MITSUBISHI ELECTRIC
Changes for the Better

The 39th Annual AAAI Conference on Artificial Intelligence
FEBRUARY 25 – MARCH 4, 2025 | PHILADELPHIA, PENNSYLVANIA, USA

Quantum Diffusion Models for Few-Shot Learning

Ruhan Wang

Ye Wang

Jing Liu

Toshiaki Koike-Akino

MITSUBISHI ELECTRIC RESEARCH LABORATORIES (MERL)
Cambridge, Massachusetts, USA
http://www.merl.com

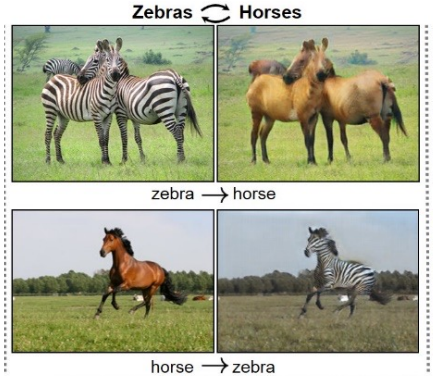
Outline

- Background of Generative AI
- Overview of Quantum AI
- Introduction of Quantum Diffusion Models
- Proposed Methods for Quantum Few-Shot Learning
- Experiments
  - 5 to 30% accuracy improvement
- Summary

Generative AI: Surpassing Human-Level Creativity

- VAE
- GAN
- Diffusion

[Karras et al, 2018]



StyleGAN [Karras et al, 2019]

Diffusion [Dhariwal et al., 2021]

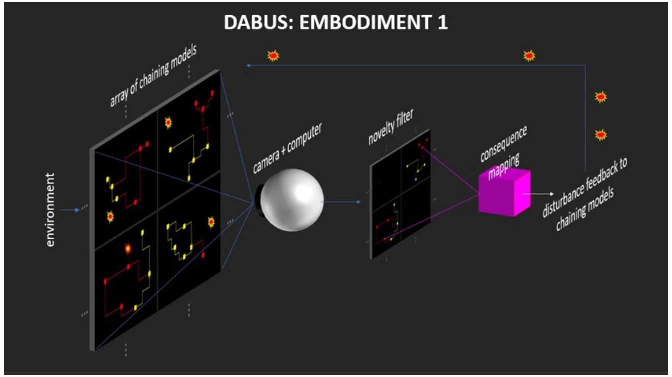
Dall-E 2 [Ramesh et al. 2022]

Imagen [Saharia et al. 2022]

StableDiffusion [Rombatch et al. 2022]

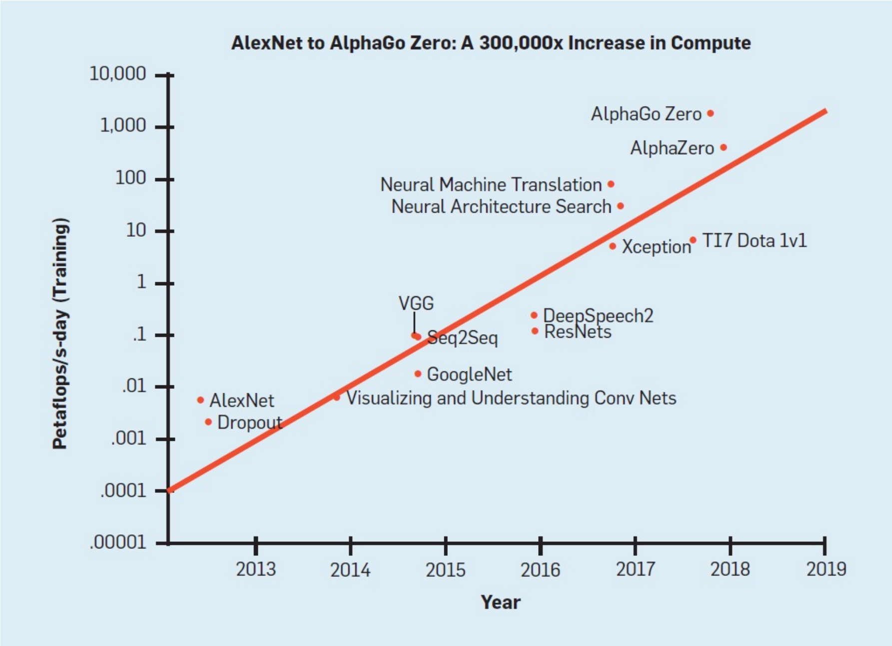
Generative AI: Surpassed Human-Level Creativity

- AI won some art prizes and inventions
  - Kamome Ashizawa’s **AI-generated novel**, “Are you there?”, took the Hoshi-Shinichi Literary Award in Nikkei press, Feb. 2022
  - Jason Allen’s **AI-generated painting**, “Théâtre D’opéra Spatial,” took first place in the digital category at the Colorado State Fair. Sep. 2022
  - DABUS **AI-generated patents** granted in South Africa, Jul. 2022
  - Boris Eldagsen’s **AI-generated photo**, “The Electrician” came top in open competition at the World Photography Organization’s Sony World Photography Awards, Apr. 2023.

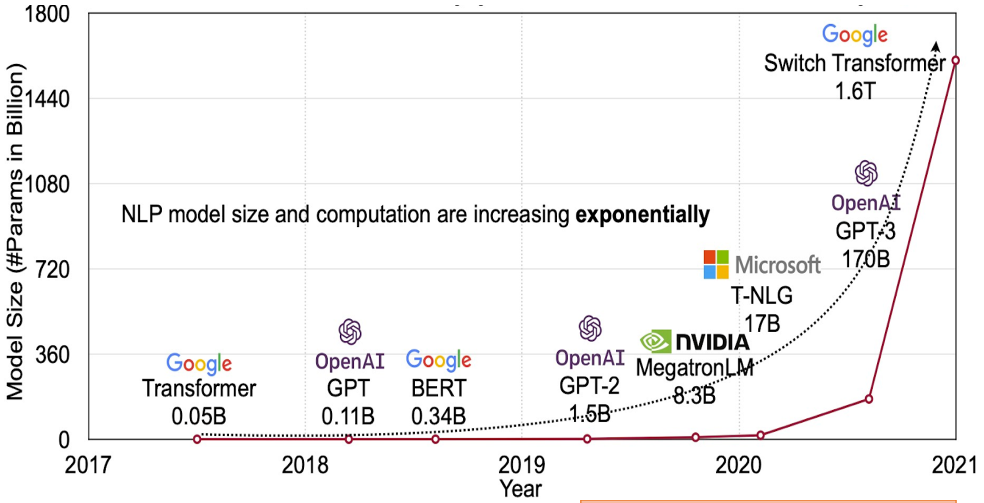


Green vs. Red AI

- Roy Schwartz, Jesse Dodge, Noah A. Smith, Oren Etzioni, “Green AI”, Communications of the ACM, December 2020, Vol. 63 No. 12, Pages 54-63 10.1145/3381831: https://cacm.acm.org/magazines/2020/12/248800-green-ai/



