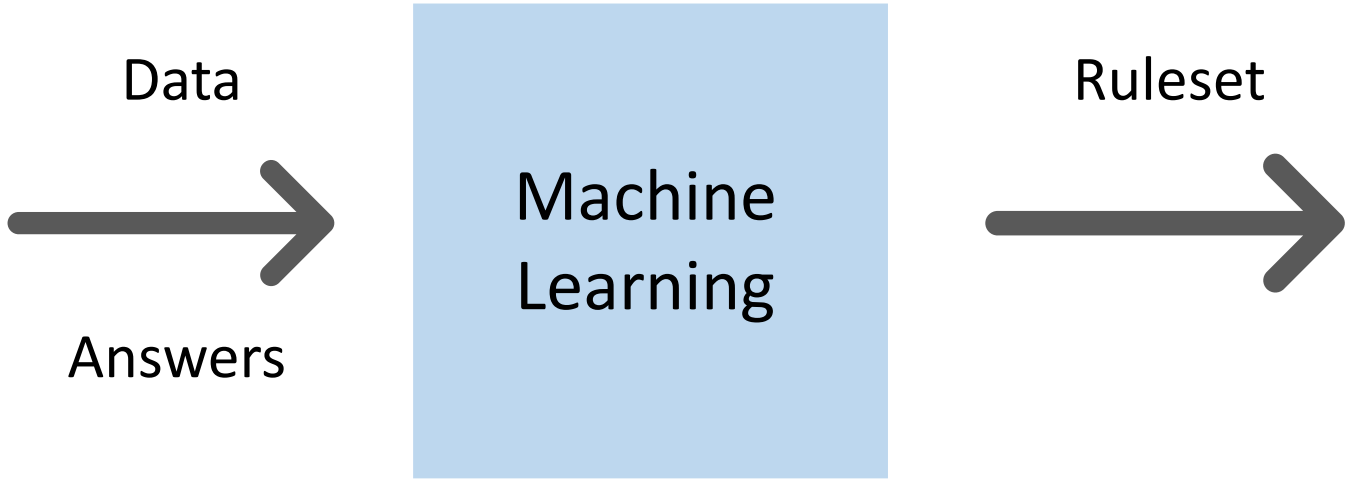
Durham University

Quantum Machine Learning
An Introduction

By Jack Streeter

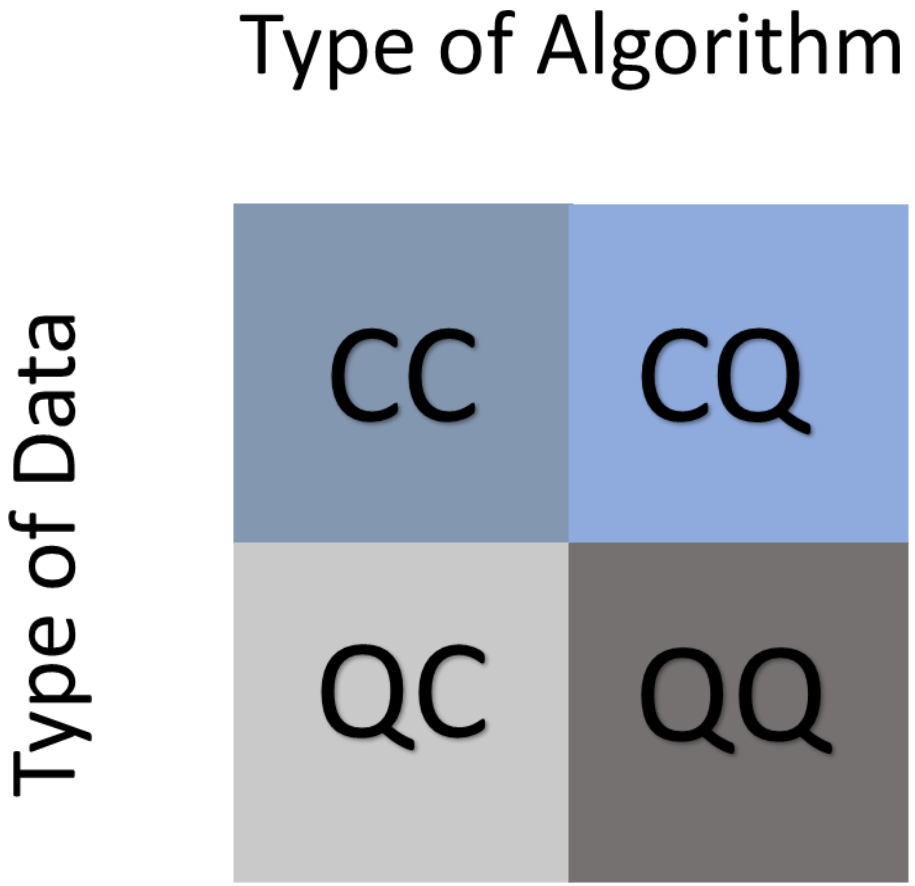
Durham University

What is Machine Learning?



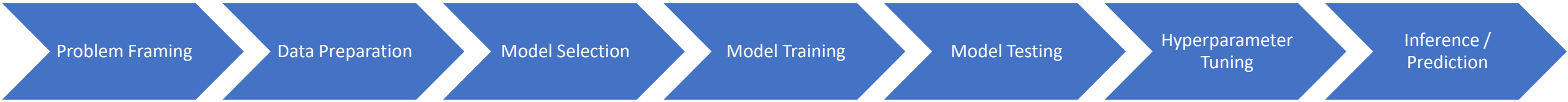
Durham University

Classical vs Quantum



