Machine learning with quantum circuits

Maria Schuld
Mark Fingerhuth, Francesco Petruccione

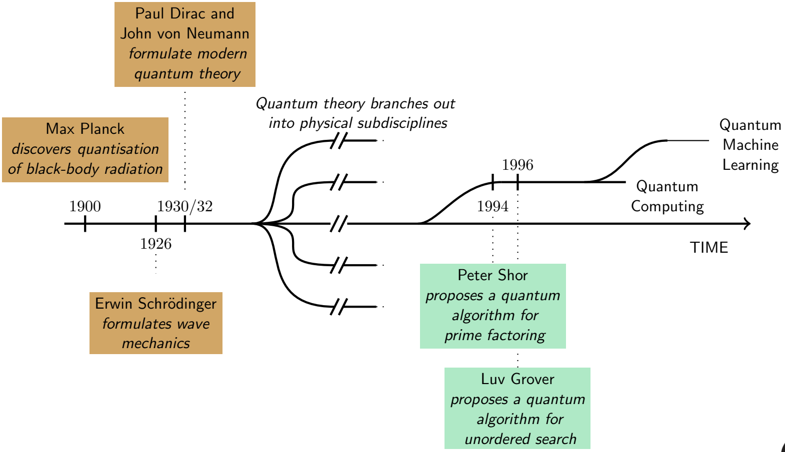
University of KwaZulu-Natal, Durban, South Africa

Beijing, 4th July 2017

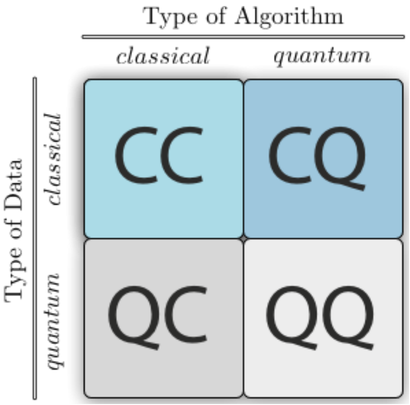
OUTLINE

1. MOTIVATION - Why should you care?
2. OVERVIEW - Where are we in quantum machine learning?
3. DEMONSTRATION - An interference circuit as a binary classifier

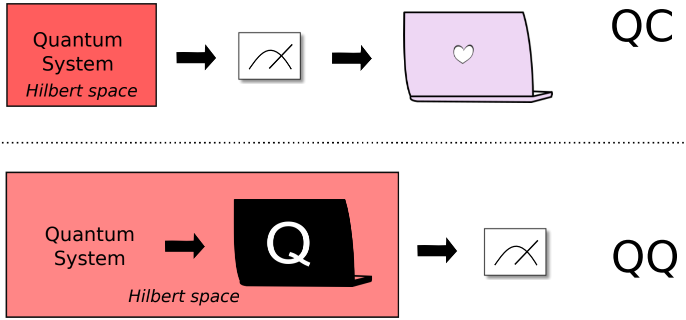
QUANTUM COMPUTING TIMELINE



COMBINING QC AND ML



Aimeur, Brassard and Gambs (2006)


Quantum
System
Hilbert space
QC
Quantum
System
Q
Hilbert space
QQ

Pitch for the open session on Thursday:

Are there problems for which feeding the wave function into a machine learning model could be of advantage?

IMPORTANT ASPECTS OF QML

1. Are quantum methods better?
2. Data encoding and readout
3. Prediction - Implementing the model
4. Training - Solving the optimisation problem

1. ARE QUANTUM METHODS BETTER?

model complexity

sample complexity

computational complexity

(Thanks to Vedran Dunjko)

model complexity

sample complexity

computational complexity

"[F]or any learning problem, if there is a quantum learning algorithm which uses polynomially many [samples] then there must also exist a classical learning algorithm which uses polynomially many [samples]."
Servedio & Gortler 2004

