Quantum vs Classical Machine Learning for Vector Boson Scattering Background Reduction at the Large Hadron Collider

Dario Gerace

Dipartimento di Fisica, Università di Pavia

D. Cugini, A. Perego, D. Valsecchi*, P. Govoni*

Dipartimento di Fisica, Università di Milano Bicocca

**CERN, Geneva (CH)*

V Workshop on HPC and Quantum Computing,
CINECA – Bologna, 15/12/2022

Acknowledgments & Collaborators

Pietro Govoni

Davide Valsecchi (PhD)

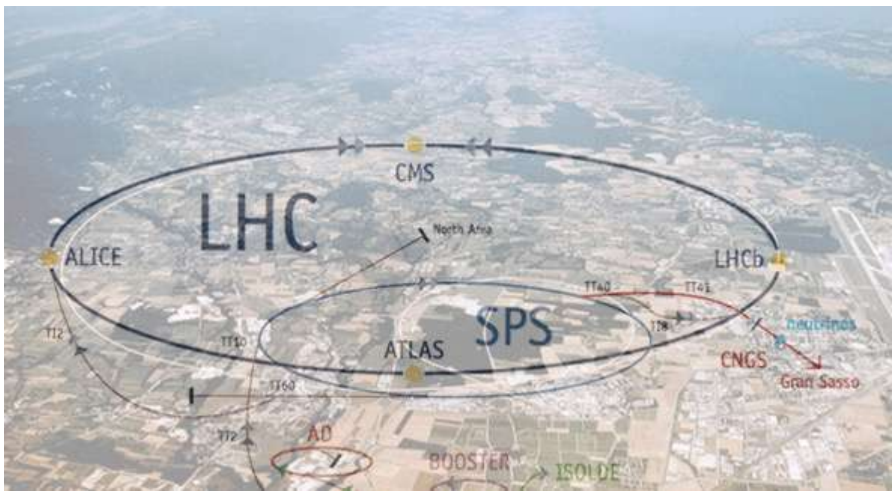
Davide Cugini

Aurora Perego

The Large Hadron Collider

The largest particle accelerator in the world

~8.6 km diameter ring ~100 m below Geneva



P. Higgs at the Compact Muon Solenoid (**CMS**)

The Big Data challenge at LHC

- LHC produces huge amount of data, coming from 40 million p-p collisions/second, producing PB/hour data streams at each detector, requiring sophisticated computing tools

The World LHC Computational Grid (WLCG) project...
~1.4M cores, ~1.5 EB storage,
from 170 sites in 42 countries

The Big Data challenge at LHC

- LHC produces huge amount of data, coming from 40 million p-p collisions/second, producing PB/hour data streams at each detector, requiring sophisticated computing tools

