// Maria Schuld

XANADU

How to rethink quantum machine learning

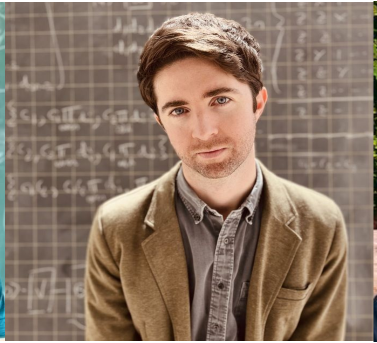
IPAM, October 2023

// Credits

Xanadu's quantum machine learning team

Dr Maria Schuld
QML Research lead

Dr Richard East
Researcher



Dr Joseph Bowles
Researcher

Dr David Wakeham
Researcher

Dr Shahnawaz Ahmed
Researcher

Dr Nathan Killoran
CTO Software

Dr Chae-Yeun Park
Researcher

Dr David Wierichs
Researcher

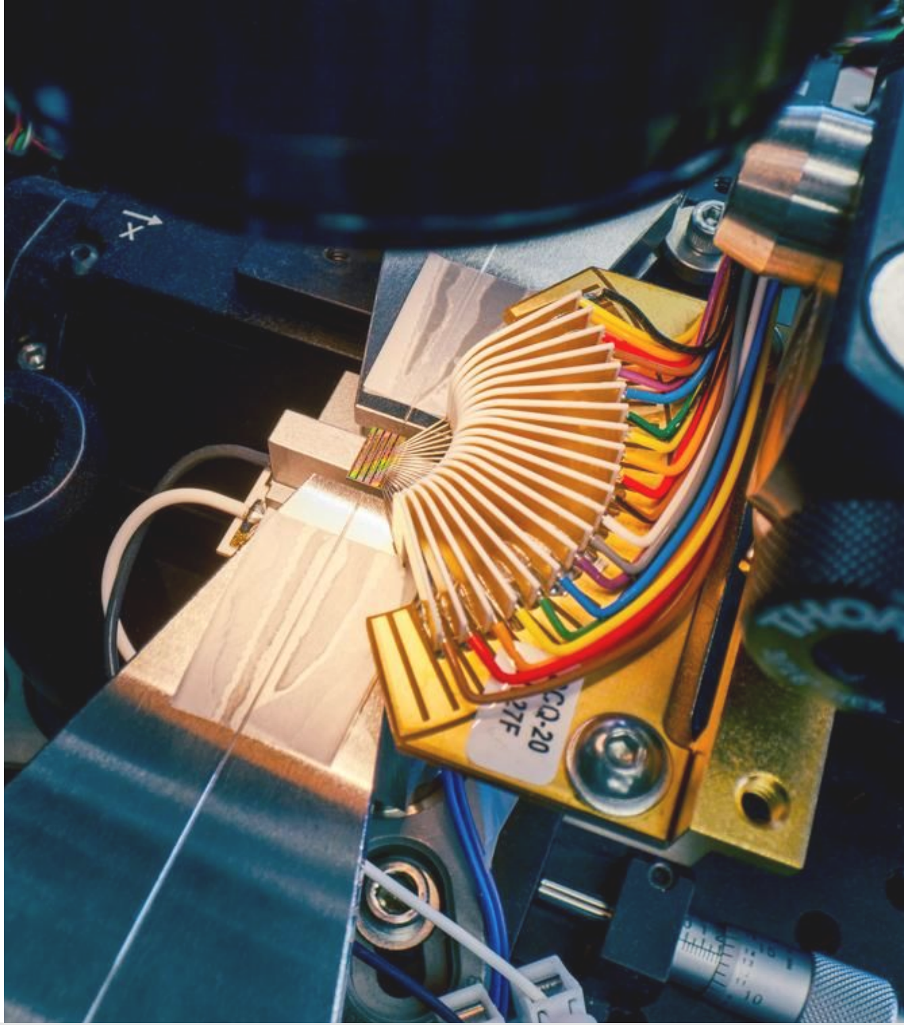
Dr Korbinian Kottmann
Researcher

We're hiring!

// Xanadu's mission

To build quantum computers that are useful and available to people everywhere

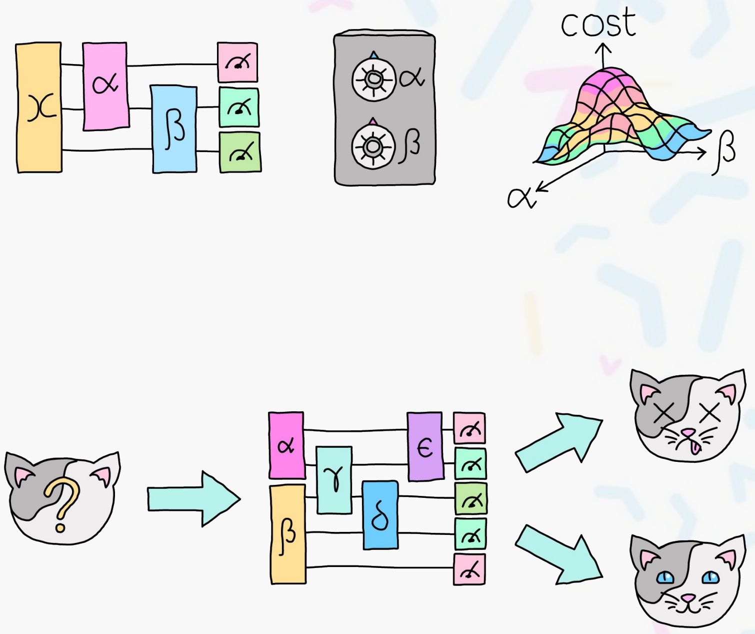
Founded	Headquarters	People
2016	**Toronto**	**170+**



XANADU

// The QML team objective

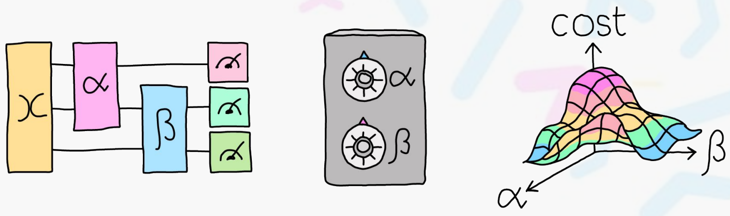
Make quantum computers useful for machine learning

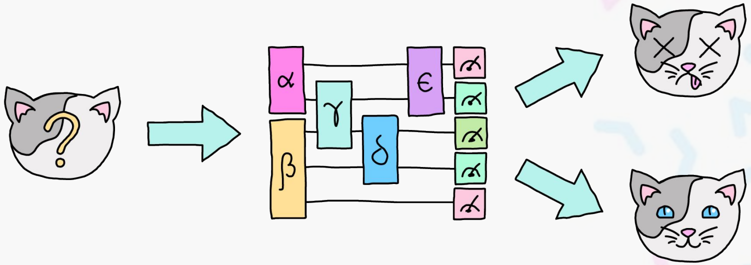


// The QML team objective

Make quantum computers useful for machine learning

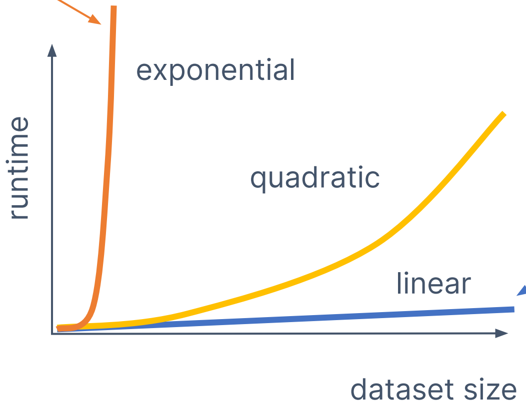
For this to happen we need to change some things in our approach to research





“We prove an exponential speedup for QML…”

