Diffusion Transformers for Tabular Data Time Series Generation

Fabrizio Garuti[1,3], Enver Sangineto[3], Simone Luetto[2], Lorenzo Forni[1], Rita Cucchiara[3]

1 Prometeia Associazione, Bologna | 2 Prometeia SpA, Bologna | 3 University of Modena and Reggio Emilia

Singapore, 24 April 2025

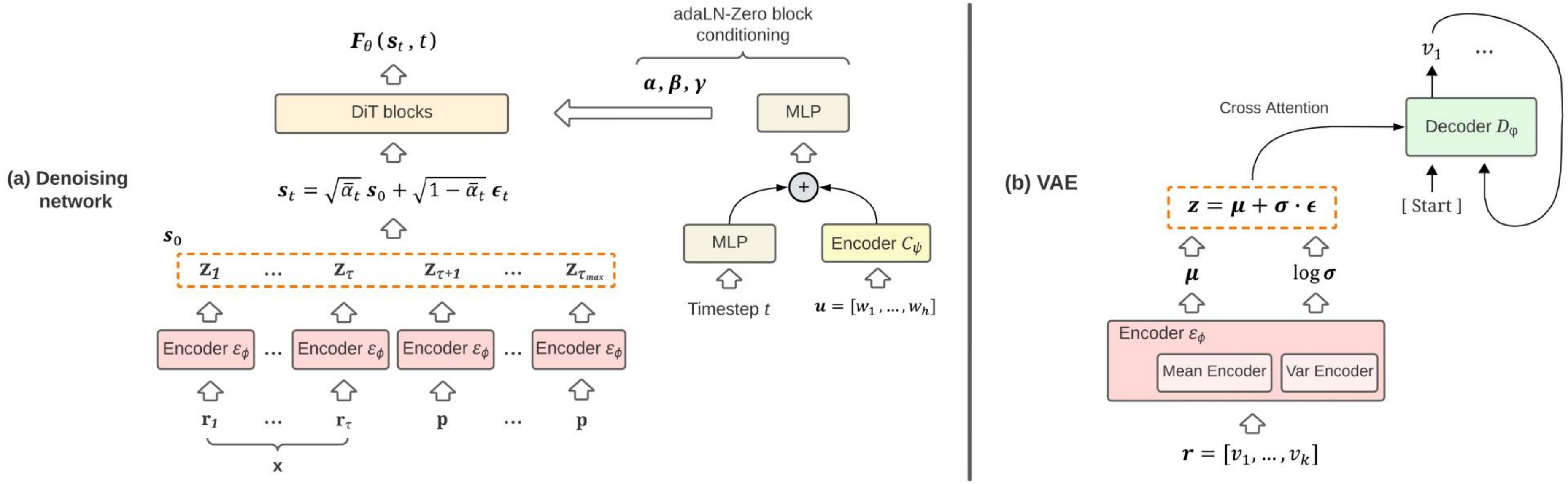
Introduction & Motivation

- **Tabular time series data** is crucial in many real-life applications (e.g., finance, healthcare).
- Real-world datasets are often unavailable due to **privacy** and legal constraints.
- Traditional generative models struggle with tabular data **heterogeneity** and **variable sequence lengths**.
- GANs and VAEs methods cannot easily represent mixed numerical and categorical data.
- AR Transformer models like TabGPT [1] and REaLTabFormer [2] lack diversity in generation.
- Current methods fail to generate fully **unconditional**, high-quality tabular time series.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	User	Card	Year	Month	Day	Time	Amount	Use Chip	Merchant Name	Merchant City	Merchant State	Zip	MCC	Is Fraud?	Errors?
2	0	0	2019	4	1	4:08	$39.58	Chip Transaction	Kelly Auto Repair	Brandon	FL	33510	7538	No	
3	0	0	2019	4	2	4:12	$36.98	Chip Transaction	Kelly Auto Repair	Brandon	FL	33510	7538	No	
4	0	0	2019	4	4	4:20	$40.36	Chip Transaction	Kelly Auto Repair	Brandon	FL	33510	7538	No	
5	0	0	2019	4	4	10:54	$7.98	Chip Transaction	Walmart	Brandon	FL	33510	5311	No	
6	0	0	2019	4	5	4:46	$49.87	Swipe Transaction	Nissan Service	Brandon	FL	33510	7538	No	
7	0	0	2019	4	5	10:59	$11.27	Chip Transaction	Applebees	Brandon	FL	33511	5812	No	

[1] Padhi et al. *"Tabular transformers for modeling multivariate time series". ICASSP, 2021.*
[2] Solatorio & Dupriez. *"REaLTabFormer: Generating realistic relational and tabular data using transformers". arXiv, 2023.*

Our Approach - TabDiT (Tabular Diffusion Transformer)



- Leverages **Latent-space Diffusion Model** (LDM) with a Transformer denoising network.
- Uses an **AutoRegressive Variational Autoencoder** (VAE) to compress individual tabular rows into a single embedding vector and to preserve the intra-row dependencies.
- The Transformer denoising network generates diverse and realistic sequences while handling **variable-length** time series.
- Novel **Numerical Representation**: uses a variable-range decimal representation to accurately generate values ranging from a few cents to tens of millions.