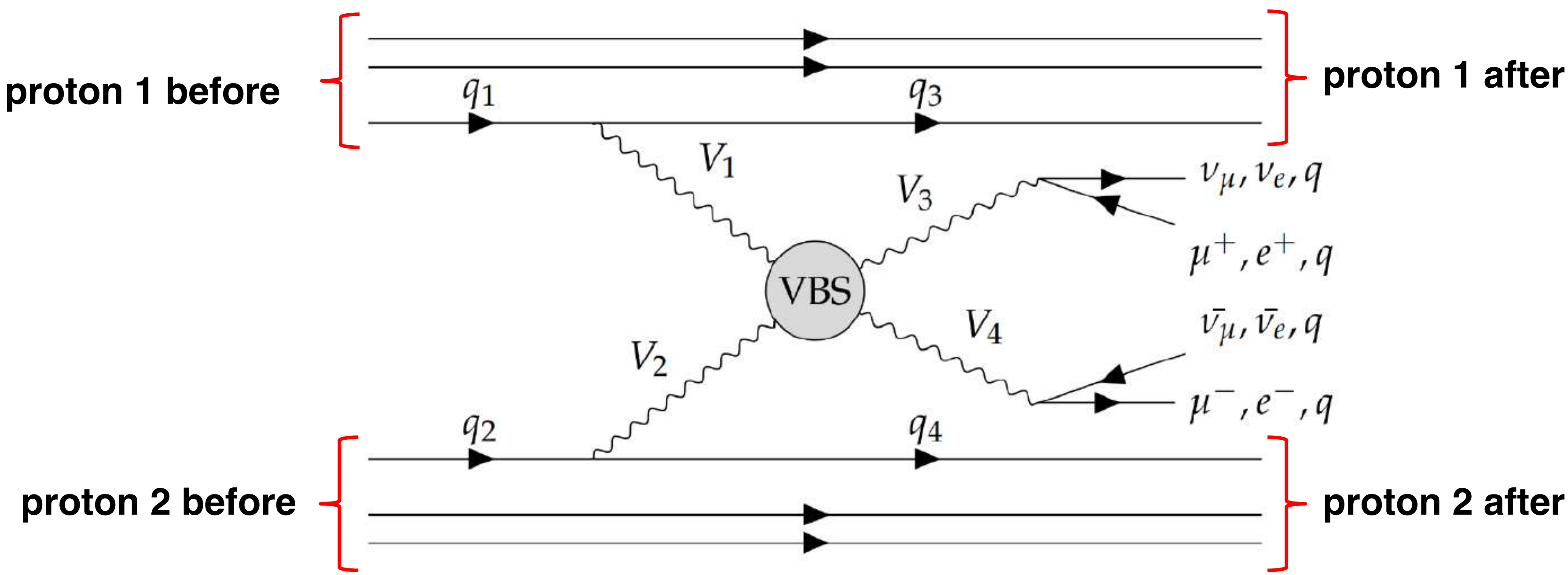
The World LHC Computational Grid (WLCG) project...
~1.4M cores, ~1.5 EB storage,
from 170 sites in 42 countries

- Machine Learning (ML) tools are becoming increasingly relevant in different aspects, but particularly to select the targeted events from the background of other highly probable and indistinguishable events

Grossi, Novak, Kersevan, Rebuzzi, Eur. Phys. J. C **80**, 1144 (2020)

The Vector Boson Scattering

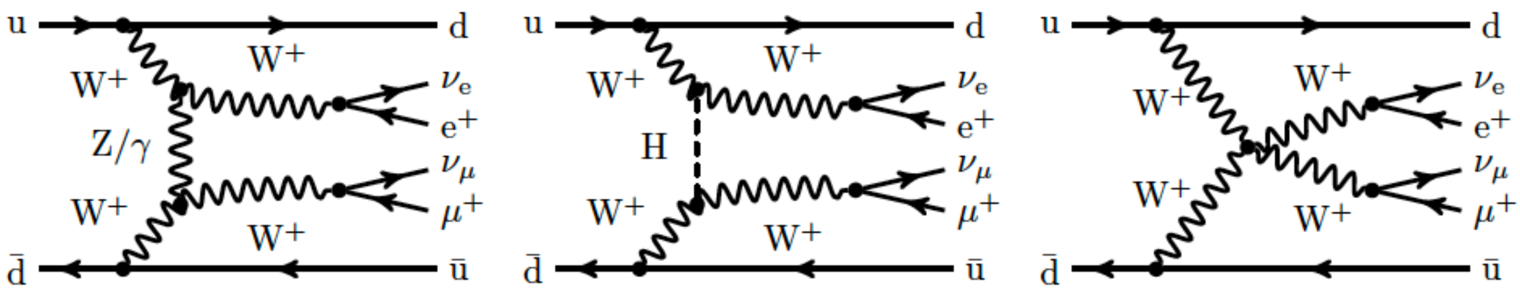
Generic Feynman diagram from p-p collision and involving vector boson scattering (VBS), with possible products