UNIVERSITY OF KWAZULU-NATAL

1. ARE QUANTUM METHODS BETTER?

model complexity

sample complexity

computational complexity

UNIVERSITY OF KWAZULU-NATAL

1. ARE QUANTUM METHODS BETTER?

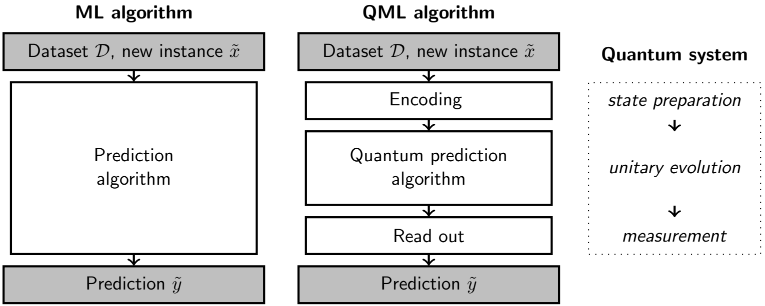
model
complexity

sample
complexity

computational
complexity

UNIVERSITY OF
KWAZULU-NATAL

2. DATA ENCODING AND READOUT



MS, Sinayskiy, Petruccione *Encyclopaedia of ML* (2016)

UNIVERSITY OF KWAZULU-NATAL

index	prob.	amplitude	state
0	$\lvert a_0\rvert^2$	a_0	$\lvert 000\rangle$
1	$\lvert a_1\rvert^2$	a_1	$\lvert 001\rangle$
2	$\lvert a_2\rvert^2$	a_2	$\lvert 010\rangle$
3	$\lvert a_3\rvert^2$	a_3	$\lvert 011\rangle$
4	$\lvert a_4\rvert^2$	a_4	$\lvert 100\rangle$
5	$\lvert a_5\rvert^2$	a_5	$\lvert 101\rangle$
6	$\lvert a_6\rvert^2$	a_6	$\lvert 110\rangle$
7	$\lvert a_7\rvert^2$	a_7	$\lvert 111\rangle$

UNIVERSITY OF KWAZULU-NATAL

Dynamic encoding of a Hermitian matrix

$$U = e^{iHt}$$

basis encoding of binary string (10), *e.g. to represent integer* 2

$$|\psi\rangle = \alpha_0 \,|00\rangle + \alpha_1 \,|01\rangle + \alpha_2 \,|10\rangle + \alpha_3 \,|11\rangle$$

amplitude encoding of unit-length complex vector $(a_0, a_1, a_2, a_3)^T$

UNIVERSITY OF
KWAZULU-NATAL

How to implement $f(\mathbf{x};\theta)$, $f(\mathbf{x};\mathcal{D})$, $p(\mathbf{x}, y)$, $p(y|\mathbf{x})$... ?

UNIVERSITY OF KWAZULU-NATAL

How to implement $f(\mathbf{x};\theta)$, $f(\mathbf{x};\mathcal{D})$, $p(\mathbf{x},y)$, $p(y|\mathbf{x})$... ?

Feed-forward neural network:

$$f(\mathbf{x};\mathbf{W}_1,\mathbf{W}_2,...) = ...\varphi_2(\mathbf{W}_2\varphi_1(\mathbf{W}_1\mathbf{x}))...$$

MS, Sinayskiy, Petruccione *Quant. Inf. Proc.* (2014), *Phys. Lett. A* (2015), and many others...

How to implement $f(\mathbf{x};\theta)$, $f(\mathbf{x};\mathcal{D})$, $p(\mathbf{x},y)$, $p(y|\mathbf{x})$... ?

Feed-forward neural network:

$$f(\mathbf{x};\mathbf{W}_1,\mathbf{W}_2,...) = ...\varphi_2(\mathbf{W}_2\varphi_1(\mathbf{W}_1\mathbf{x}))...$$

RBM:

$$p(\mathbf{x};\mathbf{W}) = \frac{1}{Z}\sum_h e^{-E(x,h;\mathbf{W})}$$

Wiebe, Svore, Kapoor ArXiv1412.3489 (2014)

UNIVERSITY OF KWAZULU-NATAL

How to implement $f(\mathbf{x};\theta)$, $f(\mathbf{x};\mathcal{D})$, $p(\mathbf{x},y)$, $p(y|\mathbf{x})$... ?