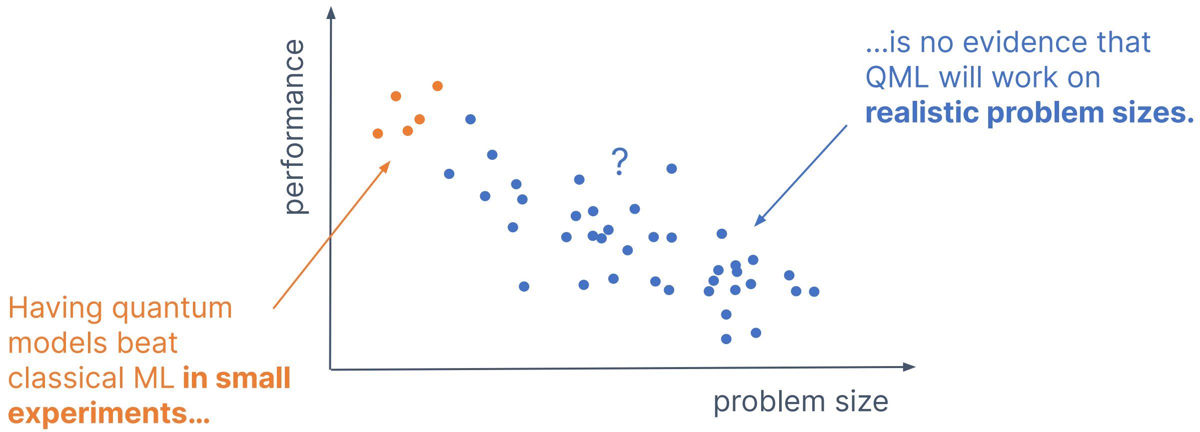
The problem that neural networks solve is **exponentially hard in theory…**



…but in practice neural nets run in **linear time.**

Our performance measures are not meaningful for (mainstream) ML.

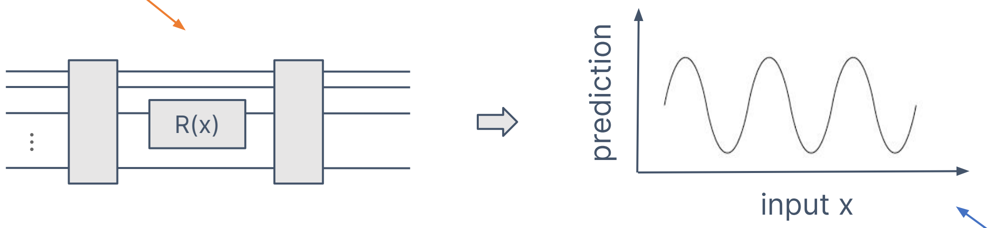
"Our quantum model does better on MNIST..."



Our experiments do not probe the right regimes yet.

"We use an ansatz of Pauli gates and entanglers…"

An impressive quantum circuit
(of 10,000 qubits and 1 billion parameters with universal, classically intractable unitary evolutions imitating a deep learning technique)…



…can be a **useless ML model.**

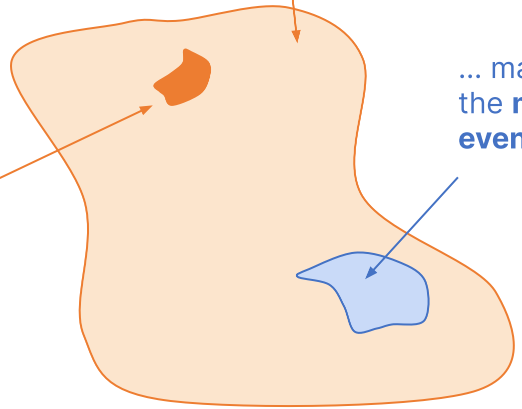
Our circuit designs should be motivated better.

// Checking the compass: Model design

“Quantum models generalise/train better/worse…”

...or about a large class of quantum models...

Theorems about a very specific class of models...



... may not be true for the **models we eventually care about.**

We don’t know if our theory targets relevant questions.

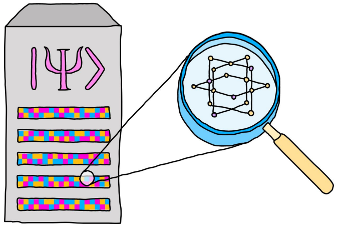
How to rethink Quantum Machine Learning

Assess how good quantum models *really* are

Systematically benchmark popular ideas in QML

Design models from first principles

Find a different approach to quantum model design



Assess how good quantum models *really* are

// Assess how good quantum models really are

What is the best benchmark design we can come up with?

Model selection

- Arxiv papers >2018 with keywords “classif”, “learn”, “supervised”, “MNIST” [3500 papers]
- >=30 Google Scholar citations [561 papers]
- New NISQ quantum model for classification on conventional classical data [29 papers]
- In random subset of 15 papers
- Not vetoed at closer inspection [12 papers]

QNN

Lloyd et al. "Quantum embeddings for machine learning." 2001.03622

[Zhao et al. "Building quantum neural networks based on a swap test." 1904.12697v2]

[Zhang et al. "Toward trainability of quantum neural networks." 2011.06258v2]

Schuld et al. "Circuit-centric quantum classifiers." 1804.00633v1

Pérez-Salinas et al. "Data re-uploading for a universal quantum classifier." 1907.02085v3

Mari et al. "Transfer learning in hybrid classical-quantum neural networks." 1912.08278v2

Havlíček et al. "Supervised learning with quantum-enhanced feature spaces." 1804.11326v2

QKernel

Havlíček et al. "Supervised learning with quantum-enhanced feature spaces." 1804.11326v2

Wilson et al. "Quantum kitchen sinks: An algorithm for ML on near-term…." 1806.08321v2

Huang et al. “Power of data in quantum machine learning.” 2011.01938v2

QConv

Henderson et al. "Quanvolutional neural networks: powering image recognition.." 1904.04767v1

Wei "A quantum convolutional neural network on NISQ devices." 2104.06918v3

QGenerative

[Zoufal et al.. "Variational quantum Boltzmann machines." 2006.06004v1]

// Assess how good quantum models really are

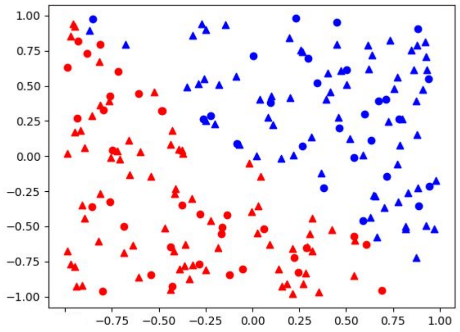
What is the best benchmark design we can come up with?

Tasks

- Binary classification
- Figure of merit: accuracy
- 4 datasets of variable dimension:

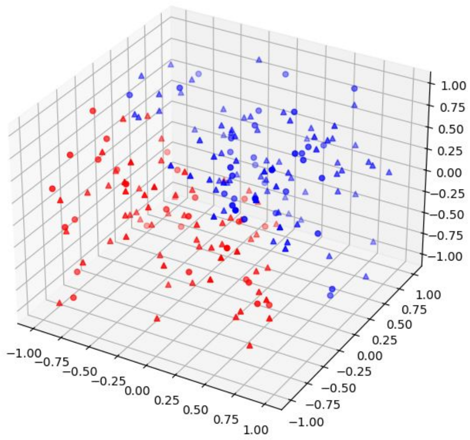
// Assess how good quantum models really are

What is the best benchmark design we can come up with?