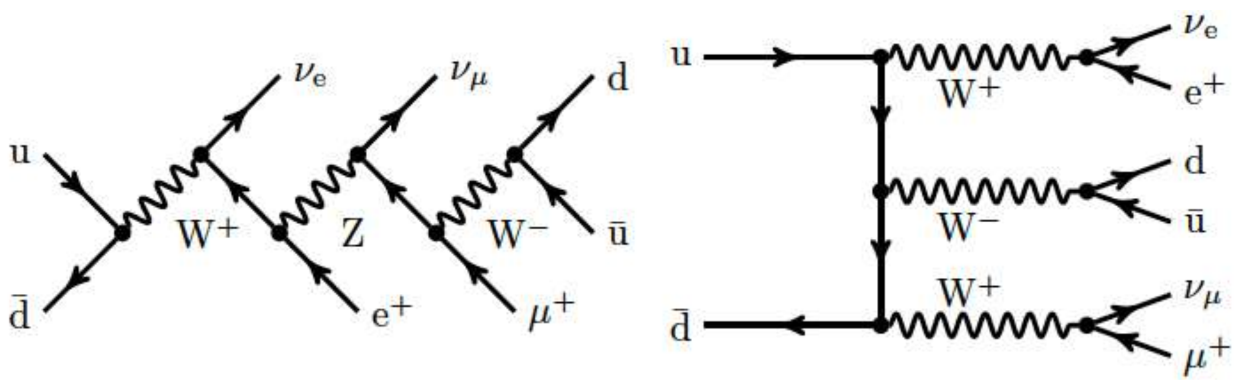
Covarelli, Pellen, Zaro, arxiv:2102.10991 (2021)

Signal vs background events

- some VBS events contributing to <u>signal</u>:



- possible non-VBS events contributing to <u>background</u>:



Covarelli, Pellen, Zaro, arxiv:2102.10991 (2021)

Simulated data set (CMS collaboration)

- VBS events used in this work were generated with MadGraph5_aMC@NLO (Montecarlo generator fully developer within the CMS collaboration)

- CMS detector simulated with Geant4

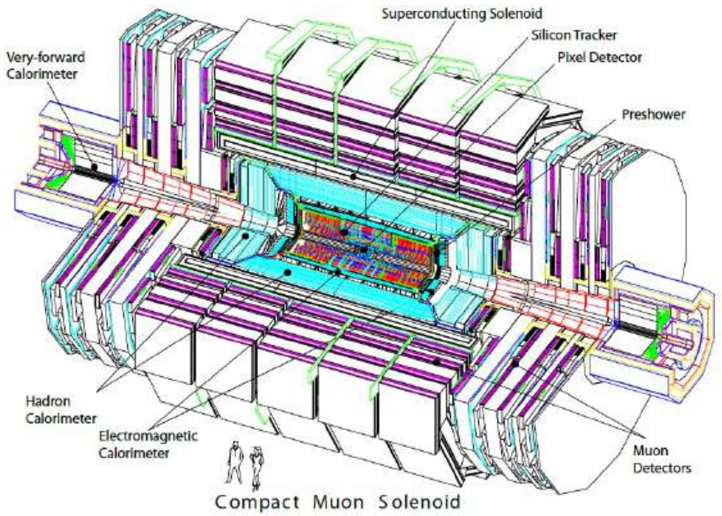
Simulated data set (CMS collaboration)

- VBS events used in this work were generated with MadGraph5_aMC@NLO (Montecarlo generator fully developer within the CMS collaboration)

- CMS detector simulated with Geant4

- Background events generated with either MadGraph or Phantom, taking into account the known ratio between different processes ($1/10^4$ is the signal/background occurence)

	Processo	XS (pb)
Segnale	VBS	2.2
Fondi principali	W+jets	$6.1 \cdot 10^4$
	WW	114.7
	ZZ	16.5
	Top	974