Feed-forward neural network:

$$f(\mathbf{x};\mathbf{W}_1,\mathbf{W}_2,...) = ...\varphi_2(\mathbf{W}_2\varphi_1(\mathbf{W}_1\mathbf{x}))...$$

RBM:

$$p(\mathbf{x};\mathbf{W}) = \frac{1}{Z}\sum_h e^{-E(x,h;\mathbf{W})}$$

Linear model:

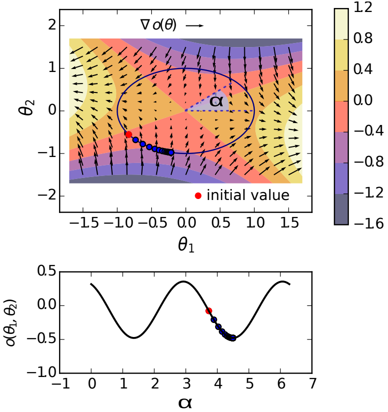
$$f(\mathbf{x};\mathbf{w}) = \mathbf{w}^T\mathbf{x}$$

Rebentrost, Mohseni, Lloyd *PRL* (2014)

Problems	QIP tools	Applied to
Simulating linear algebra calculus with qubits		
matrix inversion, inner products, eigenvalue decomposition, singular value decomposition	quantum phase estimation, postselective amplitude update, Hamiltonian simulation, density matrix exponentiation	support vector machines, Gaussian processes, linear regression, discriminant analysis, recommendation systems, principal component analysis
Optimisation with Grover search		
finding closest neighbours, Markov chains	amplitude amplification, quantum walks	k-nearest neighbour, page ranking, clustering, associative memory, perceptrons, active learning agents, natural language processing
Sampling from quantum states		
sampling from model distribution	quantum annealing, quantum rejection sampling	Boltzmann machines, Bayesian nets, Bayesian inference
Optimisation with ground states of Hamiltonians		
combinatorial optimisation	adiabatic quantum computing, quantum annealing, quantum simulation	associative memory, boosting, debugging, variational Bayes inference, Bayesian networks, perceptron, EM algorithm, clustering

UNIVERSITY OF KWAZULU-NATAL

What about gradient descent?

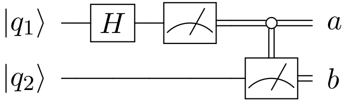


Rebentrost, MS, Petruccione, Lloyd (2016) ArXiv1612.01789

Goal:
Show how a simple interference circuit can realise a distance-based binary classifier.

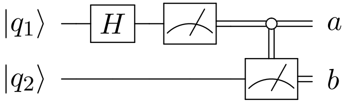
MS, Fingerhuth, Petruccione (2017) ArXiv1703.10793

UNIVERSITY OF KWAZULU-NATAL

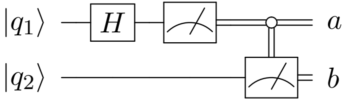


$|q_1\rangle$
H
a
$|q_2\rangle$
b

INTERFERENCE CIRCUIT

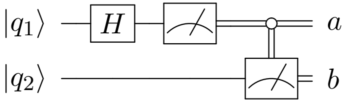


probability	amplitude	$\|q_1, q_2\rangle$
$\|a_0\|^2$	a_0	$\|00\rangle$
$\|a_1\|^2$	a_1	$\|01\rangle$
$\|a_2\|^2$	a_2	$\|10\rangle$
$\|a_3\|^2$	a_3	$\|11\rangle$



probability	amplitude	$\|q_1, q_2\rangle$
$\frac{1}{2}\|a_0 + a_2\|^2$	$\frac{1}{\sqrt{2}}(a_0 + a_2)$	$\|00\rangle$
$\frac{1}{2}\|a_1 + a_3\|^2$	$\frac{1}{\sqrt{2}}(a_1 + a_3)$	$\|01\rangle$
$\frac{1}{2}\|a_0 - a_2\|^2$	$\frac{1}{\sqrt{2}}(a_0 - a_2)$	$\|10\rangle$
$\frac{1}{2}\|a_1 - a_3\|^2$	$\frac{1}{\sqrt{2}}(a_1 - a_3)$	$\|11\rangle$