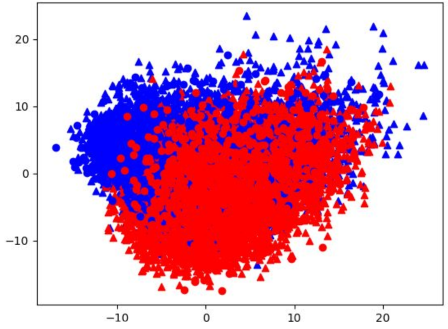
Tasks

- Binary classification
- Figure of merit: accuracy
- 4 datasets of variable dimension:
  - **SIMPLE:** Linearly separated points in hypercube



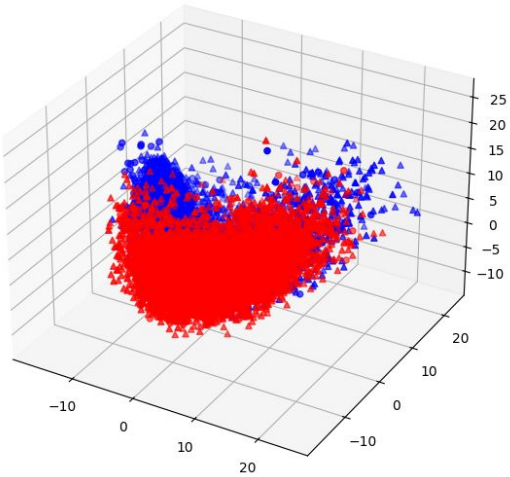
// Assess how good quantum models really are

What is the best benchmark design we can come up with?



Tasks

- Binary classification
- Figure of merit: accuracy
- 4 datasets of variable dimension:
    - **SIMPLE:** Linearly separated points in hypercube
    - **WIDELY USED:** pre-processed MNIST

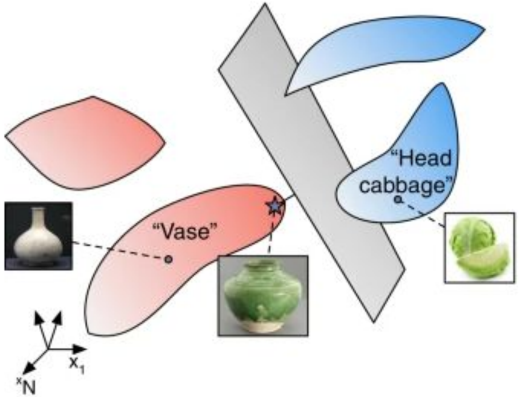


// Assess how good quantum models really are

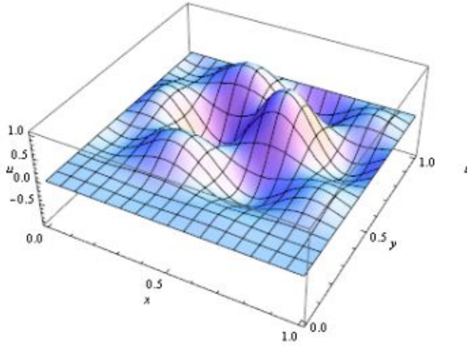
What is the best benchmark design we can come up with?

Tasks

- Binary classification
- Figure of merit: accuracy
- 4 datasets of variable dimension:
    - **SIMPLE:** Linearly separated points in hypercube
    - **WIDELY USED:** pre-processed MNIST
    - **[REALISTIC:** Low-dimensional manifolds (Goldt 2019, Buchanan 2020)]
    - **[TAILORMADE:** Multi-dimensional Fourier series]