Experimental Results

- **Conditional and unconditional** tasks evaluated on **six public datasets** from finance, healthcare and retail.
- Novel metric: **Machine Learning Detection** (MLD) score evaluates the temporal coherence of an entire time series. Previous metrics evaluate the generated data only at the individual row level.
- Using both our metrics and traditional metrics, TabDiT consistently and significantly outperforms state-of-the-art models, achieving higher realism and diversity.

Table 2: Unconditional generation results on *Rossmann*, *Airbnb* and *PKDD'99*.

Method	Rossmann		Airbnb		PKDD'99 Financial	
	MLD-TS ↓	LD-SR ↑	MLD-TS ↓	LD-SR ↑	MLD-TS ↓	LD-SR ↑
AR Baseline (ours)	97.80±2.20	49.97±3.26	77.23±1.46	56.43±2.80	92.87±1.68	71.93±1.34
TabDiT (ours)	**82.60**±3.92	**77.07**±5.37	**55.07**±3.52	**78.07**±2.77	**85.53**±4.18	**79.10**±6.09

Table 3: Unconditional generation results on *Age2*, *Age1* and *Leaving*.

Method	Age2		Age1		Leaving	
	MLD-TS ↓	LD-SR ↑	MLD-TS ↓	LD-SR ↑	MLD-TS ↓	LD-SR ↑
AR Baseline (ours)	67.53±0.75	83.47±0.38	91.20±0.46	74.23±1.06	69.43±4.02	75.33±2.86
TabDiT (ours)	**50.43**±1.85	**87.00**±1.54	**63.93**±3.20	**76.00**±4.25	**62.33**±0.99	**75.63**±4.20

Table 4: Conditional generation results.

Method	Task	Rossmann		Airbnb		Age2		PKDD'99 Financial	
		MLD-TS ↓	LD-SR ↑	MLD-TS ↓	LD-SR ↑	MLD-TS ↓	LD-SR ↑	MLD-TS ↓	LD-SR ↑
SDV	child	99.63 ±0.64	6.53*±0.39	93.30 ±0.61	0.00*±0.00	96.03 ±0.11	44.80±1.73	97.95 ±1.42	6.53±0.58
	merged	100.00 ±0.00	2.80*±0.25	94.40 ±1.65	0.00*±0.00	96.27 ±0.06	37.63±1.47	98.12 ±1.17	8.77±0.59
REaLTabFormer	child gt-cond	98.90±1.10	60.63±2.65	63.63±1.20	86.17 ±1.29	66.77 ±0.42	77.90±0.85	97.87±0.59	21.97±0.55
	child	64.83 ±1.33	52.08 *±0.89	57.77 ±0.67	30.48*±0.79	52.97 ±2.32	77.30 ±0.92	59.33 ±3.82	21.50±0.72
	merged	74.43 ±8.85	28.33 *±2.31	76.97 ±2.04	21.43 *±1.10	52.10 ±2.17	75.53 ±0.65	58.77 ±3.05	26.00±1.61
AR Baseline (ours)	child gt-cond	95.57 ±1.96	71.60 ±2.42	57.97 ±1.72	82.77±0.49	69.97±0.90	80.73 ±0.59	68.33 ±4.37	81.13 ±1.51
	child	99.63±0.64	36.03±8.79	82.33±1.53	62.53 ±3.93	79.03±1.62	65.83±3.95	79.07±6.23	67.60 ±4.36
	merged	99.63±0.64	19.70±6.80	93.50±1.30	8.53±2.49	81.30±0.78	48.03±3.61	83.33±5.75	38.73 ±4.22
TabDiT (ours)	child gt-cond	**72.20**±1.10	**82.90**±1.32	**51.10**±2.60	**98.07**±0.25	**51.40**±2.95	**84.60**±1.87	**59.50**±10.53	**81.20**±2.71
	child	**64.03**±0.64	**80.13**±3.02	**49.33**±1.18	**81.10**±0.98	**50.47**±1.71	**84.70**±1.21	**51.80**±6.44	**79.13**±3.04
	merged	**71.83**±2.77	**38.63**±1.04	**54.63**±0.85	**47.37**±2.68	**51.53**±3.04	**78.93**±0.64	**54.03**±3.95	**53.20**±0.66

prometeia

[1] Solatorio & Dupriez. *"REaLTabFormer: Generating realistic relational and tabular data using transformers". arXiv, 2023.*

Conclusion

- We presented TabDiT, an **LDM approach for tabular data** time series generation.
- Using extensive experiments with **six different public datasets**, we showed the effectiveness of TabDiT, which largely outperfoms the other baselines in both conditional and unconditional tasks.
- Specifically, TabDiT is the first network showing results for **unconditional** generation of time series of tabular data with **heterogeneous** field values.

Thank you!

For any questions: fabrizio.garuti@prometeia.com