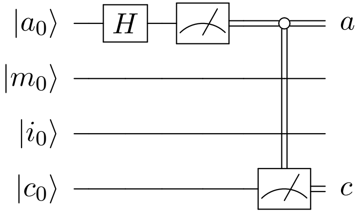
UNIVERSITY OF KWAZULU-NATAL

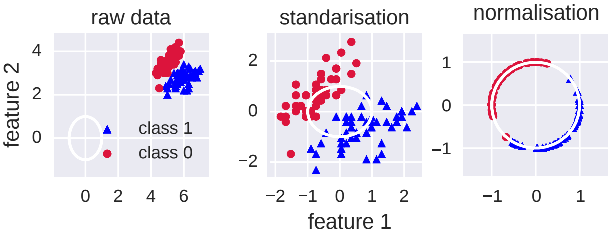


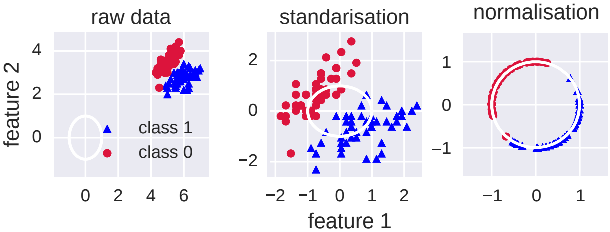
probability	amplitude	$\|q_1, q_2\rangle$
0	0	$\|00\rangle$
0	0	$\|01\rangle$
$\frac{1}{2p_{q_1=1}}\|a_0 - a_2\|^2$	$\frac{1}{\sqrt{2p_{q_1=1}}}(a_0 - a_2)$	$\|10\rangle$
$\frac{1}{2p_{q_1=1}}\|a_1 - a_3\|^2$	$\frac{1}{\sqrt{2p_{q_1=1}}}(a_1 - a_3)$	$\|11\rangle$

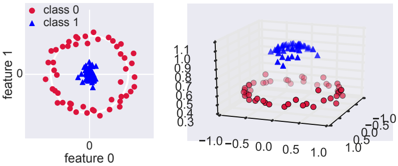
State preparation +

1 x single qubit Hadamard
1 x single qubit postselective measurement
1 x single qubit measurement

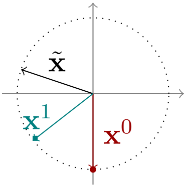



raw data
standarisation
normalisation
feature 2
feature 1
class 1
class 0
4
2
0
0
2
4
6
2
0
−2
−2
−1
0
1
2
1
0
−1
−1
0
1


raw data
standarisation
normalisation
feature 2
feature 1
class 1
class 0


class 0
class 1
feature 1
feature 0

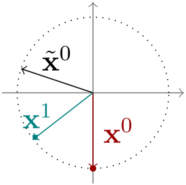
$$|\psi_{\mathcal{D}}\rangle = \frac{1}{\sqrt{2M}} \sum_{m=0}^{M-1} |m\rangle \Big(|0\rangle|\tilde{\mathbf{x}}\rangle + |1\rangle|\mathbf{x}^m\rangle \Big) |y^m\rangle$$



amplitude	$\|m\rangle$	ancilla	$\|i\rangle$	target
$\frac{1}{2} \cdot -0.948$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$
$\frac{1}{2} \cdot 0.318$	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$
0	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$
$\frac{1}{2} \cdot -1.000$	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$
$\frac{1}{2} \cdot -0.948$	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$
$\frac{1}{2} \cdot 0.318$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$
$\frac{1}{2} \cdot -0.789$	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$
$\frac{1}{2} \cdot -0.615$	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$