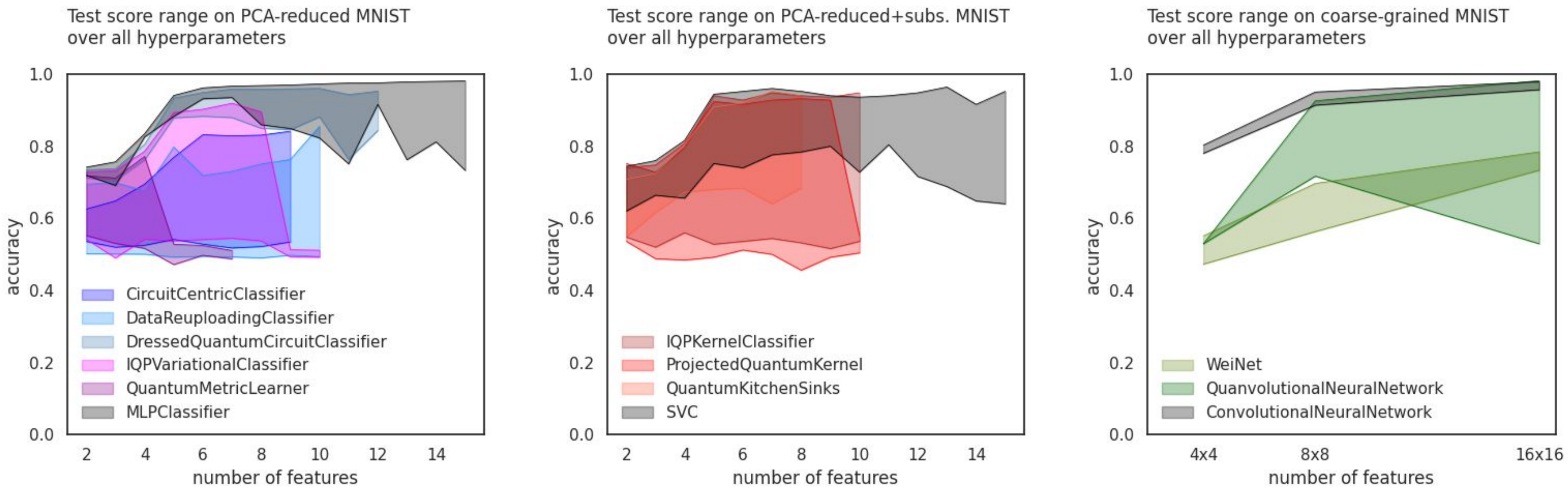


https://www.nature.com/articles/s41467-020-14578-5



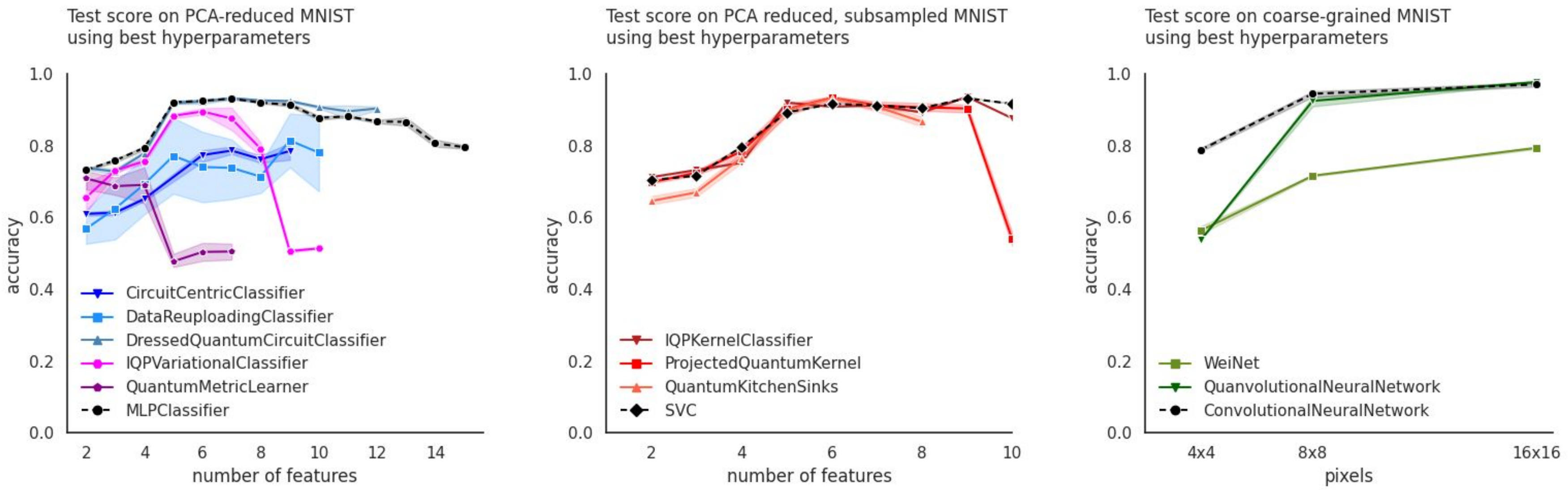
// Preliminary results

Hyperparameters matter



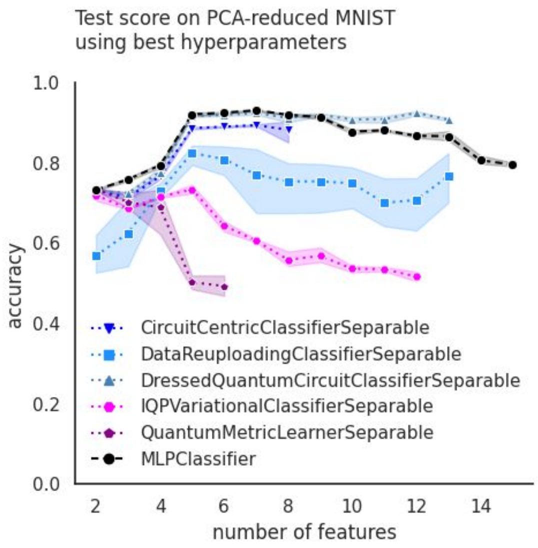
// Preliminary results

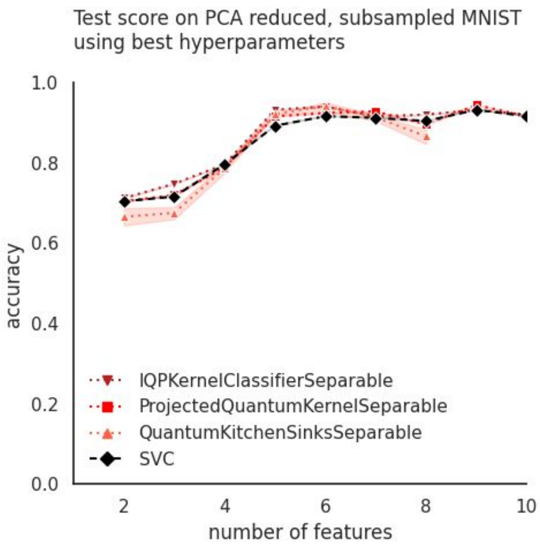
Out-of-the box classical models are not easily beaten

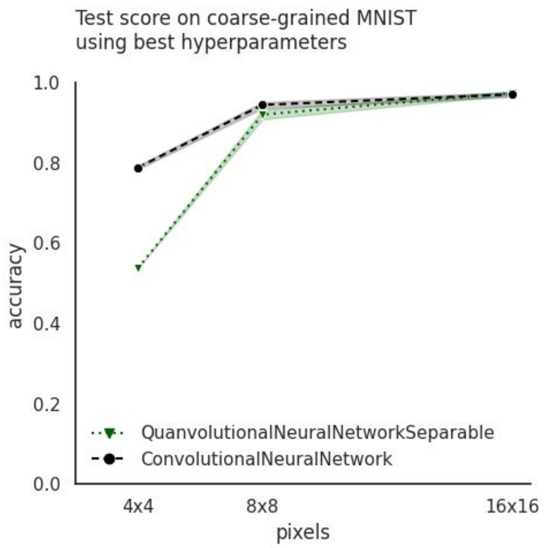


// Preliminary results

Separable circuits perform the same

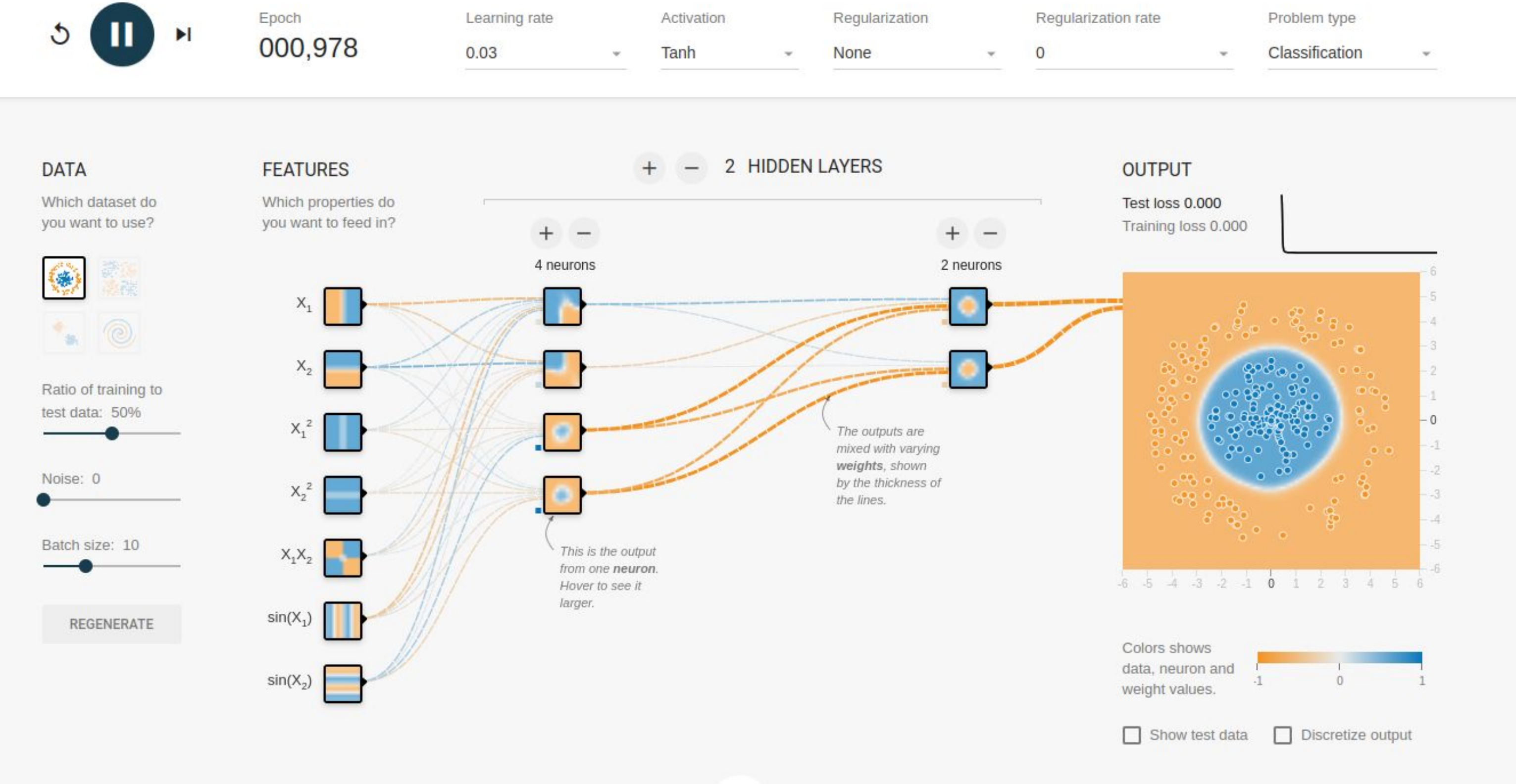






// Preliminary results

Are we just building trigonometric/polynomial features?



https://playground.tensorflow.org/

Design models from first principles

How we implicitly think about QML

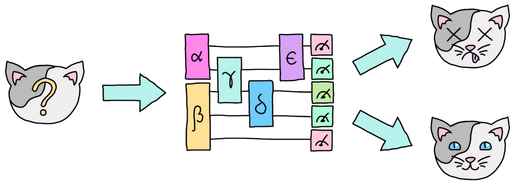
Quantum computing

provable quantum advantage,
NISQ circuits

Machine learning

deep learning
(i.e., optimising large models with gradient descent)

Quantum Neural Nets



- They inherit speedups
- They use Pauli-gates
- They follow the blueprint of deep learning