Durham University

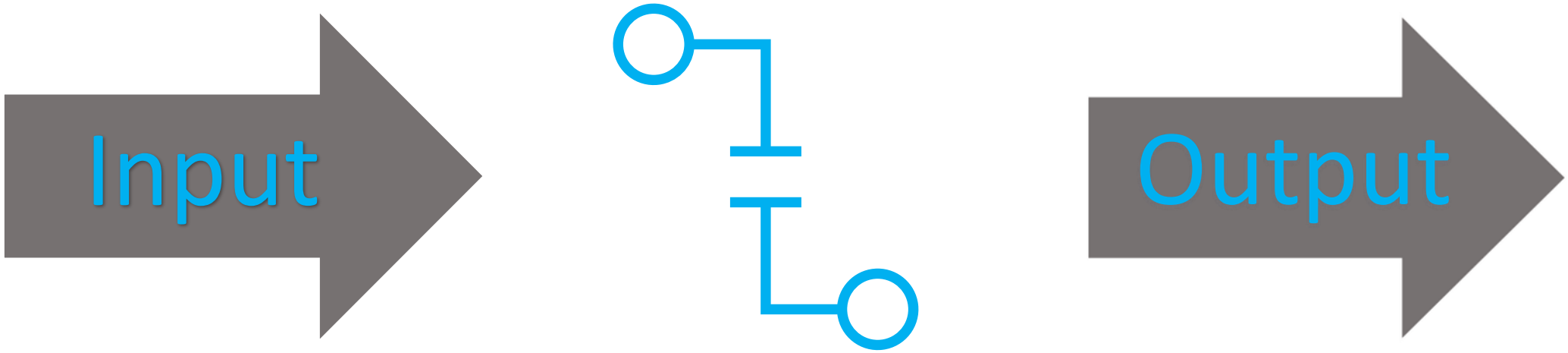
Workflow



Durham University

Supervised Learning

Durham University



$$\hat{y} = h(x)$$

Durham University

Credit Score =

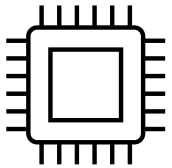
f(Total Credit, Short Term Loan, Credit Utilisation, Missing Payments)