UNIVERSITY OF KWAZULU-NATAL

$$\frac{1}{2\sqrt{M}} \sum_{m=0}^{M-1} |m\rangle \Big(|0\rangle \left[|\tilde{\mathbf{x}}\rangle + |\mathbf{x}^m\rangle \right] + |1\rangle \left[|\tilde{\mathbf{x}}\rangle - |\mathbf{x}^m\rangle \right] \Big) |y^m\rangle$$

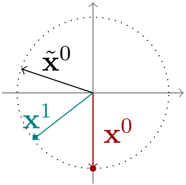


amplitude	$\|m\rangle$	ancilla	$\|i\rangle$	target
$\frac{1}{2\sqrt{2}}(0.948+0)$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$
$\frac{1}{2\sqrt{2}}(0.318+1.000)$	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$
$\frac{1}{2\sqrt{2}}(0.948-0)$	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$
$\frac{1}{2\sqrt{2}}(0.318-1.000)$	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$
$\frac{1}{2\sqrt{2}}(0.948+0.789)$	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$
$\frac{1}{2\sqrt{2}}(0.615+0.318)$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$
$\frac{1}{2\sqrt{2}}(0.948-0.789)$	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$
$\frac{1}{2\sqrt{2}}(0.615-0.318)$	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$

UNIVERSITY OF KWAZULU-NATAL

POSTSELECTIVE MEASUREMENT

$$\frac{1}{2\sqrt{Mp}}\sum_{m=1}^{M}|m\rangle\sum_{i=0}^{N-1}\left(\tilde{x}_i + x_i^m\right)|i\rangle|y^m\rangle$$

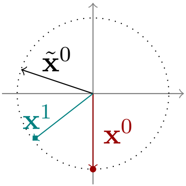


amplitude	$\|m\rangle$	ancilla	$\|i\rangle$	target
$\frac{1}{2\sqrt{2p}}(0.948+0)$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$
$\frac{1}{2\sqrt{2p}}(0.318+1.000)$	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$
0	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$
0	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$
0	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$
$\frac{1}{2\sqrt{2p}}(0.948+0.789)$	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$	$\|0\rangle$
$\frac{1}{2\sqrt{2p}}(0.615+0.318)$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$	$\|0\rangle$
0	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$
0	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$
0	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$	$\|1\rangle$

UNIVERSITY OF KWAZULU-NATAL

POSTSELECTIVE MEASUREMENT

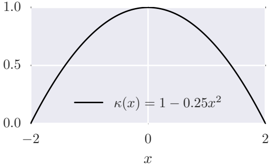
$$\sum_{m=1}^{M} |m\rangle \sum_{i=0}^{N-1} \sqrt{1 - \frac{1}{4Mp}\left(\tilde{x}_i - x_i^m\right)^2} |i\rangle |y^m\rangle$$