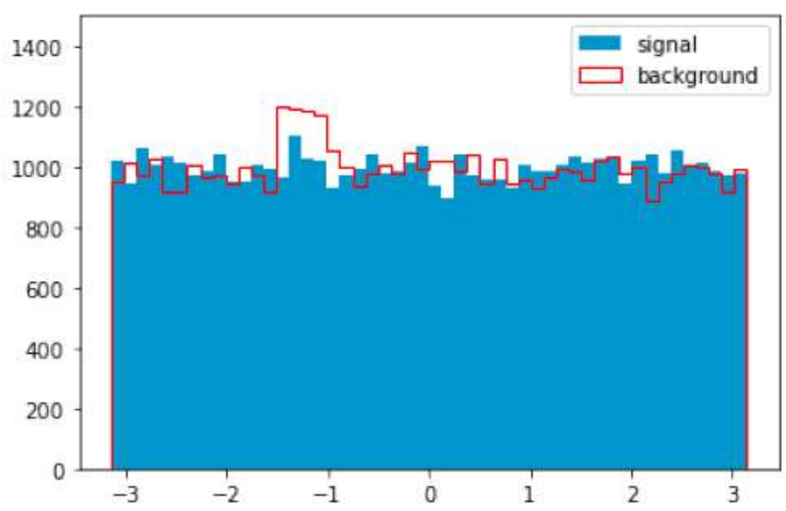


Variables selection

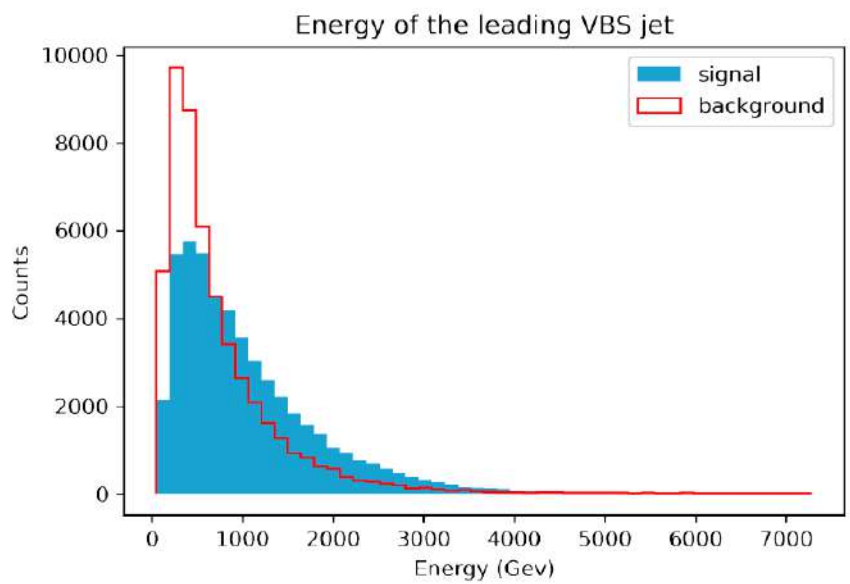
UNIVERSITÀ DI PAVIA

- The variables chosen for the training are selected by looking at the ones with the largest differences between statistical distributions of signal and background

