Centre for
Frontier AI
Research

CFAR

Generative Adversarial Ranking Nets

Yao Yinghua

Research Scientist, CFAR, IHPC, A*STAR, Singapore

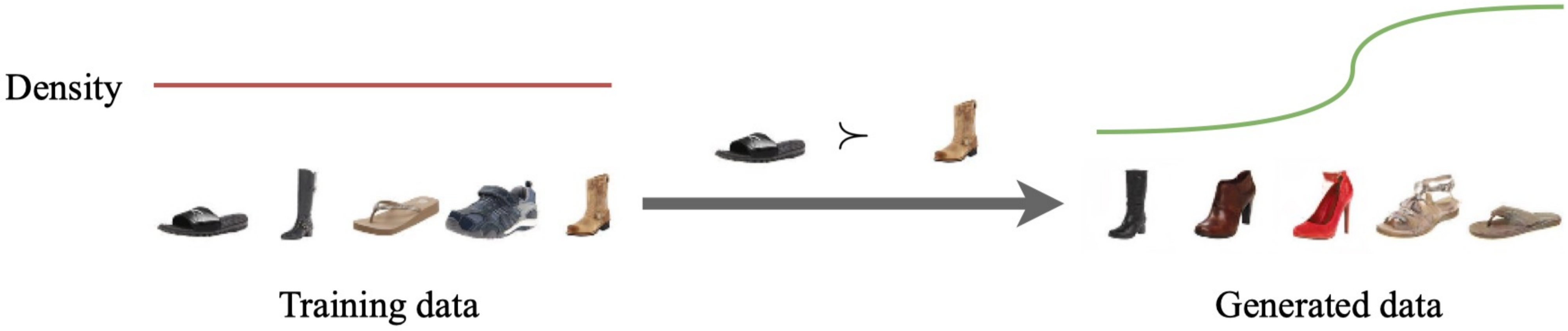
Joint work with Yuangang Pan, Jing Li, Ivor W. Tsang, Xin Yao

CREATING GROWTH, ENHANCING LIVES

Generative Modeling conditioned on Human Preferences

Model a conditional distribution $p(x|C)$: C is a single condition ✓

How about conditioning on a comparison $S \coloneqq s^1 > s^2 > \cdots > s^l, s^i \in X = \{x_n\}_{n=1}^{\mathrm{N}}, l < \mathrm{N}$?



Given: training samples $\mathrm{X} = \{x_n\}_{n=1}^{N}$; listwise preferences $\mathrm{S} = \{s_m\}_{m=1}^{M}$

$$s := s^1 > s^2 > \ldots > s^l, \quad s^i \in \mathrm{X}, \ \forall i = 1, 2, \ldots, l, l \ll N. \tag{1}$$

Learning goal: the distribution learned by the generative model $P_{\mathrm{g}}(x)$ is identical to the distribution of the desired data $P_{\mathrm{u}}(x)$, i.e., $P_{\mathrm{g}}(x) = P_{\mathrm{u}}(x)$.

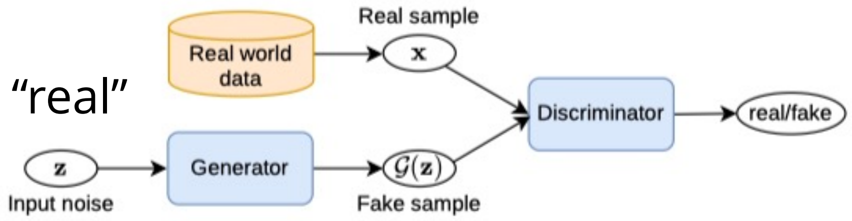
- $P_{\mathrm{u}}(x)$ allocates high density to high-ranked data while low or zero density to low-ranked data.

Y. Yao, Y. Pan, J. Li, I. W. Tsang, X. Yao. Generative Adversarial Ranking Nets. Journal of Machine Learning Research (2024).

