Generative Adversarial Networks for Time Series

Autonomous Multisensor Systems Group
Institute for Intelligent Cooperating Systems
Faculty of Computer Science
Otto von Guericke University, Magdeburg

25.09.2024

Team Members: Archana Yadav, Gowtham Premkumar, Shweta Bambal

Contents

Evaluation Metrics for GAN

Evaluation metrics ensure that synthetic data closely resembles real data in terms of realism and feature distribution.

- **Qualitative evaluation:** Human visual inspection of the generated data.
    - PCA Plot
    - t-SNE Plot
- **Quantitative evaluation:** Use of metrics associated with statistical measures used for time series analytics
    - Discriminative Score
    - Predictive Score
    - Frechet Inception Score

Discriminative Score Methodology

Train-Test Split

- Combined and labeled the real and synthetic data
- Labelled real data as 1 and synthetic as 0
- Splitted it into training and testing sets.

Discriminator Training

- Trained a discriminator to classify the data based on the labels
- Used the same discriminator model as the one used for synthetic data generation
- Used binary cross entropy loss

Score Calculation

- Calculated the discriminator's accuracy on the test set
- Computed **discriminative score** as **|accuracy - 0.5|**
- Lower score indicated better synthetic data quality

Predictive Score Methodology

Training on Synthetic Data → **Testing on Real Data** → **Score Calculation**

Training on Synthetic Data

- Trained a RNN on synthetic data

Testing on Real Data

- Predicted the value of the last feature of the next sequence for the real data using the trained model

Score Calculation

- Calculated the mean absolute error (MEA) between the predicted and the original feature value of the real data
- **Average MEA** is the **predictive score**

Frechet Inception Distance (FID) Score Methodology

Extract Features → **Calculate Mean & Covariance** → **Calculate Frechet Distance**

Extract Features

- Used **InceptionTime** neural network popular for time series data.
- The goal is to capture distributions of the extracted features.

Calculate Mean & Covariance

- **Mean** represents the **central tendency** of the feature vectors
- Mean helps to assess how closely two distributions are centered around the same point in feature space.
- **Covariance matrix** gives information about **shape and spread** of the feature distribution.

Calculate Frechet Distance

- **A large Mean difference** Synthetic data not centered around the same values as the real data
- **A Large Cov differences** Synthetic data doesn’t match the diversity or variation of the real data

$$FID(\mu_r, \Sigma_r, \mu_g, \Sigma_g) = \|\mu_r - \mu_g\|_2^2 + Tr\left(\Sigma_r + \Sigma_g - 2\left(\Sigma_r \Sigma_g\right)^{\frac{1}{2}}\right)$$

[3]

Our Results

Model	Predictive Score	Discriminative Score	FID Score
TimeGAN (GRU)	0.0111	0.4033	0.1189
TimeGAN (LSTM)	0.0075	0.3634	0.0678
TransformerGAN	0.5024	0.5148	0.5134
WGAN	0.5645	0.5841	0.6867

Conclusion

- **TimeGAN (LSTM)** achieves the best overall performance for all three metrics.
- In conclusion, the summary of evaluation metrics is as follows:

Criterion	Discriminative Score	Predictive Score	FID Score
Definition	How well a discriminator distinguishes real from fake samples	How well a model predicts any particular features in the next time step based on the history and other features in current timestamp.	Measures the distance between the distributions of generated and real images.
Purpose	Assess quality of the discriminator in a GAN.	Evaluate the semantic correctness of generated samples.	Quantify the similarity of generated and real data distributions
Common Use Cases	Evaluating GANs during training to adjust parameters.	Fine-tuning and comparison of GANs	Benchmarking and comparing different GAN models in research.

Way Forward

1. Further **tuning of the hyperparameter** for better model optimisation and performance.
2. Development of **interactive visualizations** to compare the real and generated time-series data.

References

1. Ricardo de Deijn, Aishwarya Batra, Brandon Koch, Naseef Mansoor and Hema Makkena. (2024). Reviewing FID and SID metrics on Generative Adversarial Networks. https://arxiv.org/pdf/2402.03654
2. Xiaomin Li, Vangelis Metsis, Huangyingrui Wang, Anne Hee Hiong Ngu. (*2022*) TTS-GAN: A Transformer-based Time-Series Generative Adversarial Network. https://arxiv.org/abs/2202.02691
3. Yoon, Jinsung & Jarrett, Daniel & Schaar, Mihaela. (2019). Time-series Generative Adversarial Networks. https://www.researchgate.net/publication/344464212_Time-series_Generative_Adversarial_Networks
4. Padmanaba Srinivasan, William J. Knottenbelt. (2022) Time-series Transformer Generative Adversarial Networks https://arxiv.org/abs/2205.11164
5. EOIN BROPHY, ZHENGWEI WANG, QI SHE, TOMÁS WARD. (2023) Generative Adversarial Networks in Time Series: A Systematic Literature Review https://dl.acm.org/doi/fullHtml/10.1145/3559540
6. H. Arnout, J. Bronner and T. Runkler, "Evaluation of Generative Adversarial Networks for Time Series Data," 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 2021, pp. 1-7, doi: 10.1109/IJCNN52387.2021.9534373.
7. Brophy, Eoin et al. “Generative adversarial networks in time series: A survey and taxonomy.” ArXiv abs/2107.11098 (2021): n. pag.,
8. https://www.techtarget.com/searchenterpriseai/tip/GAN-vs-transformer-models-Comparing-architectures-and-uses
9. https://towardsdatascience.com/evaluation-of-synthetic-time-series-1b4fc4e2be39https://machinelearningmastery.com/how-to-implement-the-frechet-inception-distance-fid-from-scratch/
10. https://wandb.ai/ayush-thakur/gan-evaluation/reports/How-to-Evaluate-GANs-using-Frechet-Inception-Distance-FID---Vmlldzo0MTAxOTI

Generative Adversarial Networks for Time Series

Thank You For Your Attention!

Autonomous Multisensor Systems