example

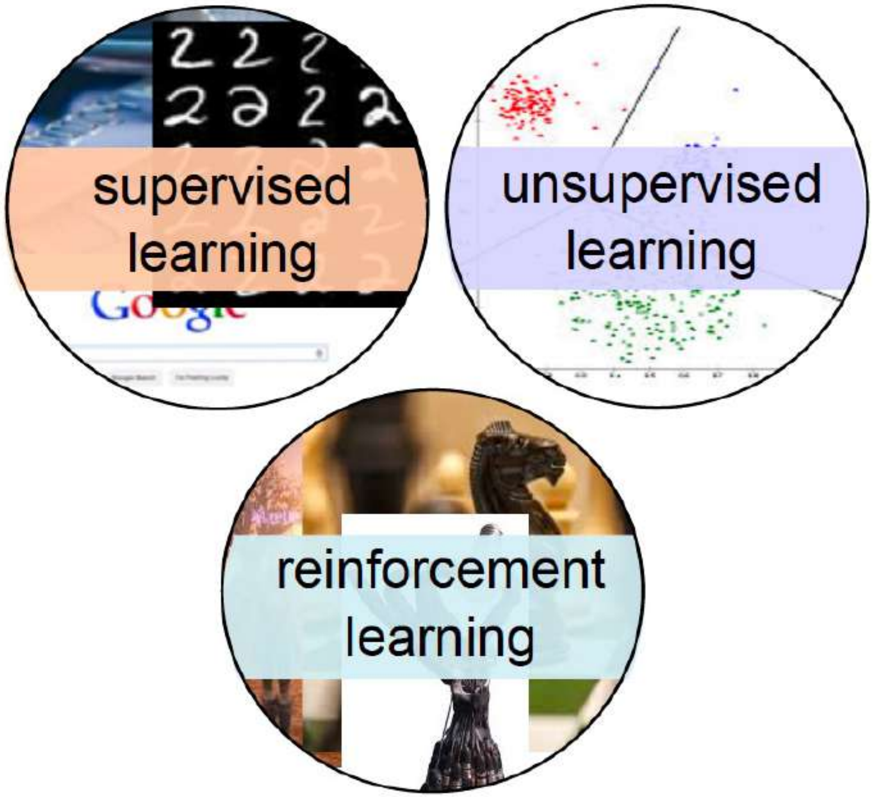


(a) *Azimuth angle of lepton (8)*



Machine Learning

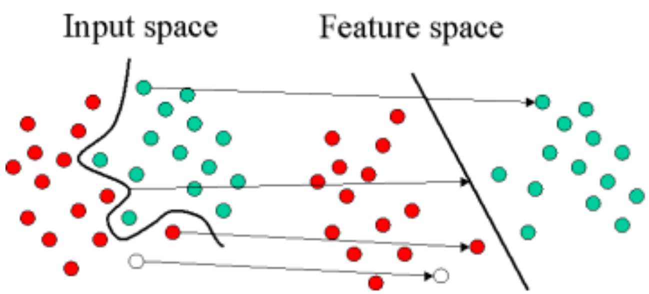
ML is based on finding suitable mathematical models (functions) mapping **input data** into **output predictions**