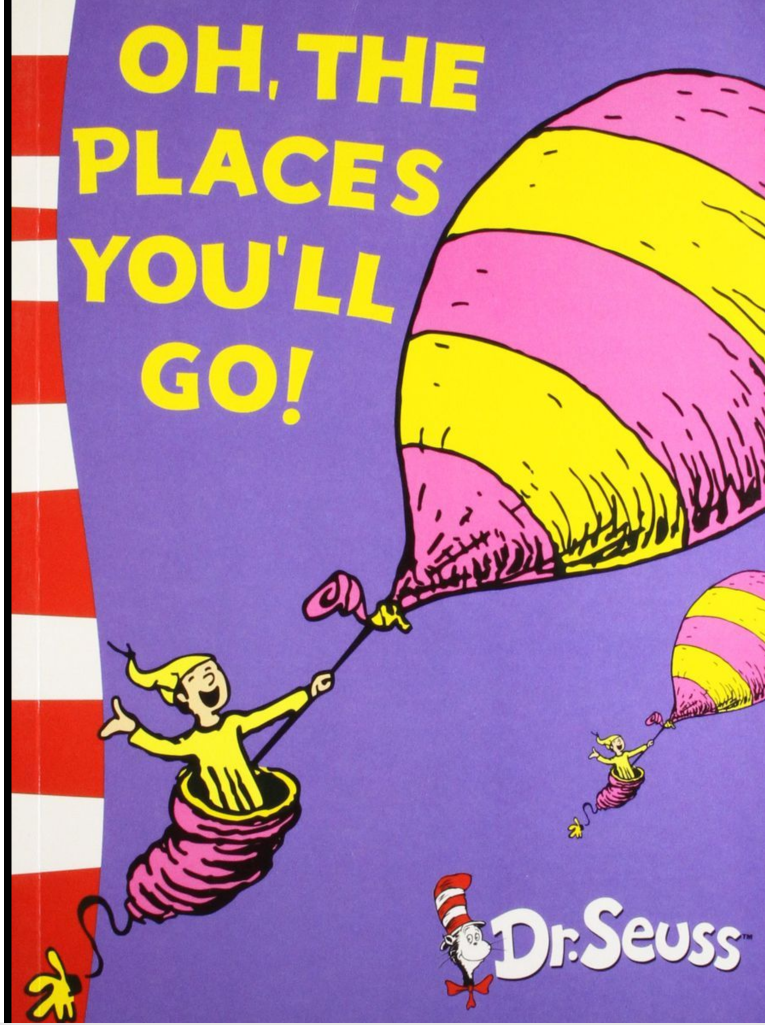
A different starting point...

Quantum computing

solving highly structured problems with interference

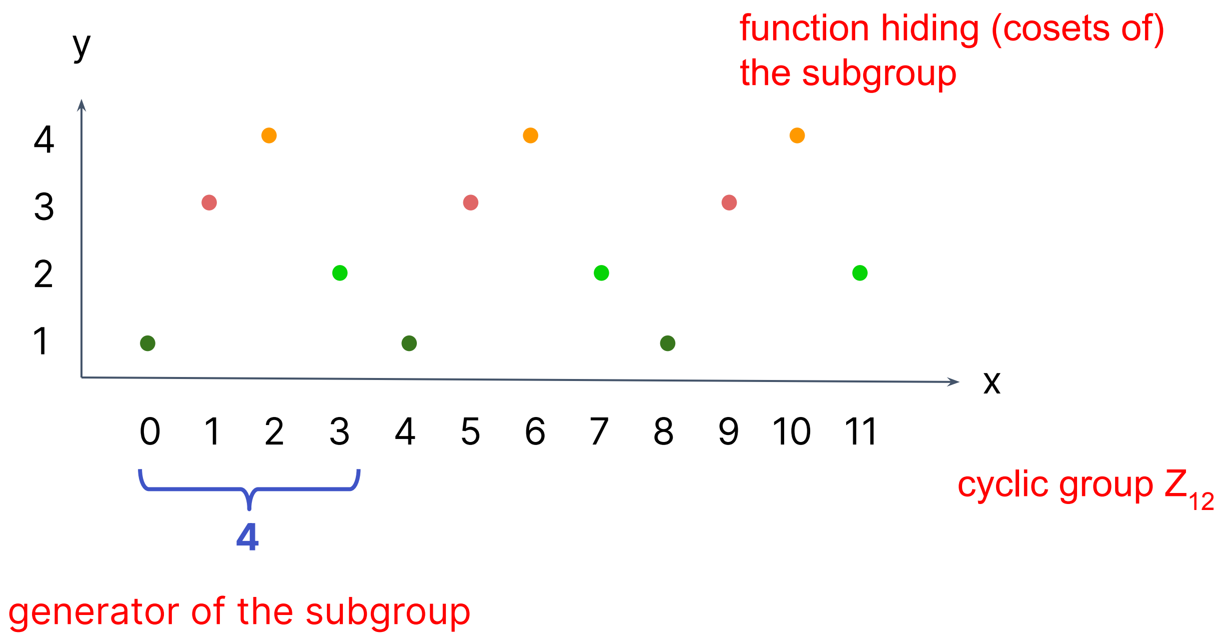
Machine learning

generalising from samples



// Our playground

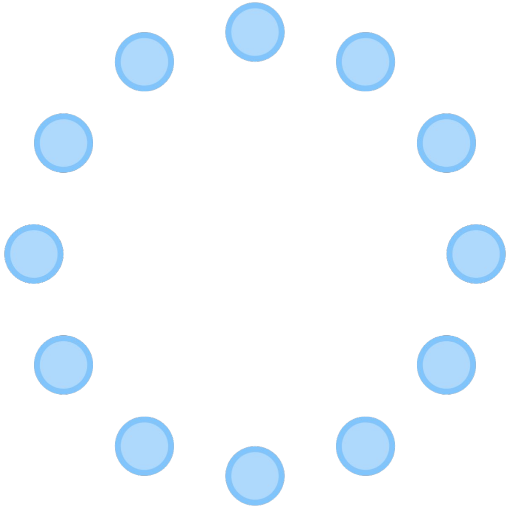
Period finding: A hidden subgroup problem



// Our playground

Algorithm for the hidden subgroup problem: Initial state

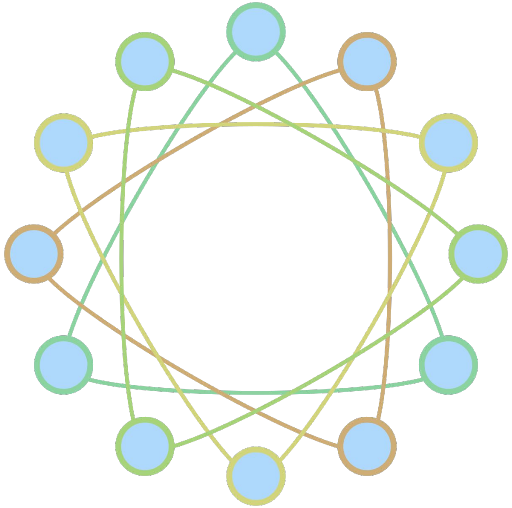
$$
\begin{aligned}
&|0\rangle \\
+&|1\rangle \\
+&|2\rangle \\
+&|3\rangle \\
+&|4\rangle \\
+&|5\rangle \\
+&|6\rangle \\
+&|7\rangle \\
+&|8\rangle \\
+&|9\rangle \\
+&|10\rangle \\
+&|11\rangle
\end{aligned}
$$



// Our playground

Algorithm for the hidden subgroup problem: Oracle

$|0\rangle\ |1\rangle$
$+|1\rangle\ |3\rangle$
$+|2\rangle\ |4\rangle$
$+|3\rangle\ |2\rangle$
$+|4\rangle\ |1\rangle$
$+|5\rangle\ |3\rangle$
$+|6\rangle\ |4\rangle$
$+|7\rangle\ |2\rangle$
$+|8\rangle\ |1\rangle$
$+|9\rangle\ |3\rangle$
$+|10\rangle\ |4\rangle$
$+|11\rangle\ |2\rangle$