The computation used to train deep learning models has increased 300,000x in six years: nearly 10x annually



GPT-3: 175B params
GPT-4: ~1.7T params

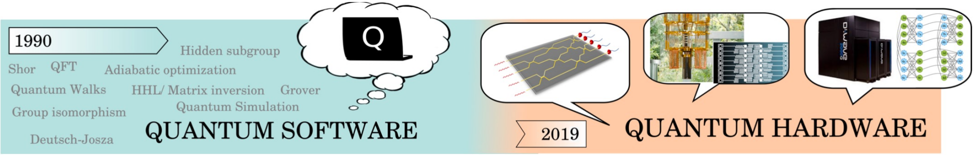
Deep Learning Crisis for Sustainable Growth

- Escalating power consumption of DNN training
  - [Strubell et al. Energy and policy considerations for deep learning in NLP. 2019]
  - DNN training with network architecture search (NAS) on GPUs requires **5-fold higher** carbon emission of single car lifetime!
- Therefore, we should consider **Green AI**
  - Efficient, fast, low-power, lightweight AI
  - New computing modality alternative to CPU/GPU/TPU: Natural computing (**Quantum**)

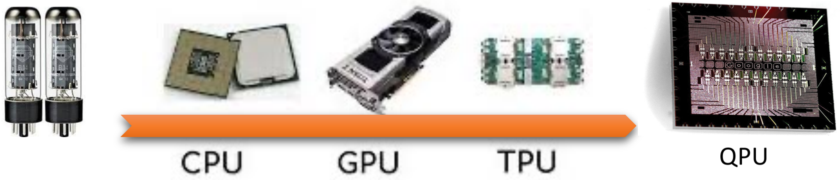
Consumption	**CO_2e (lbs)**
Air travel, 1 passenger, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000
Training one model (GPU)	
NLP pipeline (parsing, SRL)	39
w/ tuning & experimentation	78,468
Transformer (big)	192
w/ neural architecture search	626,155

Quantum Computing

- Morgan Stanley: Quantum tech. can drive **4th industrial revolution**
- Escalating government funds: National Quantum Initiative
- Quantum processor providers: **IBM**, **Google**, **Microsoft**, **Honeywell**, Intel, IONQ, rigetti, …
- Quantum cloud services: IBMQ, Amazon Bracket, …
- Free libraries to evaluate quantum computing on realistic simulators or real devices

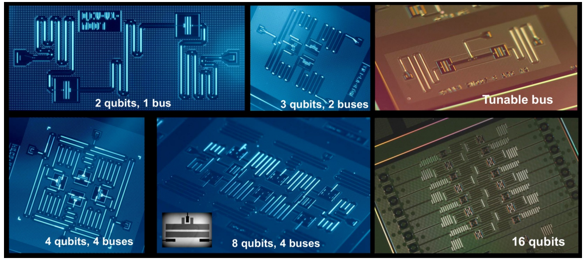


QISKit
Quantum Information Software Kit
PYQUIL
Cirq
QuTiP
Quantum Toolbox in Python
PROJECTQ
PENNYLANE

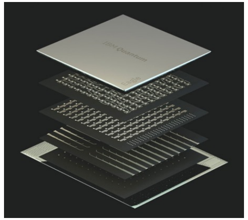


intel Microsoft
rigetti Honeywell IBM
Google

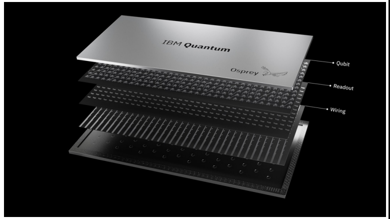
Evolution of Quantum Processing Unit (QPU)



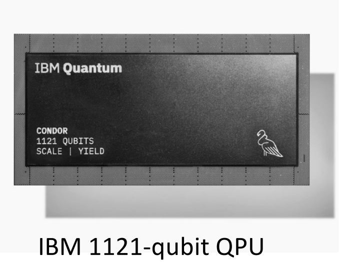
- Rapid QPU development to allow many qubits
  - IBM released **127-qubit** QPUs in Nov. 2021
  - IBM released **433-qubit** QPUs in Nov. 2022
  - IBM released **1121-qubit** QPUs in Dec. 2023 (**4158-qubits** by 2025)



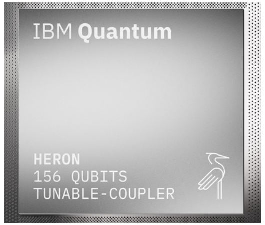
IBM 127-qubit QPU
(Nov. 2021)