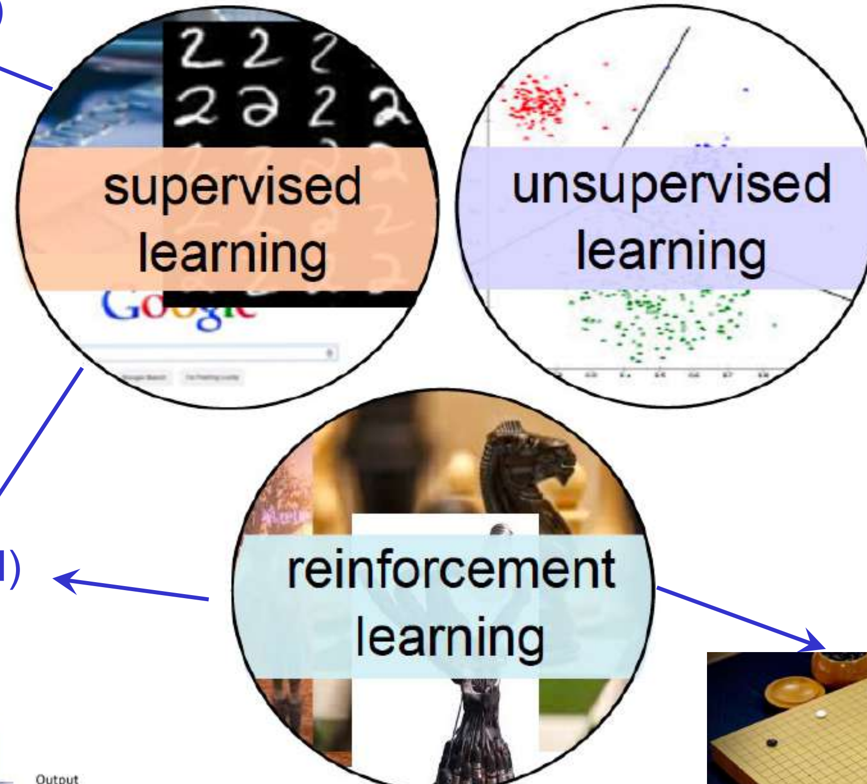


Machine Learning

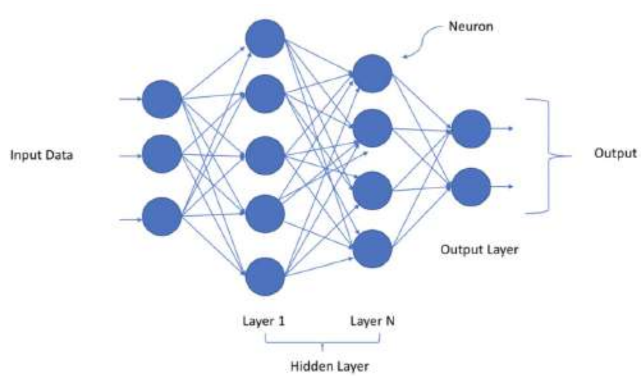
ML is based on finding suitable mathematical models (functions) mapping **input data** into **output predictions**

E.g., support vector machines (SVM)





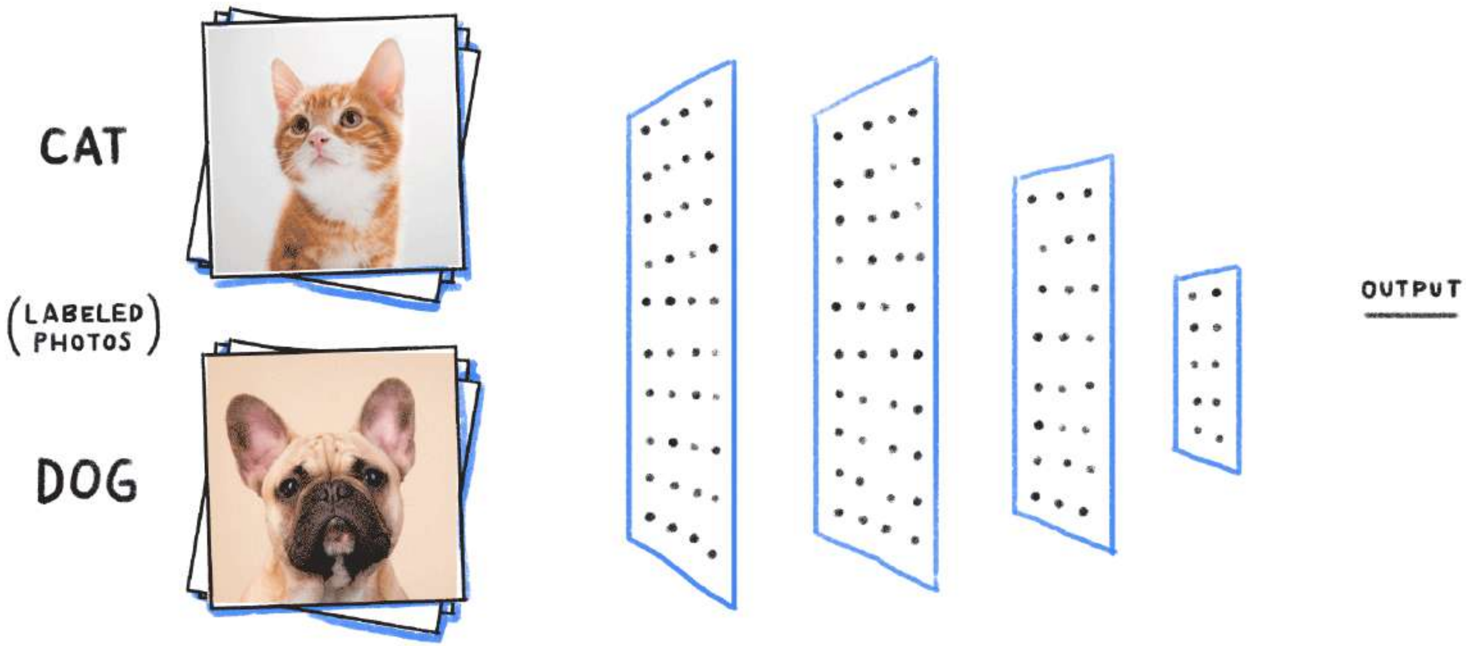
E.g., deep neural networks (DNN)