Durham University

Classical Algorithm: Regression

Durham University

Input x / independent variable



Model Function

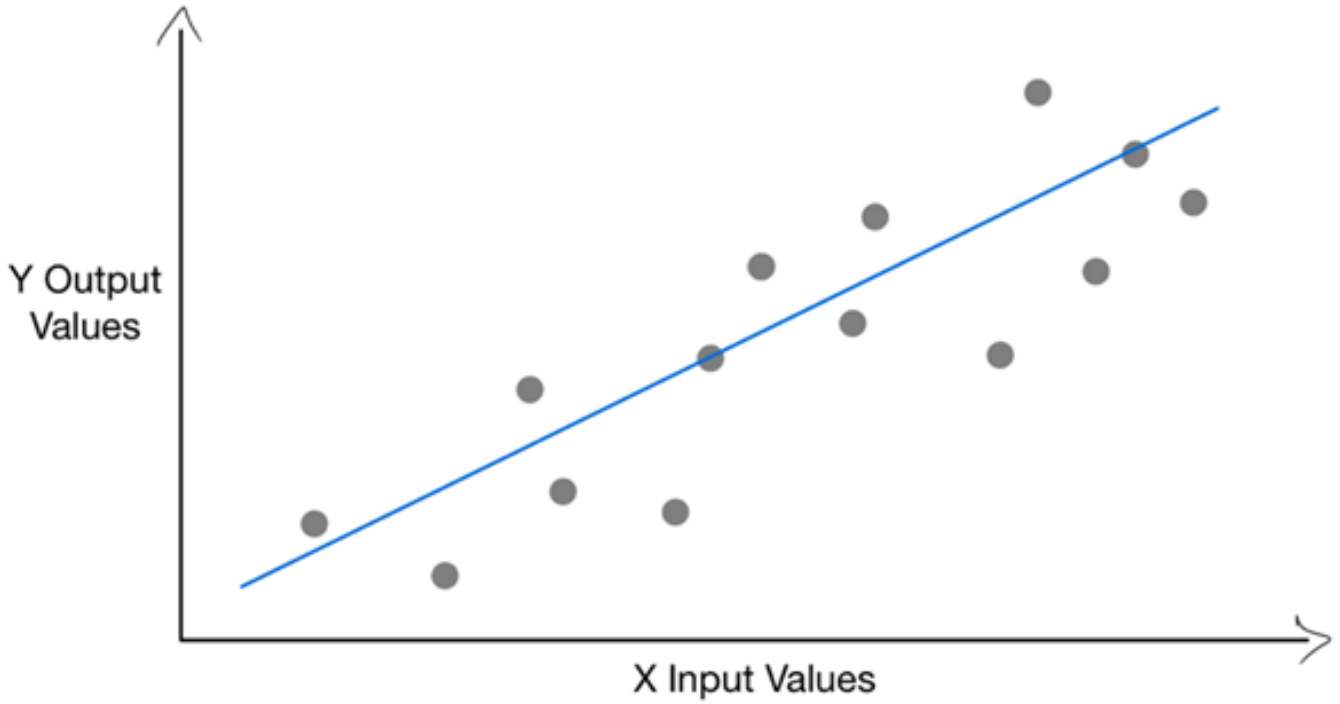
Output y / dependant variable

Hypothesis function:

$$h_\theta(x) = \theta_0 + \theta_1 x$$

Durham University

Instance	x	y
1	x_1	y_1
2	x_2	y_2
…	…	…
i	x_i	y_i
…	…	…
m	x_m	y_m



x – input

y – output

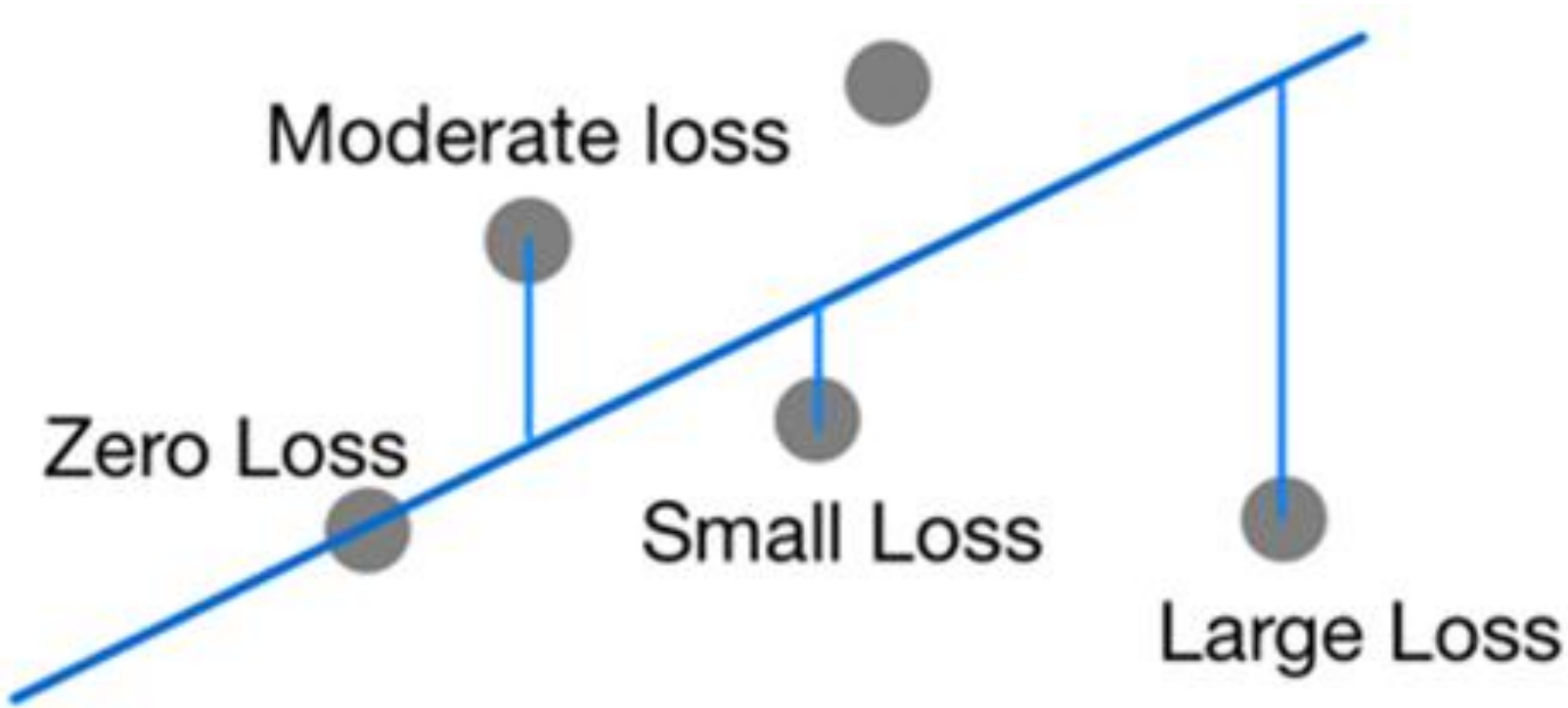
m – number of instances

Durham University

Cost Function

$$\mathrm{J} = \frac{1}{2m}\sum_{i=1}^{m}\left(y^{(i)} - \hat{y}^{(i)}\right)^2$$