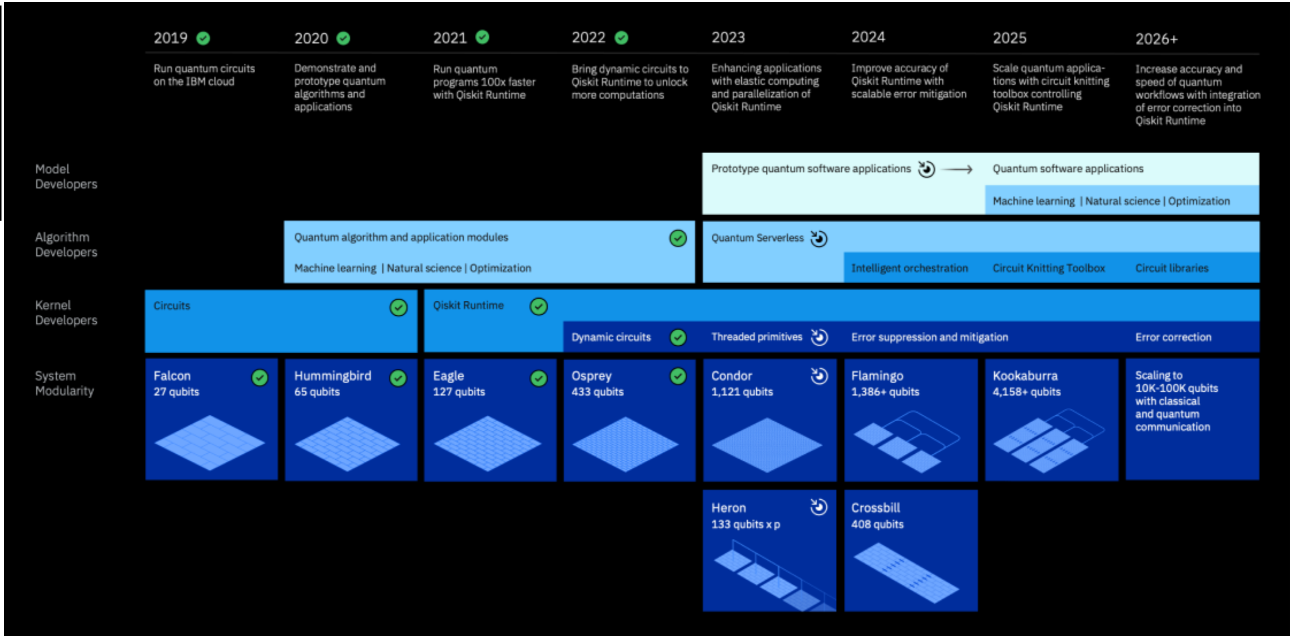
IBM 433-qubit QPU
(Nov. 2022)



IBM 1121-qubit QPU
(Dec. 2023)

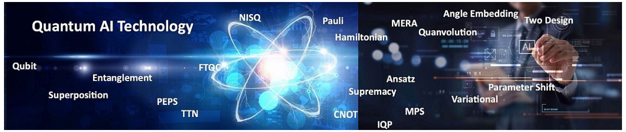


IBM 156-qubit QPU
(Nov. 2024): 5000 gates

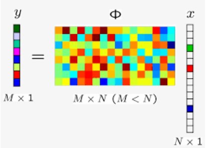


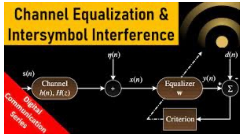
IBM QPU development roadmap (as of 2022)

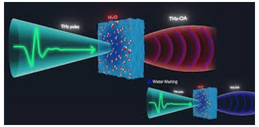
MERL's QML Research

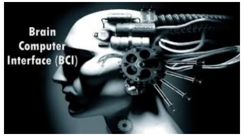


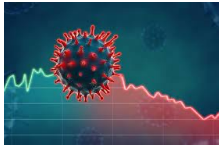
- MERL QML research highlight:
  - https://www.merl.com/research/highlights/quantum-ai
  - QHack: 2022 AWS award; IBM award; 2023 Nvidia award; 3rd prize
- Various Industrial QML Applications
  - Matsumine, T., Koike-Akino, T., Wang, Y., **"Channel Decoding with Quantum Approximate Optimization Algorithm"**, *ISIT*, July 2019.
  - Koike-Akino, T., Matsumine, T., Wang, Y., Millar, D.S., Kojima, K., Parsons, K., **"Variational Quantum Demodulation for Coherent Optical Multi-Dimensional QAM"**, *OFC/NFOEC*, March 2020, pp. T3D.6.
  - Koike-Akino, T., Wang, P., Wang, Y., "**Quantum Transfer Learning for Wi-Fi Sensing**", ICC, June 2022.
  - Liu, B., Koike-Akino, T., Wang, Y., Parsons, K., "**Variational Quantum Compressed Sensing for Joint User and Channel State Acquisition in Grant-Free Device Access Systems** ", ICC, June 2022.
  - Koike-Akino, T., Wang, P., Wang, Y., "**AutoQML: Automated Quantum Machine Learning for Wi-Fi Integrated Sensing and Communications**", SAM, Aug. 2022.
  - Koike-Akino, T., et al., "**Quantum Feature Extraction for THz Multi-Layer Imaging**", IRMMW-THz, Aug. 2022.
  - Koike-Akino, T., Wang, Y., "**quEEGNet: Quantum AI for Biosignal Processing**", BHI, Sep. 2022.
  - Koike-Akino, T., Wang. P., "**Post-Deep Learning Era: Emerging Quantum Machine Learning for Sensing and Communications**", GLOBECOM, Dec. 2022.
  - Liu, B., Koike-Akino, T., Wang, Y., Parsons, K., "**Learning to Learn Quantum Turbo MIMO Detection**", arXiv, 2022.
  - Koike-Akino, T., "COVID-19 Quantum Forecasting", QHack, Mar. 2022.
  - Koike-Akino, T., "Quantum mixed reality (XR)", QHack, Mar. 2023.
  - Ahmed, M.R., Koike-Akino, T., Parsons, K., Wang, Y., "**AutoHLS: Learning to Accelerate Design Space Exploration for HLS Designs**", MWSCAS, Aug. 2023.





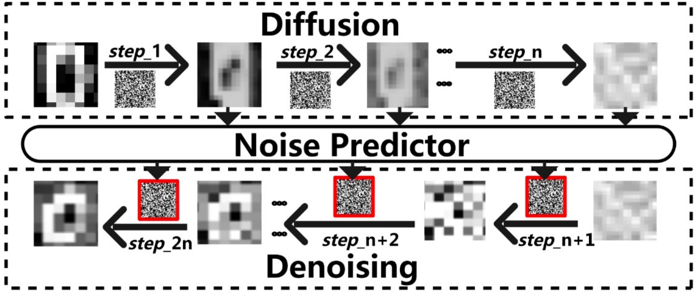


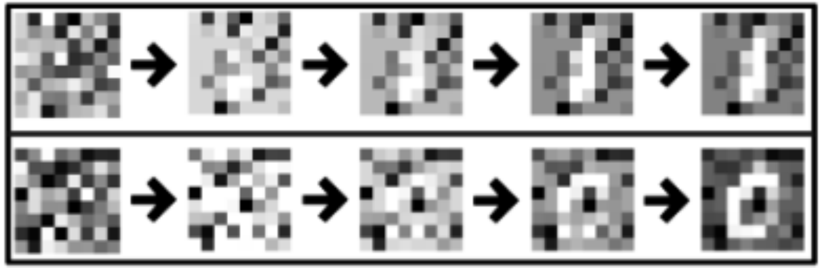




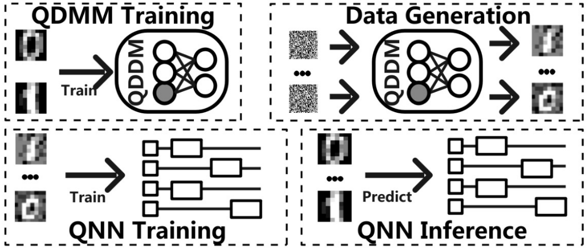
Quantum Diffusion Models

- Quantum denoising diffusion models (QDDM)