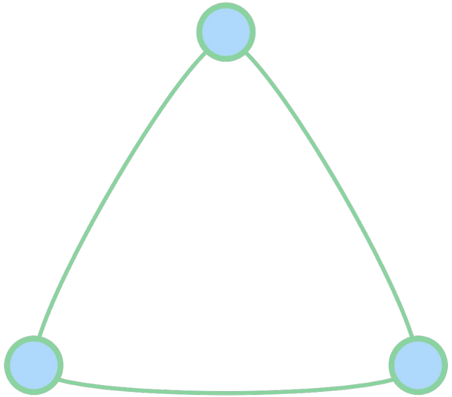
// Our playground

Algorithm for the hidden subgroup problem: Measurement

$|0\rangle$

$+|4\rangle$

$+|8\rangle$



// Our playground

Algorithm for the hidden subgroup problem: QFT

$$|0\rangle + |4\rangle + |8\rangle \Rightarrow$$

$$
\begin{aligned}
&(e^{2\pi i\, 0*0/12} + e^{2\pi i\, 4*0/12} + e^{2\pi i\, 8*0/12})\,|0\rangle \\
+\;&(e^{2\pi i\, 0*1/12} + e^{2\pi i\, 4*1/12} + e^{2\pi i\, 8*1/12})\,|1\rangle \\
+\;&(e^{2\pi i\, 0*2/12} + e^{2\pi i\, 4*2/12} + e^{2\pi i\, 8*2/12})\,|2\rangle \\
+\;&(e^{2\pi i\, 0*3/12} + e^{2\pi i\, 4*3/12} + e^{2\pi i\, 8*3/12})\,|3\rangle \\
+\;&(e^{2\pi i\, 0*4/12} + e^{2\pi i\, 4*4/12} + e^{2\pi i\, 8*4/12})\,|4\rangle \\
+\;&(e^{2\pi i\, 0*5/12} + e^{2\pi i\, 4*5/12} + e^{2\pi i\, 8*5/12})\,|5\rangle \\
+\;&(e^{2\pi i\, 0*6/12} + e^{2\pi i\, 4*6/12} + e^{2\pi i\, 8*6/12})\,|6\rangle \\
+\;&(e^{2\pi i\, 0*7/12} + e^{2\pi i\, 4*7/12} + e^{2\pi i\, 8*7/12})\,|7\rangle \\
+\;&(e^{2\pi i\, 0*8/12} + e^{2\pi i\, 4*8/12} + e^{2\pi i\, 8*8/12})\,|8\rangle \\
+\;&(e^{2\pi i\, 0*9/12} + e^{2\pi i\, 4*9/12} + e^{2\pi i\, 8*9/12})\,|9\rangle \\
+\;&(e^{2\pi i\, 0*10/12} + e^{2\pi i\, 4*10/12} + e^{2\pi i\, 8*10/12})\,|10\rangle \\
+\;&(e^{2\pi i\, 0*11/12} + e^{2\pi i\, 4*11/12} + e^{2\pi i\, 8*11/12})\,|11\rangle
\end{aligned}
$$

// Our playground

Algorithm for the hidden subgroup problem: QFT

$$|0\rangle + |4\rangle + |8\rangle$$

➡

$$
\begin{aligned}
& (e^{2\pi i\, 0} + e^{2\pi i\, 0} + e^{2\pi i\, 0})\,|0\rangle \\
+ & (e^{2\pi i\, 0/3} + e^{2\pi i\, 1/3} + e^{2\pi i\, 2/3})\,|1\rangle \\
+ & (e^{2\pi i\, 0/3} + e^{2\pi i\, 2/3} + e^{2\pi i\, 1'1/3})\,|2\rangle \\
+ & (e^{2\pi i\, 0} + e^{2\pi i\, 1} + e^{2\pi i\, 2})\,|3\rangle \\
+ & (e^{2\pi i\, 0/3} + e^{2\pi i\, 1'1/3} + e^{2\pi i\, 2'2/3})\,|4\rangle \\
+ & (e^{2\pi i\, 0/3} + e^{2\pi i\, 1'2/3} + e^{2\pi i\, 3'1/3})\,|5\rangle \\
+ & (e^{2\pi i\, 0} + e^{2\pi i\, 2} + e^{2\pi i\, 4})\,|6\rangle \\
+ & (e^{2\pi i\, 0/3} + e^{2\pi i\, 2'1/3} + e^{2\pi i\, 4'2/3})\,|7\rangle \\
+ & (e^{2\pi i\, 0/3} + e^{2\pi i\, 2'2/3} + e^{2\pi i\, 5'1/3})\,|8\rangle \\
+ & (e^{2\pi i\, 0} + e^{2\pi i\, 3} + e^{2\pi i\, 9})\,|9\rangle \\
+ & (e^{2\pi i\, 0/3} + e^{2\pi i\, 3'1/3} + e^{2\pi i\, 6'2/3})\,|10\rangle \\
+ & (e^{2\pi i\, 0/3} + e^{2\pi i\, 3'2/3} + e^{2\pi i\, 7'1/3})\,|11\rangle
\end{aligned}
$$

// Our playground

Algorithm for the hidden subgroup problem: QFT

$|0\rangle$

$+|4\rangle$

$+|8\rangle$

$\Rightarrow$

$|0\rangle$

$+ \ |3\rangle$

$+ \ |6\rangle$

$+ \ |9\rangle$

Algorithm for the hidden subgroup problem: Decoding

Any state
one might measure
has the form

|k*12/4⟩

known

wanted

can be inferred from several samples

|0⟩

+ |3⟩

+ |6⟩

+ |9⟩

// A different starting point…

Analysing traditional quantum algorithms from an ML perspective

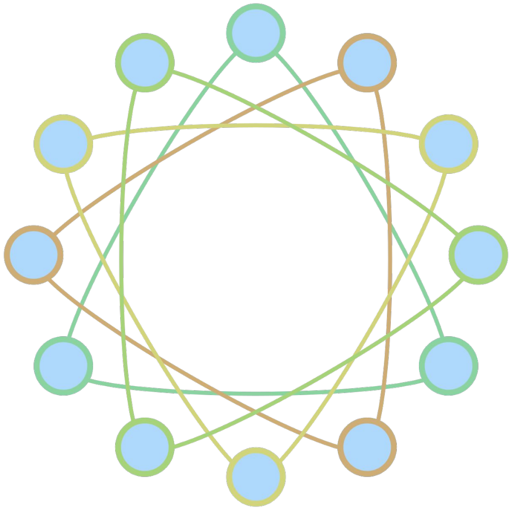
Quantum computing

solving highly structured problems with interference

Machine learning

generalising from samples

$|0\rangle\ |1\rangle$
$+|1\rangle\ |3\rangle$
$+|2\rangle\ |4\rangle$
$+|3\rangle\ |2\rangle$
$+|4\rangle\ |1\rangle$
$+|5\rangle\ |3\rangle$
$+|6\rangle\ |4\rangle$
$+|7\rangle\ |2\rangle$
$+|8\rangle\ |1\rangle$
$+|9\rangle\ |3\rangle$
$+|10\rangle\ |4\rangle$
$+|11\rangle\ |2\rangle$