Generative Adversarial Ranking Nets (GARNet)

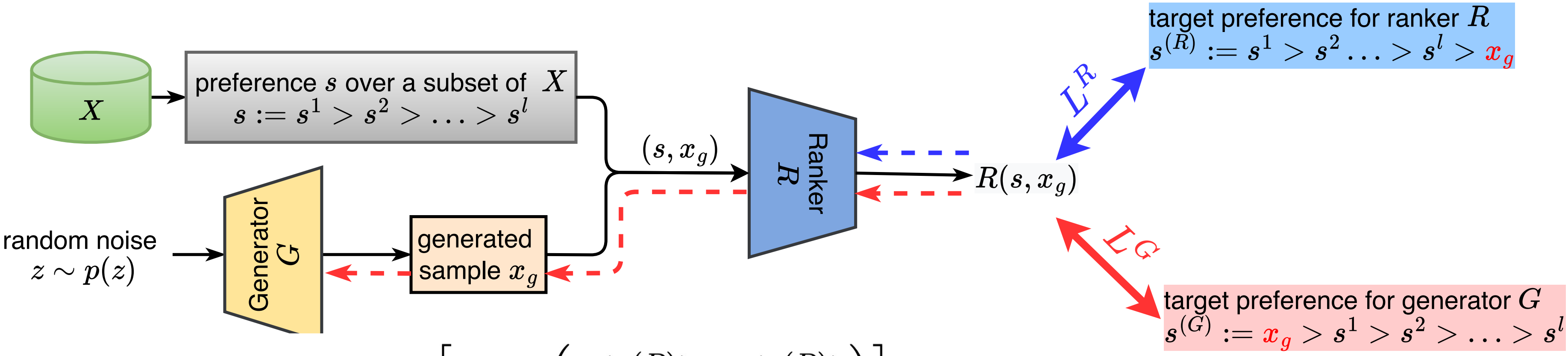
The adversarial goal in GAN:

- Generator fools discriminator to classify generated samples as “real”
- Discriminator classifies generated samples as “fake”



Motivate our generative adversarial ranking:

- Generator fools ranker to rank generated samples x_g “top”
- Ranker ranks generated samples x_g “bottom”



Ranker $\sup_{R:\mathcal{X}\to\mathbb{R}} \mathbb{E}_{\substack{s\sim\mathcal{S}\\ x_g\sim P_g}} \left[\mathcal{L}_{L2R}\left(\pi\left(s^{(R)}\right), R\left(s^{(R)}\right)\right)\right]$

Generator $\sup_{G:\mathcal{Z}\to\mathcal{X}} \mathbb{E}_{\substack{s\sim\mathcal{S}\\ x_g\sim P_g}} \left[\mathcal{L}_{L2R}\left(\pi\left(s^{(G)}\right), R\left(s^{(G)}\right)\right)\right]$

$$\mathcal{L}_{L2R}(\pi(s), R(s)) = \sum_{i=1}^{l} \sigma(\pi(s))_i \log \sigma(R(s))_i$$

GARNet Defines a Relativistic f-Divergence

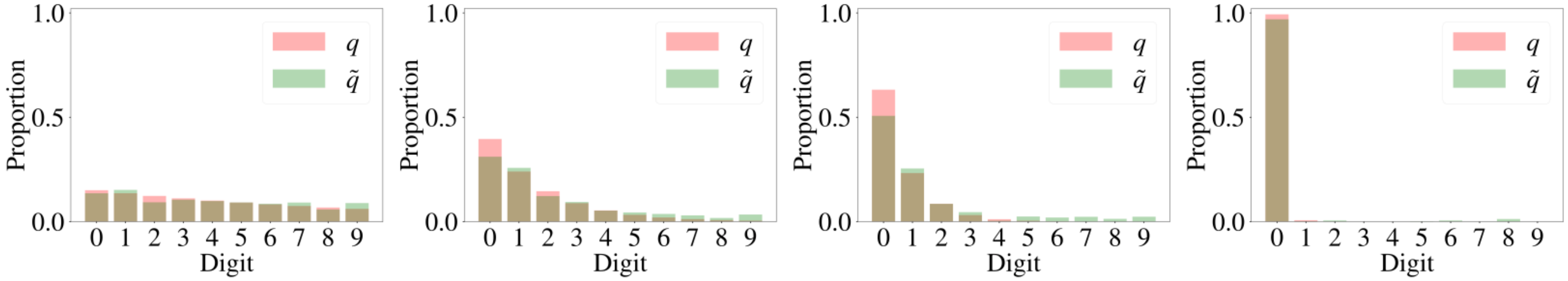
The optimal ranker in GARNet estimates a relativistic f-divergence between the user-preferred data distribution P_{u} and the generated data distribution P_{g} :

$$D_f(P_{\mathrm{u}}, P_{\mathrm{g}}) = \sup_{R:\mathrm{X}\to\mathbb{R}} \mathop{\mathbb{E}}_{\substack{s\sim\mathrm{S} \\ x_{\mathrm{g}}\sim P_{\mathrm{g}}}} \left[\mathcal{L}_{L2R}\left(\pi\left(s^{(R)}\right), R\left(s^{(R)}\right)\right)\right]$$