Durham University



Durham University

Multivariable Regression

Instance	x_1	x_2	x_3	x_4	y
1					
2					
…	…	…	…	…	…
i					
…	…	…	…	…	…
m					

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \theta_4 x_4$$

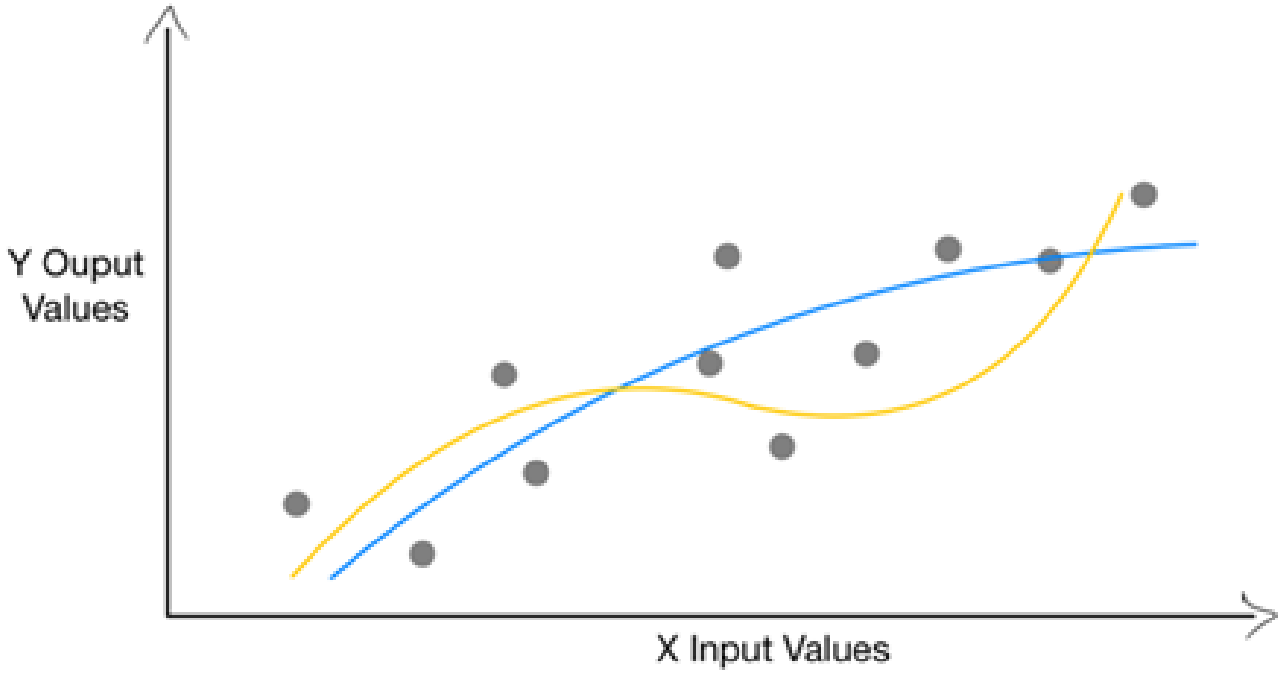
Durham University

Polynomial Regression

$$h_\theta(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \cdots + \theta_n x^k$$