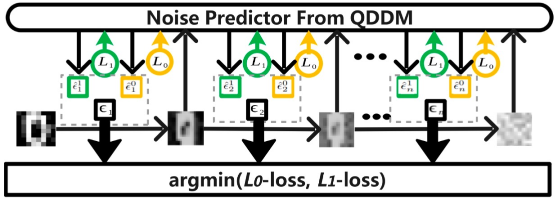




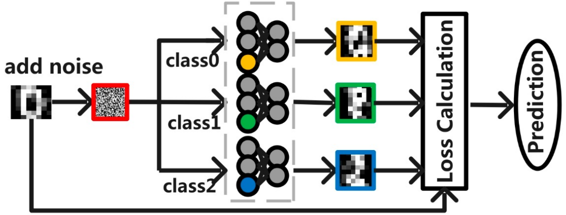
- Quantum few-shot learning: 3 different approaches



Generation inference



Diffusion inference

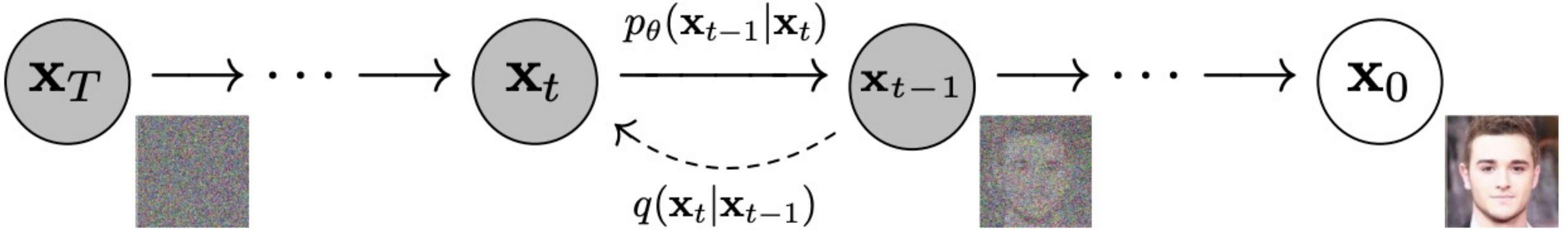


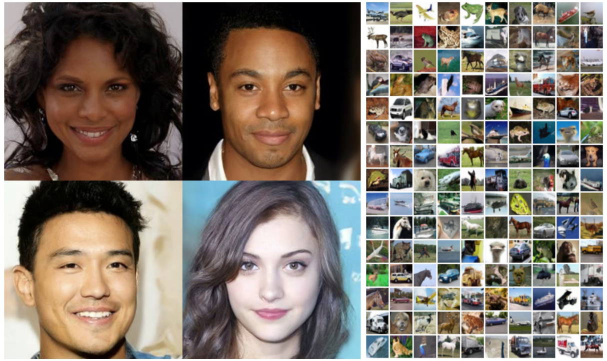
Denoising inference

- 5-30% accuracy improvement

Diffusion Models

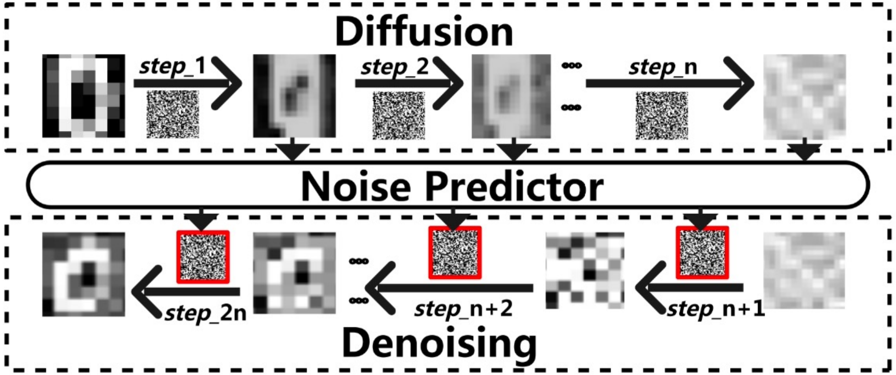
- Denoising Diffusion Probabilistic Model (DDPM) [Ho 2020] is a pioneering generative AI model, outperforming VAE and GAN
  - 2 processes: Diffusion steps; Denoising steps
  - Variants: Implicit Diffusion [Song 2020]; Latent Diffusion [Rombach 2022]; Guided Diffusion [Ho 2022]

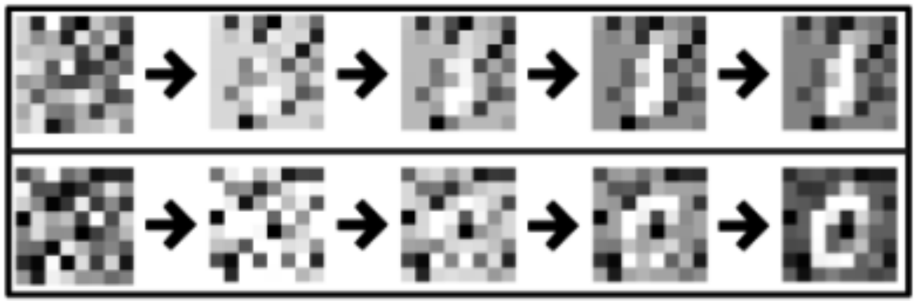




Quantum Denoising Diffusion Model (QDDM)