// A different starting point…

Analysing traditional quantum algorithms from an ML perspective

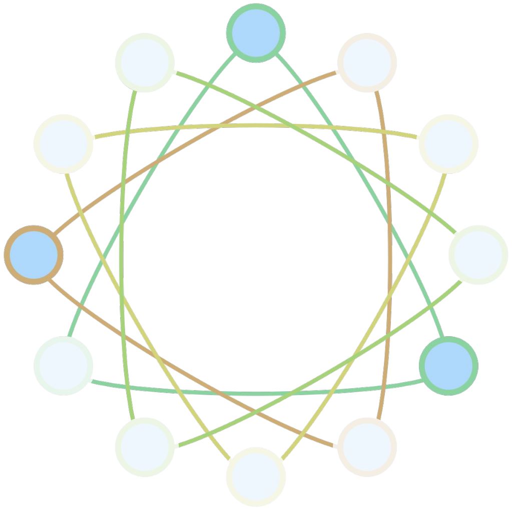
Quantum computing

solving highly structured problems with interference

Machine learning

generalising from samples

$|0\rangle|1\rangle$
$+|1\rangle|3\rangle$
$+|2\rangle|4\rangle$
$+|3\rangle|2\rangle$
$+|4\rangle|1\rangle$
$+|5\rangle|3\rangle$
$+|6\rangle|4\rangle$
$+|7\rangle|2\rangle$
$+|8\rangle|1\rangle$
$+|9\rangle|3\rangle$
$+|10\rangle|4\rangle$
$+|11\rangle|2\rangle$



// A different starting point…

Analysing traditional quantum algorithms from an ML perspective

Quantum computing

solving highly structured problems with interference

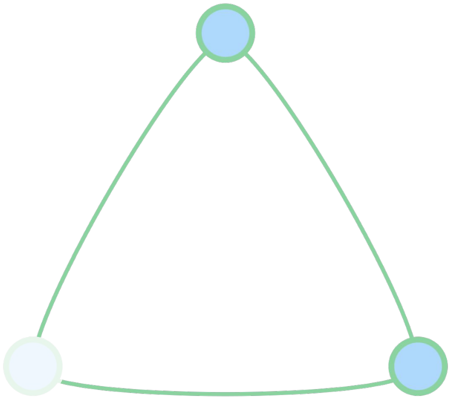
Machine learning

generalising from samples

$|0\rangle \; |1\rangle$

$+|4\rangle \; |1\rangle$

$+|8\rangle \; |1\rangle$



// A different starting point…

Analysing traditional quantum algorithms from an ML perspective

Quantum computing

solving highly structured problems with interference

Machine learning

generalising from samples

$$
\begin{array}{lllllll}
 & (e^{2\pi i\ 0} & + & e^{2\pi i\ 0} & + & e^{2\pi i\ 0} &)\ |0\rangle \\
+ & (e^{2\pi i\ 0/3} & + & e^{2\pi i\ 1/3} & + & e^{2\pi i\ 2/3} &)\ |1\rangle \\
+ & (e^{2\pi i\ 0/3} & + & e^{2\pi i\ 2/3} & + & e^{2\pi i\ 1'1/3} &)\ |2\rangle \\
+ & (e^{2\pi i\ 0} & + & e^{2\pi i\ 1} & + & e^{2\pi i\ 2} &)\ |3\rangle \\
+ & (e^{2\pi i\ 0/3} & + & e^{2\pi i\ 1'1/3} & + & e^{2\pi i\ 2'2/3} &)\ |4\rangle \\
+ & (e^{2\pi i\ 0/3} & + & e^{2\pi i\ 1'2/3} & + & e^{2\pi i\ 3'1/3} &)\ |5\rangle \\
+ & (e^{2\pi i\ 0} & + & e^{2\pi i\ 2} & + & e^{2\pi i\ 4} &)\ |6\rangle \\
+ & (e^{2\pi i\ 0/3} & + & e^{2\pi i\ 2'1/3} & + & e^{2\pi i\ 4'2/3} &)\ |7\rangle \\
+ & (e^{2\pi i\ 0/3} & + & e^{2\pi i\ 2'2/3} & + & e^{2\pi i\ 5'1/3} &)\ |8\rangle \\
+ & (e^{2\pi i\ 0} & + & e^{2\pi i\ 3} & + & e^{2\pi i\ 9} &)\ |9\rangle \\
+ & (e^{2\pi i\ 0/3} & + & e^{2\pi i\ 3'1/3} & + & e^{2\pi i\ 6'2/3} &)\ |10\rangle \\
+ & (e^{2\pi i\ 0/3} & + & e^{2\pi i\ 3'2/3} & + & e^{2\pi i\ 7'1/3} &)\ |11\rangle
\end{array}
$$

// A different starting point…

Analysing traditional quantum algorithms from an ML perspective

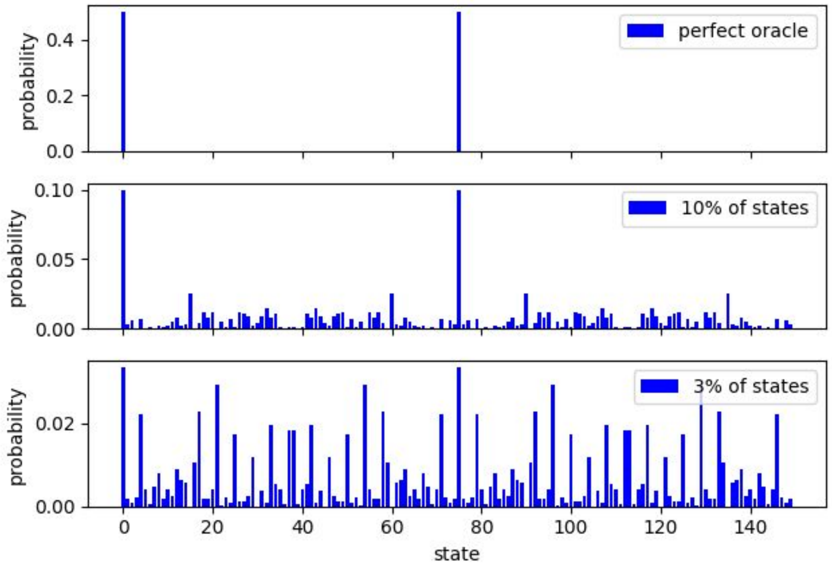
Quantum computing

solving highly structured problems with interference

Machine learning

generalising from samples

Z150, subgroup: {0,2,..., 146, 148}



// Research question

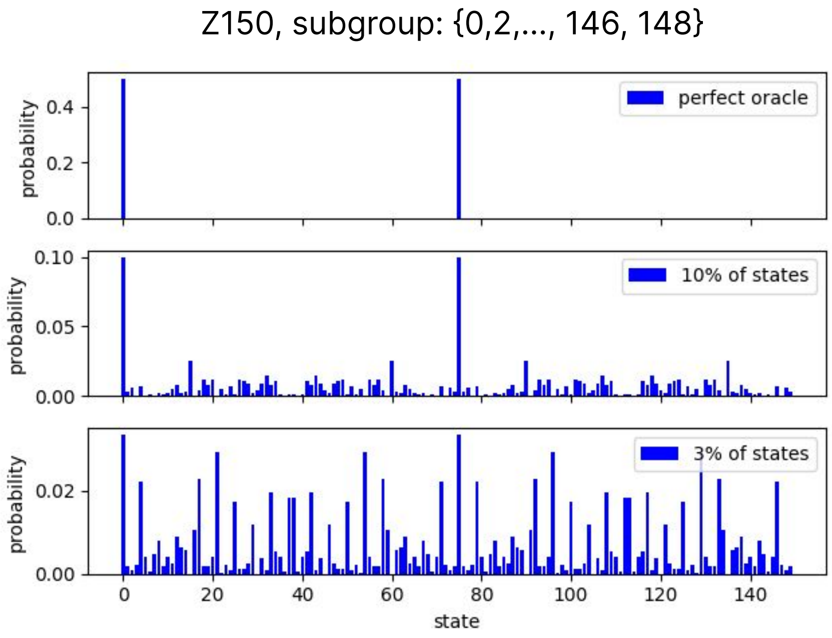
Can we amplify the signal in Fourier sampling?

Quantum computing

solving highly structured problems with interference

Machine learning

generalising from samples



// Research question

Can we learn to reconstruct the oracle from data?