Durham University

Bayesian Information Criterion



$$BIC_k =$$

$$m * ln(SS_\epsilon) + k * \ln(m)$$

Durham University

Quantum Machine Learning

Durham University

Parametrised Quantum Circuits

$$|\psi_\theta\rangle = U_\theta |\phi_0\rangle$$

Durham University

Training Parametrised Quantum Circuits

$\vec{\theta}_n$

Update Parameters

Compute

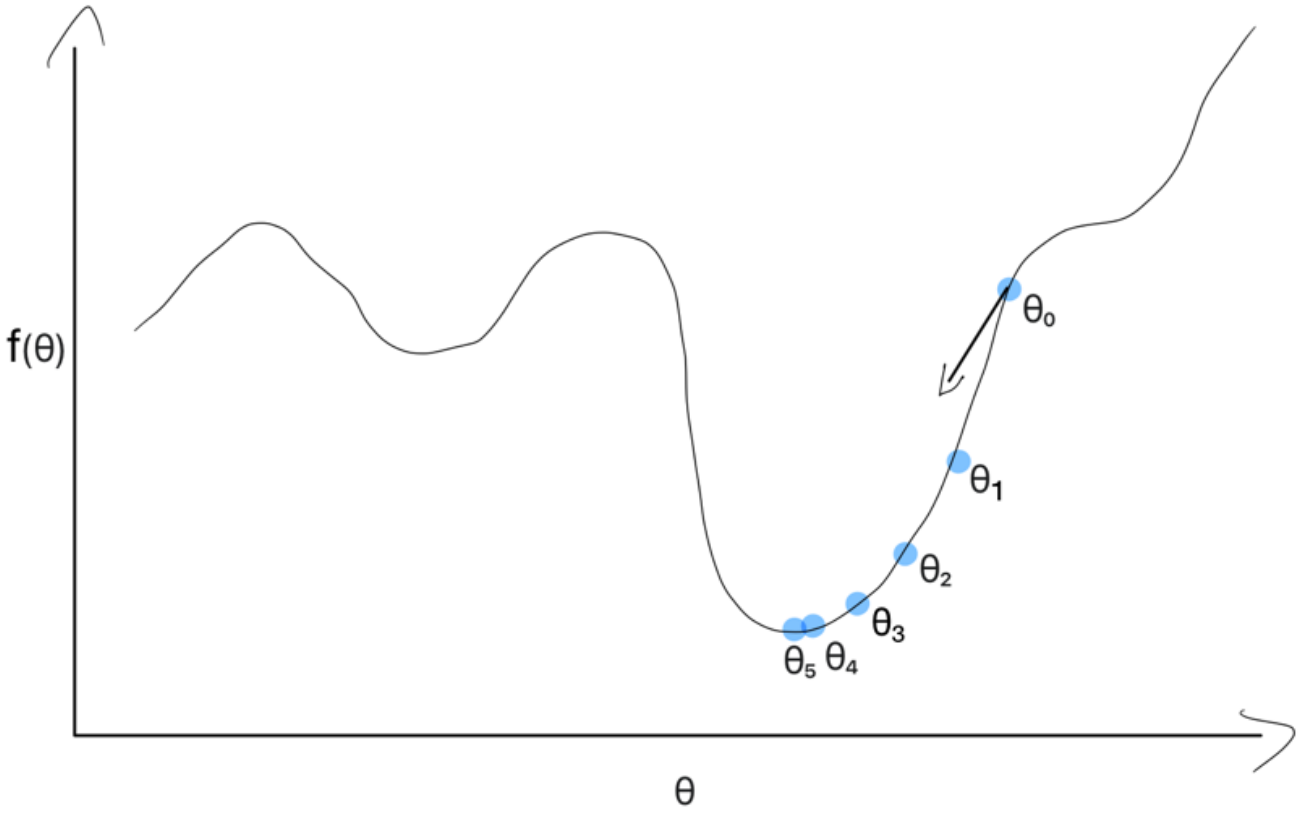
$\vec{\theta}_n \rightarrow \vec{\theta}_{n+1}$

$\langle \Psi(\vec{\theta}) \mid \hat{H} \mid \Psi(\vec{\theta}) \rangle$

Function Value

Durham University

Training Parametrised Quantum Circuits



$$\vec{\theta}_{n+1} = \vec{\theta}_n - \eta \nabla f(\vec{\theta}_n)$$

Durham University

Data Encoding

$$Data\ set\ Y = (x^{(1)}, \dots, x^{(m)}, \dots, x^{(M)})$$

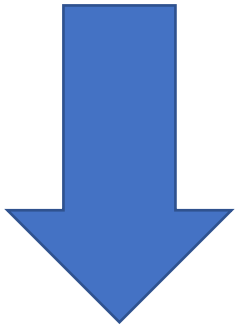
Where

$$x^{(m)} = (b_1, b_2, \dots, b_N)$$

Durham University

Data Encoding

$$x^{(m)} = (b_1,\ b_2,\ \dots, b_N)$$



$$|x^{(m)}\rangle = |b_1, b_2, \dots\ , b_N\rangle$$

Where b_N = 0 or 1

Durham University

Data Encoding

$$|Y\rangle = \frac{1}{\sqrt{N}} \sum_{m=1}^{M} |x^m\rangle$$

Durham University

Variational Classification

Durham University

Variational Classification

$$|\psi(\vec{x}_i)\rangle = U_{W(\vec{\theta})} U_{\Phi \vec{x}_i} |0\rangle$$