Different from GAN: $D_f(P_{data}, P_{\mathrm{g}})$

distribution of given training data

- P_{u} allocates higher density to higher-ranked data
- $\pi\left(s^{(R)}\right)$ is arithmetic sequence with a common difference d $\quad \pi\left(s^{(R)}\right) = [a+(T-1)d, a+(T-2)d, \ldots, a, b]$
- For a sufficient large d, GARNet converges to the distribution of top-ranked data



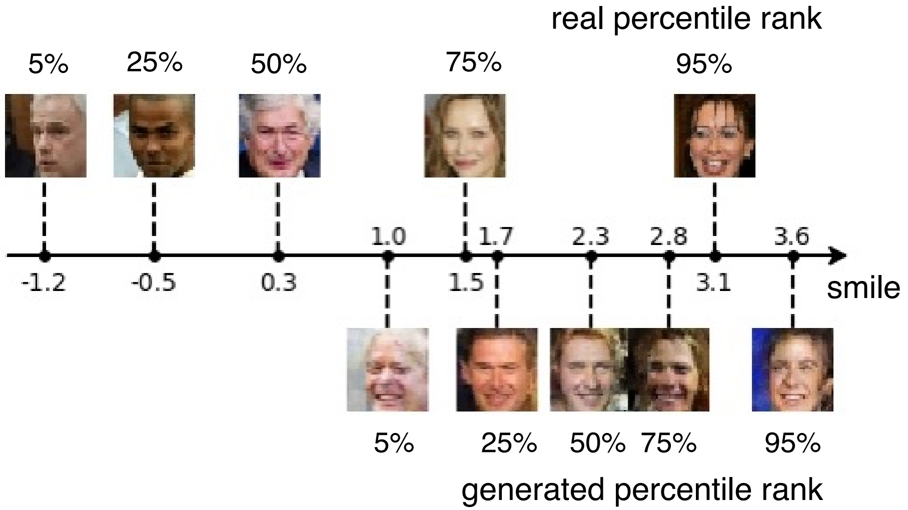
$d = 0.1$ $\quad$ $d = 0.5$ $\quad$ $d = 1$ $\quad$ $d = 5$

MNIST (preferring small digits, i.e., 0 > 1 > 2 > . . . > 9).

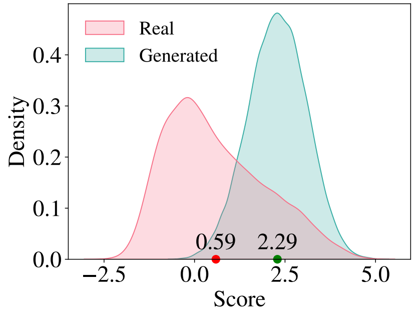
$q = \sigma(\pi(s))$ calculates the user-specified top-1 probability of each digit.

$\tilde{q}$ calculates the proportion of different digit classes for generated data.