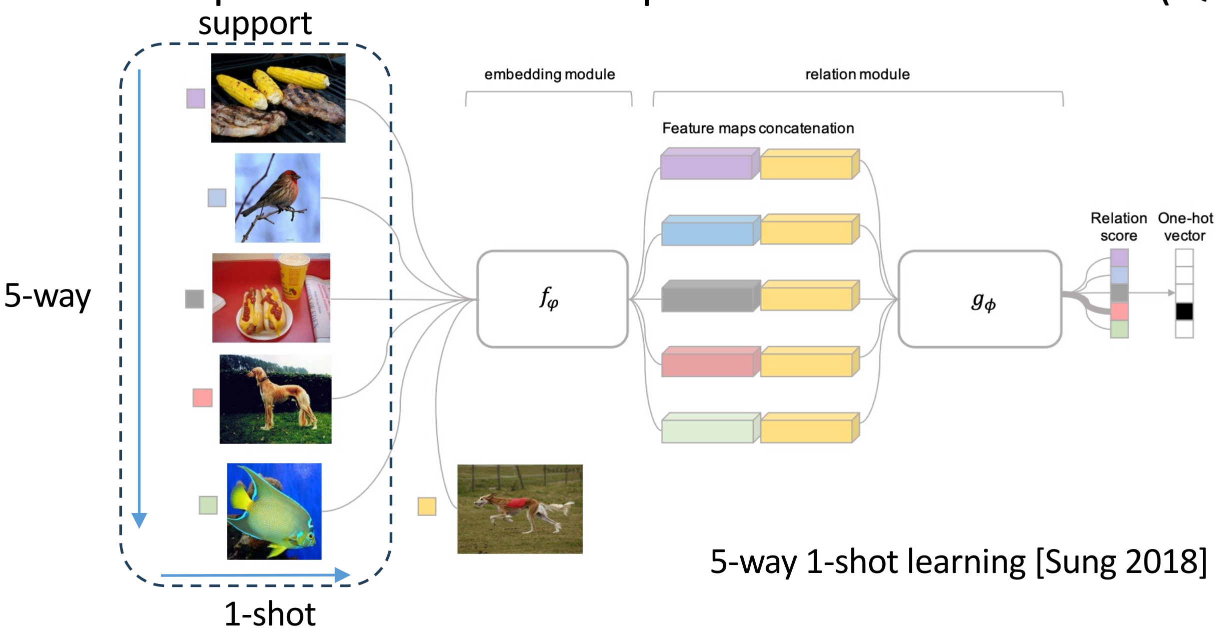
- The diffusion models were migrated to the quantum domain
  - Cacioppo 2023: Quantum diffusion models
  - Kölle 2024: Quantum denoising diffusion models
  - Zhang 2024: Generative quantum machine learning via denoising diffusion probabilistic models
  - Parigi 2024: Quantum-Noise-Driven Generative Diffusion Models
  - Kwun 2024: Mixed-State Quantum Denoising Diffusion Probabilistic Model
  - Chen 2024 Quantum generative diffusion model: a fully quantum-mechanical model for generating quantum state ensemble
- We use **QDDM** as a foundational tool for **few-shot learning (FSL)**





Quantum Few-Shot Learning

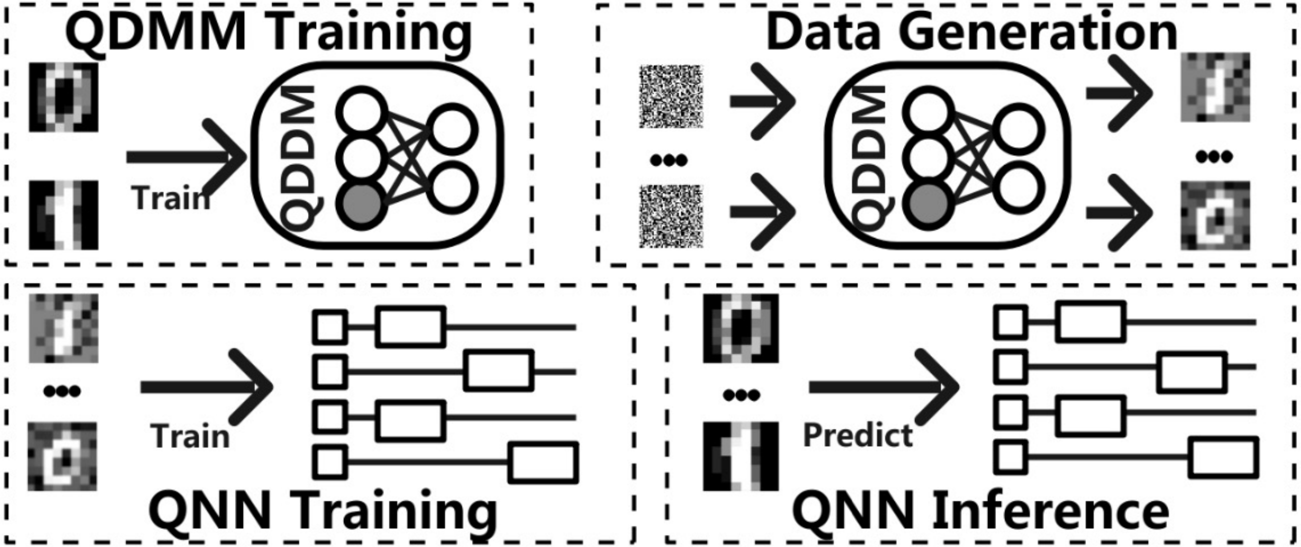
- Few-shot learning (FSL) is designed to address supervised learning challenges with a very limited number of training examples: **support set**.
  - The support set consists of a small number of labeled examples, encompassing ***n*** **classes, each with *k* examples**: called ***n-way, k-shot***.
  - We consider quantum FSL to train quantum neural networks (QNNs)



5-way 1-shot learning [Sung 2018]

