy1f0e-R-.