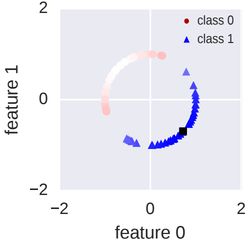


amplitude	$\lvert m\rangle$	ancilla	$\lvert i\rangle$	target
$\frac{1}{2\sqrt{2p}}(0.948 + 0)$	$\lvert 0\rangle$	$\lvert 0\rangle$	$\lvert 0\rangle$	$\lvert 0\rangle$
0	$\lvert 0\rangle$	$\lvert 0\rangle$	$\lvert 0\rangle$	$\lvert 1\rangle$
$\frac{1}{2\sqrt{2p}}(0.318 + 1.000)$	$\lvert 0\rangle$	$\lvert 0\rangle$	$\lvert 1\rangle$	$\lvert 0\rangle$
0	$\lvert 0\rangle$	$\lvert 0\rangle$	$\lvert 1\rangle$	$\lvert 1\rangle$
0	$\lvert 0\rangle$	$\lvert 1\rangle$	$\lvert 0\rangle$	$\lvert 0\rangle$
0	$\lvert 0\rangle$	$\lvert 1\rangle$	$\lvert 0\rangle$	$\lvert 1\rangle$
0	$\lvert 0\rangle$	$\lvert 1\rangle$	$\lvert 1\rangle$	$\lvert 0\rangle$
0	$\lvert 0\rangle$	$\lvert 1\rangle$	$\lvert 1\rangle$	$\lvert 1\rangle$
0	$\lvert 1\rangle$	$\lvert 0\rangle$	$\lvert 0\rangle$	$\lvert 0\rangle$
$\frac{1}{2\sqrt{2p}}(0.948 + 0.789)$	$\lvert 1\rangle$	$\lvert 0\rangle$	$\lvert 0\rangle$	$\lvert 1\rangle$
0	$\lvert 1\rangle$	$\lvert 0\rangle$	$\lvert 1\rangle$	$\lvert 0\rangle$
$\frac{1}{2\sqrt{2p}}(0.615 + 0.318)$	$\lvert 1\rangle$	$\lvert 0\rangle$	$\lvert 1\rangle$	$\lvert 1\rangle$
0	$\lvert 1\rangle$	$\lvert 1\rangle$	$\lvert 0\rangle$	$\lvert 0\rangle$
0	$\lvert 1\rangle$	$\lvert 1\rangle$	$\lvert 0\rangle$	$\lvert 1\rangle$
0	$\lvert 1\rangle$	$\lvert 1\rangle$	$\lvert 1\rangle$	$\lvert 0\rangle$
0	$\lvert 1\rangle$	$\lvert 1\rangle$	$\lvert 1\rangle$	$\lvert 1\rangle$

THE CLASSIFIER

$$\mathbb{E}[\text{target}] = \sum_{m=0}^{M-1} (2y^m - 1)\left[1 - \frac{1}{4Mp}|\mathbf{x} - \mathbf{x}^m|^2\right]$$

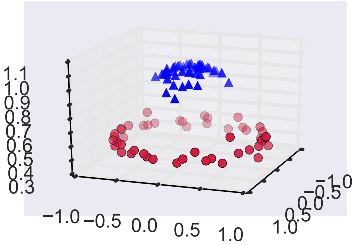




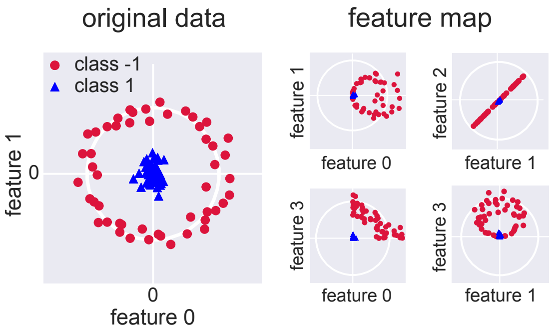
UNIVERSITY OF KWAZULU-NATAL

Dataset	test error	variance	p_{acc}
Iris class 1&2	0.00	0.000	0.50
Iris class 1&3	0.00	0.000	0.50
Iris class 2&3	0.07	0.003	0.50
Circles	0.62	0.006	0.50

Dataset	test error	variance	p_{acc}
Iris class 1&2	0.00	0.000	0.50
Iris class 1&3	0.00	0.000	0.50
Iris class 2&3	0.07	0.003	0.50
Circles	0.62	0.006	0.50



Use $|\mathbf{x}\rangle \rightarrow |\mathbf{x}\rangle \otimes |\mathbf{x}\rangle$



Dataset	test error	variance	p_{acc}
Iris class 1&2	0.00	0.000	0.50
Iris class 1&3	0.00	0.000	0.50
Iris class 2&3	0.07	0.003	0.50
Iris class 2&3, feat map	0.00	0.000	0.50
Circles	0.62	0.006	0.50
Circles, feat map	0.00	0.000	0.55

TAKE HOME

- Quantum machine learning wants to use quantum computers for learning tasks
- A lot of proposals show how to use quantum routines to make training computationally faster (BUT READ THE FINE PRINT :)
- We should start asking which kind of new models quantum computing can genuinely realise
- Quantum classifier for quantum data?

UNIVERSITY OF KWAZULU-NATAL