D. Silver et al.,
Nature **550**, 354 (2017)

Supervised learning by DNN

- Image classification after DNN training



Credits: becominghuman.ai/building-an-image-classifier-using-deep-learning-in-python-totally-from-a-beginners-perspective-be8dbaf22dd8

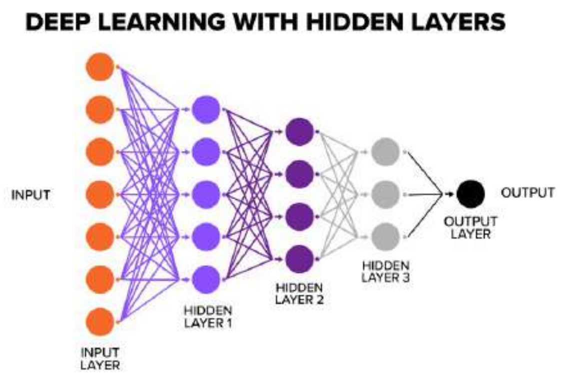
- In general:

$$\boldsymbol{x} \longrightarrow \boxed{f(\boldsymbol{x},\theta)} \longrightarrow \boldsymbol{y}$$

ML task is to learn how f maps $\boldsymbol{x}$ into $\boldsymbol{y}$, on varying θ, such that the algorithm will correctly predict $\boldsymbol{y}$ upon being fed with a previously unknown $\boldsymbol{x}$