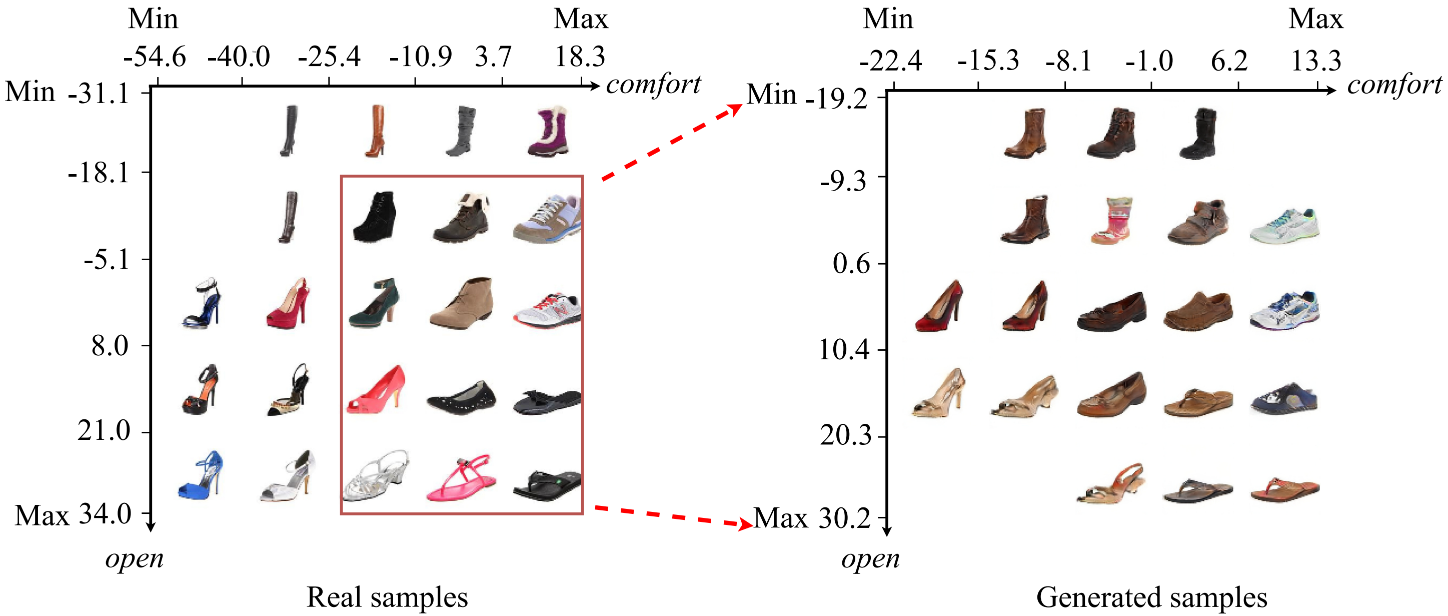
GARNet with Continuous Properties (single/multiple)



Smiling faces are preferred. The percentile rank of a given score is the percentage of scores in its frequency distribution that are less than that score.



Both open and comfortable shoes are preferred simultaneously.

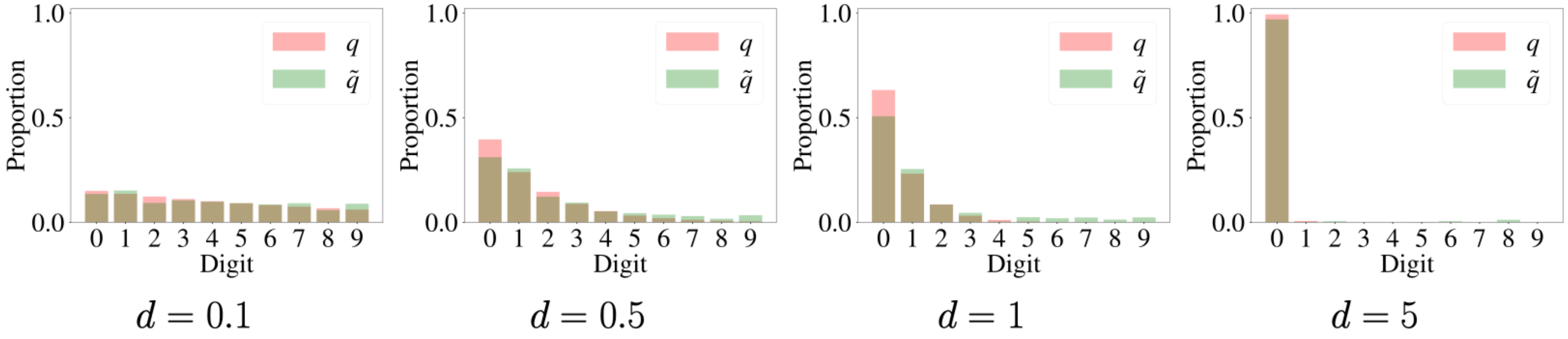


Discussions

- The generated data distribution converges to the user-preferred data distribution based on **full/partial** preferences involving **single**/**multiple discrete**/**continuous** properties, supported by theoretical guarantees.

- Diffusion model with preference alignment is proven to converge to the distribution of top-ranked data

- Our GARNet demonstrates distribution learning of user-preferred data in a finer-grained manner.



CREATING GROWTH, ENHANCING LIVES

Centre for Frontier AI Research

CFAR

THANK YOU

www.a-star.edu.sg