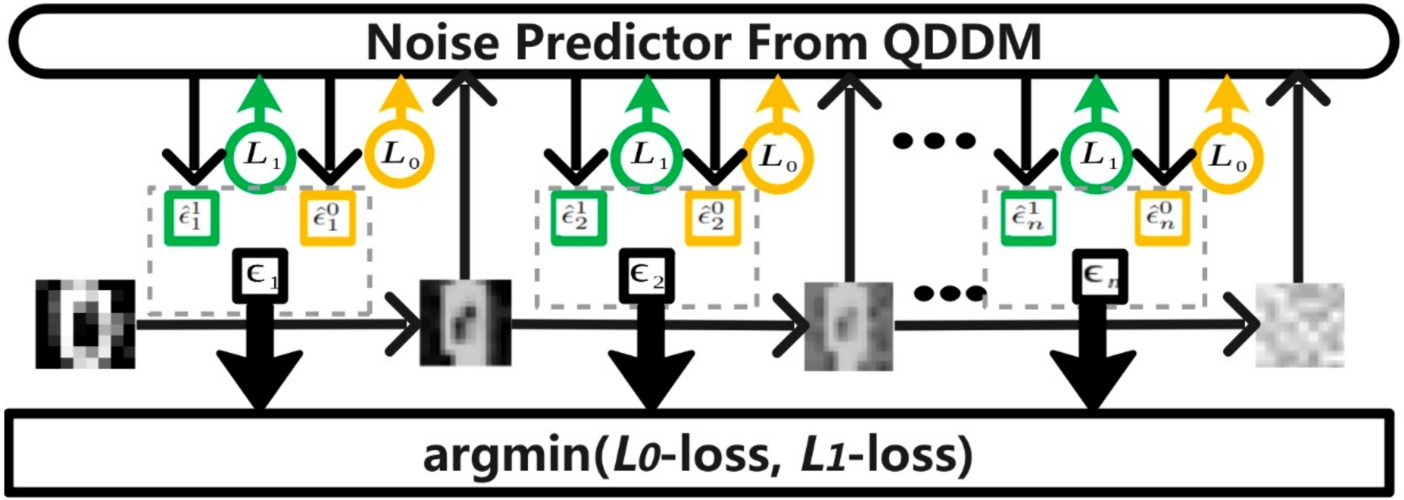
Method I: Generation Inference

- **QDDM-Based Label-Guided Generation Inference (LGGI)**
  - The primary challenge of QFSL is the limited availability of training data. Thus, expanding the training dataset can significantly enhance the performance of QFL.
  - A small amount of few-shot data is used to train the QDDM. Then, the QDDM is employed to expand the training dataset for QNN.
  - This expanded dataset is then used to train the QNN, which in turn improves its inference accuracy on real data.



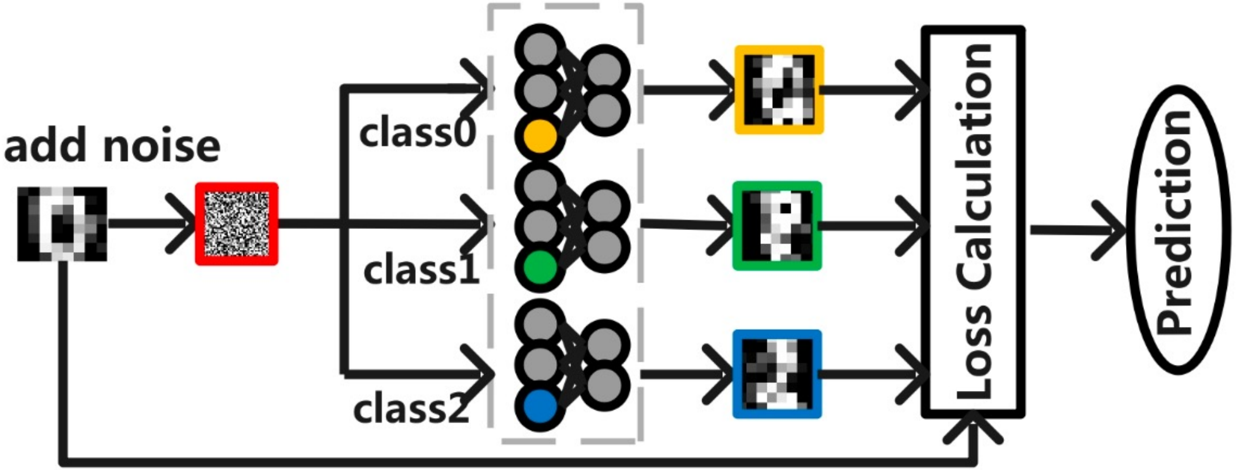
Method II: Diffusion Inference

- **QDM-Based Label-Guided Noise Addition Inference (LGNAI)**
  - The learning objective of the QDDM relies on using a noise predictor to estimate the noise in noisy data compared to the actual noise.
  - The noise predictor's estimation is guided by a label. By using the correct label for guidance, the error between the predicted noise and the actual noise may be minimized.



Method III: Denoising Inference

- **QDDM-Based Label-Guided Denoising Inference (LGDI)**
  – During the denoising phase of the QDDM, the noise predictor is utilized to estimate the noise present in the noisy data, which is subsequently subtracted.
  – The noise prediction is guided by labels. Consequently, the final generated images vary according to the guidance provided by different labels.
  – The data generated under the guidance of the true label may be close to the original data.



QDDM Training

- We use Adam for 10,000 epochs; labels are encoded with RX gate
- The training loss reflects the discrepancy between the noises predicted by the noise predictor and the actual noises during the denoising phase.

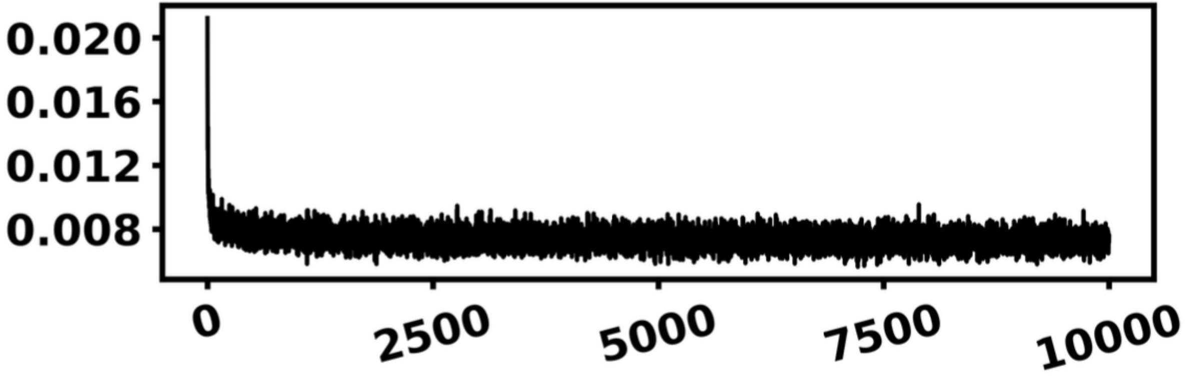


Figure 8: Training Loss Trends during QDDM Model Training.