Quantum computing

solving highly structured problems with interference

Machine learning

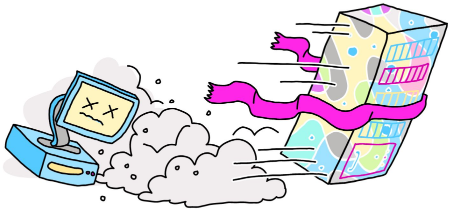
generalising from samples

$+\|0\rangle\ \|1\rangle$	$+\|0\rangle\ \|1\rangle$
$+\|1\rangle\ \|3\rangle$	$+\|1\rangle\ \|3\rangle$
$+\|2\rangle\ \|4\rangle$	$+\|2\rangle\ \|4\rangle$
$+\|3\rangle\ \|2\rangle$	$+\|3\rangle\ \|2\rangle$
$+\|4\rangle\ \|1\rangle$	$+\|4\rangle\ \|1\rangle$
$+\|5\rangle\ \|3\rangle$	$+\|5\rangle\ \|3\rangle$
$+\|6\rangle\ \|4\rangle$	$+\|6\rangle\ \|4\rangle$
$+\|7\rangle\ \|2\rangle$	$+\|7\rangle\ \|2\rangle$
$+\|8\rangle\ \|1\rangle$	$+\|8\rangle\ \|1\rangle$
$+\|9\rangle\ \|3\rangle$	$+\|9\rangle\ \|3\rangle$
$+\|10\rangle\ \|4\rangle$	$+\|10\rangle\ \|4\rangle$
$+\|11\rangle\ \|2\rangle$	$+\|11\rangle\ \|2\rangle$

How to rethink Quantum Machine Learning

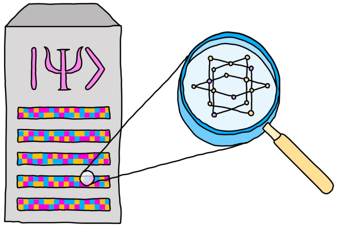
Assess how good quantum models *really* are

Put quantum model design ideas to the test



Design models from first principles

Analyse traditional quantum algorithms from an ML perspective



Thank you

XANADU

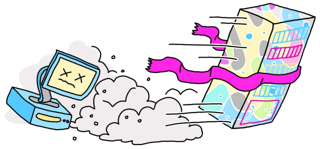
xanadu.ai
Twitter → @xanadu.ai
maria@xanadu.ai

// Assess how good quantum models really are

What is the best benchmark design we can come up with?

Tasks

- Binary classification
- Figure of merit: accuracy
- 4 datasets of variable dimension:
  - **SIMPLE:** Linearly separated points in hypercube
  - **REALISTIC:** pre-processed MNIST
  - [**ADVANCED:** Hidden Manifold Model (Goldt et al. PRX ‘20, Buchanan ICLR ‘21)]
  - [**TAILORMADE:** Multi-dimensional Fourier series]



// Past focus areas

Xanadu's work in QML

Training quantum circuits efficiently

2018

Evaluating analytic gradients on quantum hardware
M Schuld, V Bergholm, C Gogolin, J Izaac, N Killoran

Quantum machine learning in feature Hilbert spaces
M Schuld, N Killoran

2019

Quantum Natural Gradient
J Stokes, J Izaac, N Killoran, G Carleo

Continuous-variable quantum neural networks
N Killoran, TR Bromley, JM Arrazola, M Schuld, N Quesada, S Lloyd

2020

Estimating the gradient and higher-order derivatives on quantum hardware
A Mari, TR Bromley, N Killoran

Quantum embeddings for machine learning
S Lloyd, M Schuld, A Ijaz, J Izaac, N Killoran

2021

General parameter-shift rules for quantum gradients
D Wierichs, J Izaac, C Wang, C Yen-Yu Lin

The effect of data encoding on the expressive power of variational quantum machine learning models
M Schuld, R Sweke, JJ Meyer

2022

Optimizing quantum circuits with Riemannian gradient flow
R Wiersema, N Killoran

Generalization despite overfitting in quantum machine learning models
E Peters, M Schuld

2023

Backpropagation scaling in parameterised quantum circuits
J Bowles, D Wierichs, C-Y Park

Contextuality and inductive bias in quantum machine learning
J Bowles, VJ Wright, M Farkas, N Killoran, M Schuld

Understand what ML models quantum circuits are

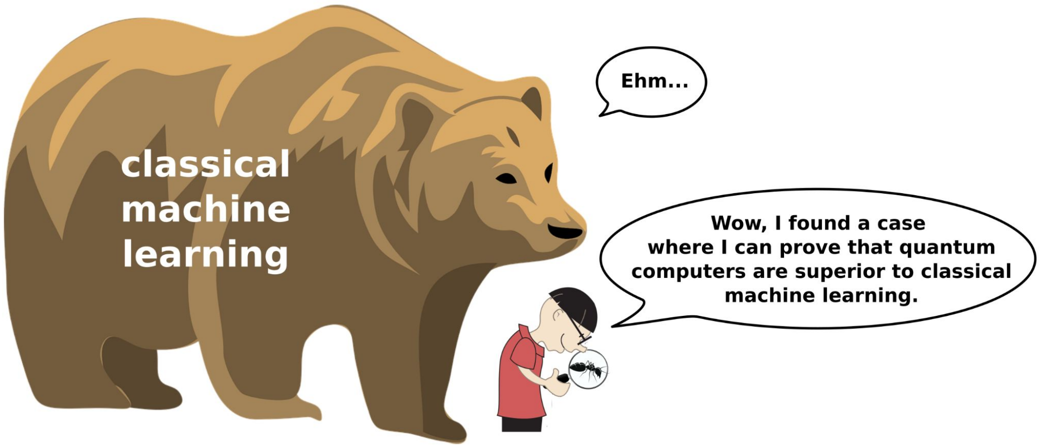
* Xanadu researcher at the time

// Checking the compass

Are we heading in the right direction?

Is quantum advantage the right goal for quantum machine learning?
M Schuld, N Killoran, 2022

PennyLane blog post, 2022
Why measuring performance is our biggest blind spot in quantum machine learning



// Assess how good quantum models really are

What is the best benchmark design we can come up with?

Crucial decisions

- Faithful implementation
  Carefully deduce main model and training procedure from paper
- Convergence criteria
  Compare averages over 2 loss intervals
- Batches in SGD
  Don’t optimize, but adapt to runtime needs
- Data preprocessing
  Always rescale data
- Hyperparameter optimisation grid
  Balance choices from paper, common sense and runtime considerations

Hinton: The Forward-forward algorithm (2022)

Sensibly-engineered convolutional neural nets with a few hidden layers typically get about 0.6% test error. In the "permutation-invariant" version of the task, the neural net is not given any information about the spatial layout of the pixels so it would perform equally well if all of the training and test images were subjected to the same random permutation of the pixels before training started. For the permutation-invariant version of the task, feed-forward neural networks with a few fully connected hidden layers of Rectified Linear Units (ReLUs) typically get about 1.4% test error[6] and they take about 20 epochs to train. This can be reduced to around 1.1% test error using a variety of regularizers such as dropout (Srivastava et al., 2014) (which makes training slower) or label smoothing (Pereyra et al., 2017) (which makes training faster). It can be further reduced by combining supervised learning of the labels with unsupervised learning that models the distribution of images.

To summarize, 1.4% test error on the permutation-invariant version of the task without using complicated regularizers, shows that, for MNIST, a learning procedure works about as well as backpropagation[7].

[5]FF uses the simplest version of layer normalization which does not subtract the mean before dividing by the length of the activity vector.

[6]I have trained thousands of different neural networks on MNIST and, in the spirit of Ramon y Cajal, I believe that it is more informative to report the performance of the non-existent typical net than to report the performance of any particular net.

[7]Some papers on biologically plausible alternatives to backpropagation report test error rates greater than 2% on permutation invariant MNIST. This indicates that the learning procedure does not work nearly as well as backpropagation or that the authors failed to tune it properly.