Performance of Quantum Few-Shot Learning

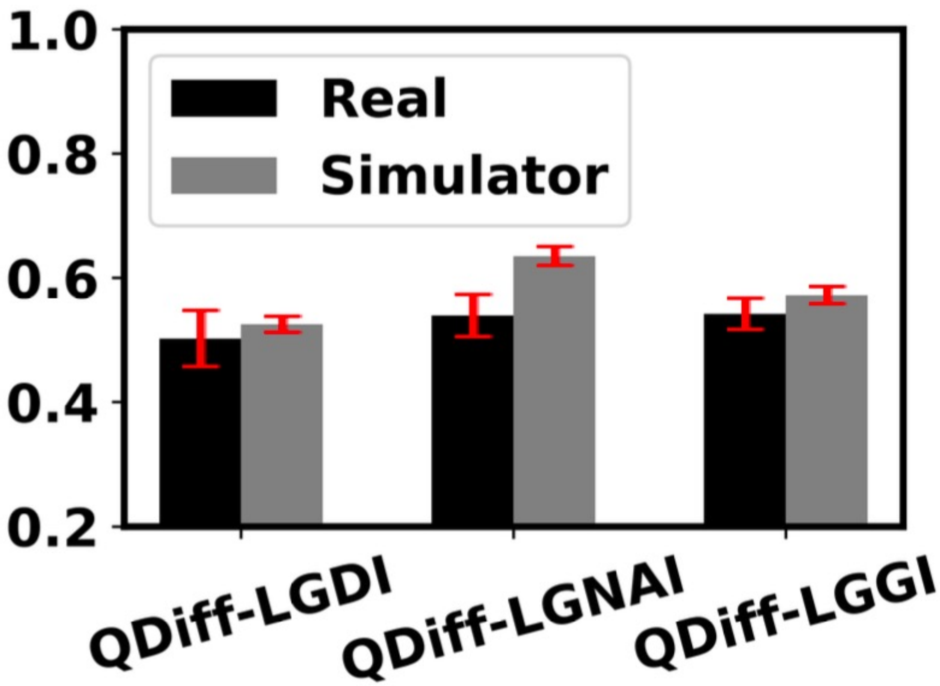
- QDDM-based FSL offers significant improvement in accuracy

Dataset	Tasks	QDiff-LGDI	QDiff-LGNAI	QDiff-LGGI	QMLP	C14	OPTIC	Quantumnat
Digits	2w-01s	0.975$_{\pm 0.059}$	0.978$_{\pm 0.003}$	0.992$_{\pm 0.009}$	0.764$_{\pm 0.108}$	0.505$_{\pm 0.175}$	0.525$_{\pm 0.133}$	0.751$_{\pm 0.147}$
	2w-10s	0.983$_{\pm 0.006}$	0.997$_{\pm 0.002}$	0.984$_{\pm 0.012}$	0.892$_{\pm 0.086}$	0.627$_{\pm 0.086}$	0.886$_{\pm 0.193}$	0.722$_{\pm 0.186}$
	3w-01s	0.525$_{\pm 0.001}$	0.635$_{\pm 0.007}$	0.573$_{\pm 0.069}$	0.338$_{\pm 0.087}$	0.447$_{\pm 0.193}$	0.475$_{\pm 0.021}$	0.555$_{\pm 0.013}$
	3w-10s	0.857$_{\pm 0.015}$	0.801$_{\pm 0.008}$	0.632$_{\pm 0.035}$	0.355$_{\pm 0.059}$	0.481$_{\pm 0.183}$	0.698$_{\pm 0.121}$	0.687$_{\pm 0.156}$
MNIST	2w-01s	0.943$_{\pm 0.002}$	0.965$_{\pm 0.003}$	0.805$_{\pm 0.093}$	0.675$_{\pm 0.067}$	0.567$_{\pm 0.064}$	0.845$_{\pm 0.149}$	0.701$_{\pm 0.162}$
	2w-10s	0.953$_{\pm 0.011}$	0.978$_{\pm 0.005}$	0.915$_{\pm 0.079}$	0.817$_{\pm 0.048}$	0.810$_{\pm 0.152}$	0.807$_{\pm 0.173}$	0.727$_{\pm 0.151}$
	3w-01s	0.475$_{\pm 0.003}$	0.505$_{\pm 0.007}$	0.428$_{\pm 0.035}$	0.325$_{\pm 0.027}$	0.503$_{\pm 0.122}$	0.477$_{\pm 0.159}$	0.501$_{\pm 0.012}$
	3w-10s	0.720$_{\pm 0.016}$	0.825$_{\pm 0.008}$	0.405$_{\pm 0.022}$	0.547$_{\pm 0.085}$	0.607$_{\pm 0.142}$	0.770$_{\pm 0.191}$	0.527$_{\pm 0.078}$
Fashion	2w-01s	0.738$_{\pm 0.007}$	0.768$_{\pm 0.007}$	0.898$_{\pm 0.036}$	0.688$_{\pm 0.064}$	0.581$_{\pm 0.187}$	0.765$_{\pm 0.149}$	0.583$_{\pm 0.181}$
	2w-10s	0.755$_{\pm 0.020}$	0.805$_{\pm 0.002}$	0.895$_{\pm 0.066}$	0.731$_{\pm 0.035}$	0.773$_{\pm 0.099}$	0.793$_{\pm 0.157}$	0.887$_{\pm 0.129}$
	3w-01s	0.453$_{\pm 0.008}$	0.433$_{\pm 0.001}$	0.483$_{\pm 0.012}$	0.331$_{\pm 0.098}$	0.332$_{\pm 0.172}$	0.473$_{\pm 0.128}$	0.622$_{\pm 0.063}$
	3w-10s	0.655$_{\pm 0.018}$	0.735$_{\pm 0.004}$	0.585$_{\pm 0.025}$	0.647$_{\pm 0.015}$	0.527$_{\pm 0.173}$	0.593$_{\pm 0.139}$	0.653$_{\pm 0.032}$
Average		0.754$_{\pm 0.015}$	0.795$_{\pm 0.004}$	0.719$_{\pm 0.045}$	0.574$_{\pm 0.060}$	0.546$_{\pm 0.140}$	0.678$_{\pm 0.150}$	0.666$_{\pm 0.120}$

QDDM-based FSL

Impact of Quantum Noise

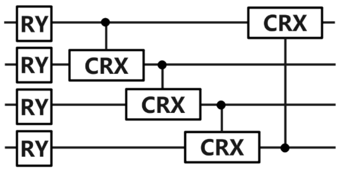
- We demonstrated robustness against quantum noise on IBM_Almaden quantum processors



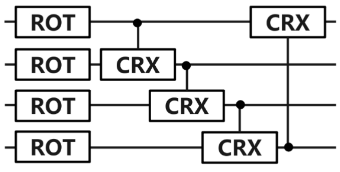
3-way, 1-shot

Impact of QNN Ansatz

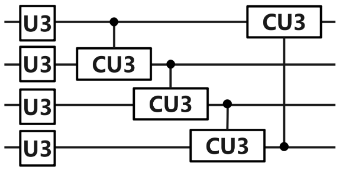
- The impact of the selection of QNNs utilized in Qdiff-LGGI
  - Different QNNs have different expressive abilities and different information extraction capabilities for input images.



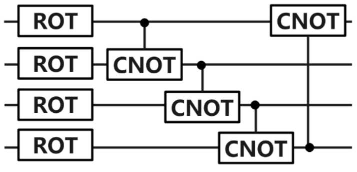
(a)C14



(b)QMLP

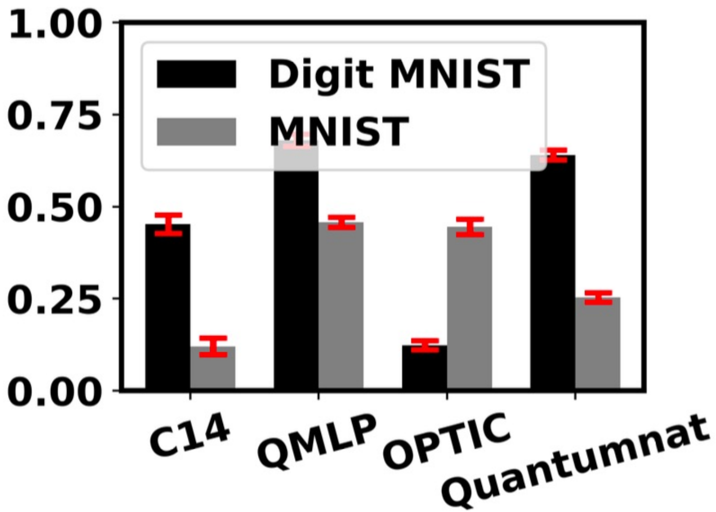


(c) Quantumnat



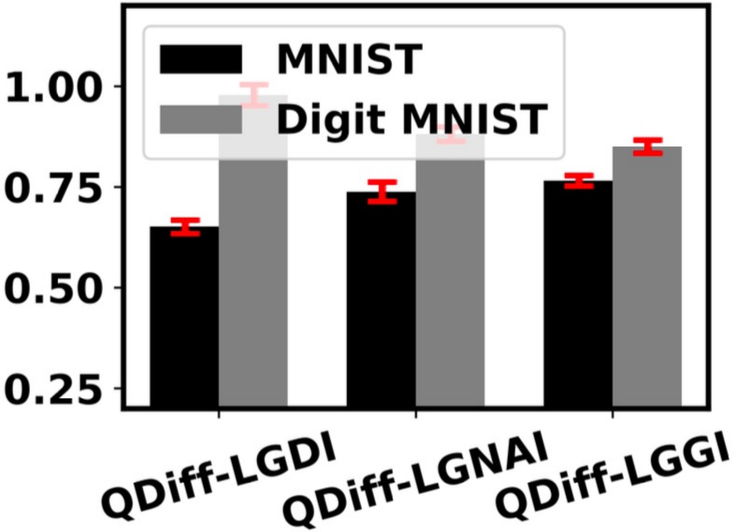
(d)OPTIC

QNN	# Qubits	1QG	2QG	# Param.
QMLP	6	ROT	CRX	$24 \times n$
C14	6	RY	CRX	$12 \times n$
OPTIC	6	ROT	CNOT	$18 \times n$
Quantumnat	6	U3	CU3	$36 \times n$

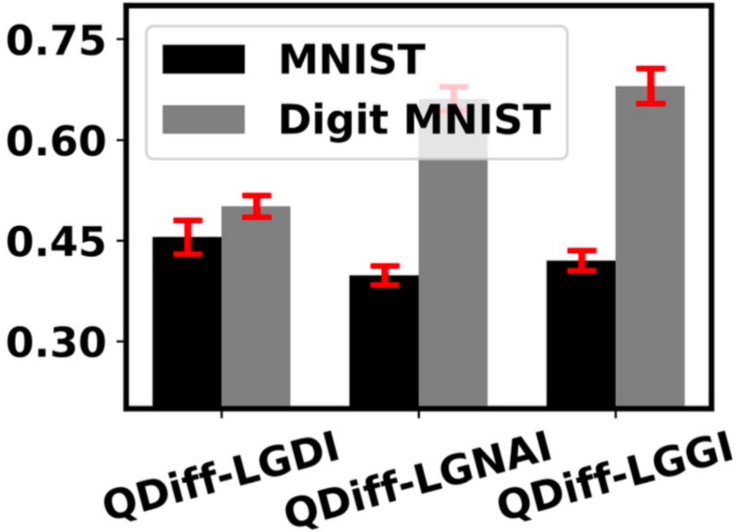


Performance of Quantum Zero-Short Learning

- We utilize the Digit MNIST dataset to train the QDDM and then use Qdiff-based algorithms to complete the inference on the MNIST dataset. The same strategy is employed for experiments on the MNIST dataset.



2-way



3-way

Summary

- We addressed overviews of generative AI and quantum AI
- We proposed new methods for quantum few-shot learning using quantum diffusion models
  - We introduced 3 different approaches based on QDDM’s generation/diffusion/denoising
  - Our methods demonstrated significant improvement up to 30% gain
  - We also validated that our method has a high resilience to the quantum noise
  - We evaluated different QNN ansatz
  - Zero-shot capability was discussed too
- Future work:
  - Enhancing the capabilities of QDM through improvements in model architecture and optimization techniques, enabling more intricate datasets with diverse and high-dimensional features for diverse real-world applications.
- Questions?
  - koike@merl.com