Classical ML tools in this work

- DNN training with a set of labelled background and signal events

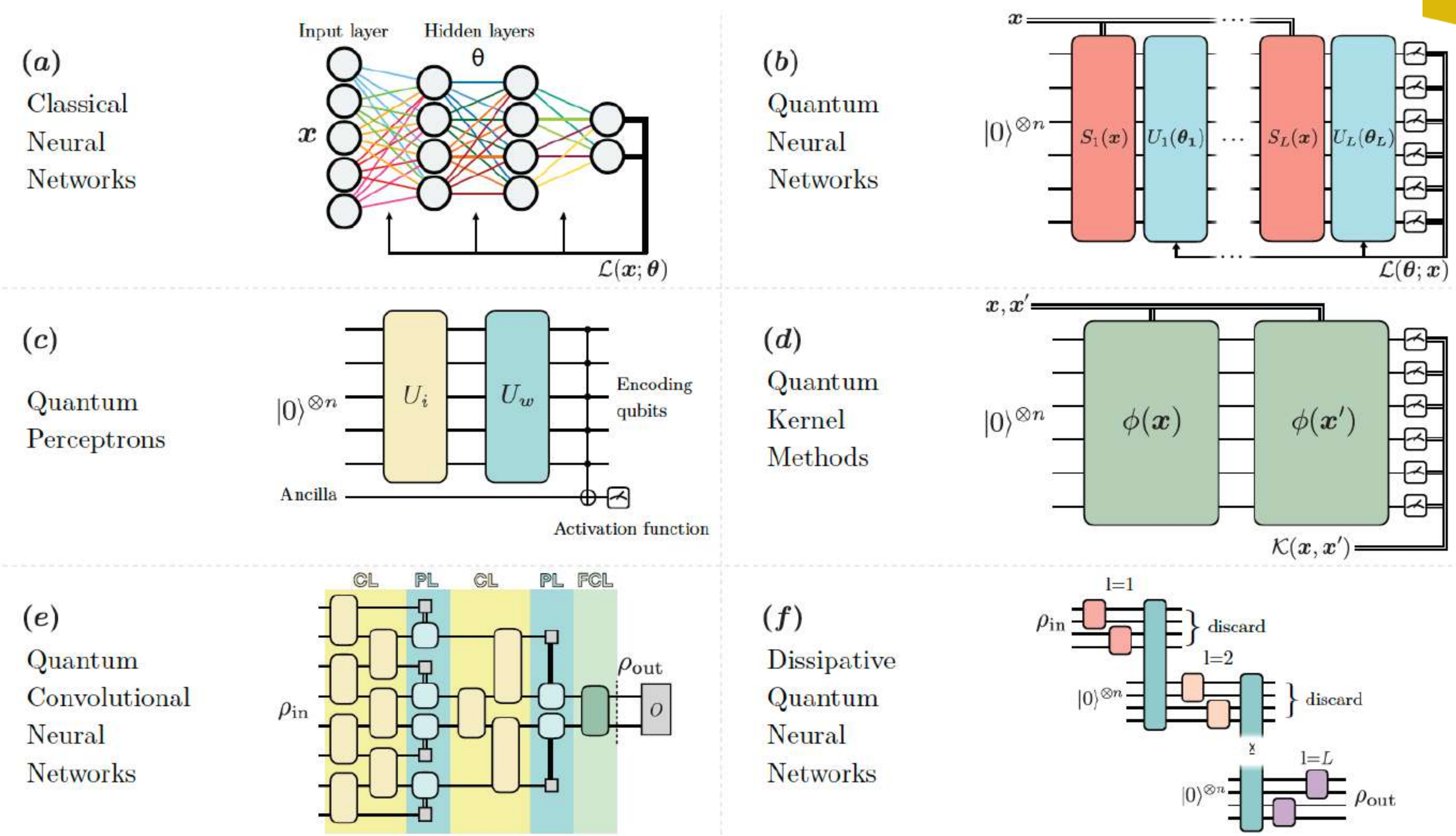


typicaly between 5000 and 50000 nodes

- minimization of cost function to find optimal weights

- performance is tested against un-labelled data

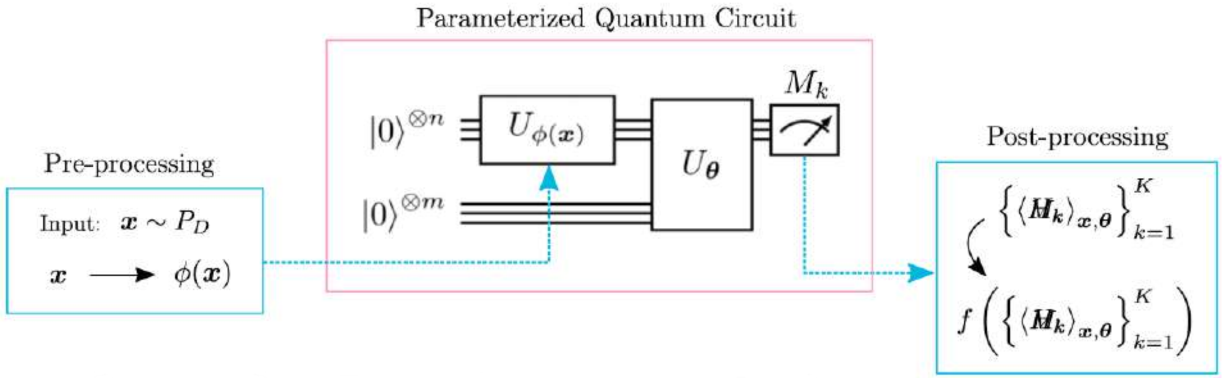
Quantum computing models for artificial neural networks



Mangini, Tacchino, Gerace, Bajoni, Macchiavello, EPL **134**, 10002 (2021)

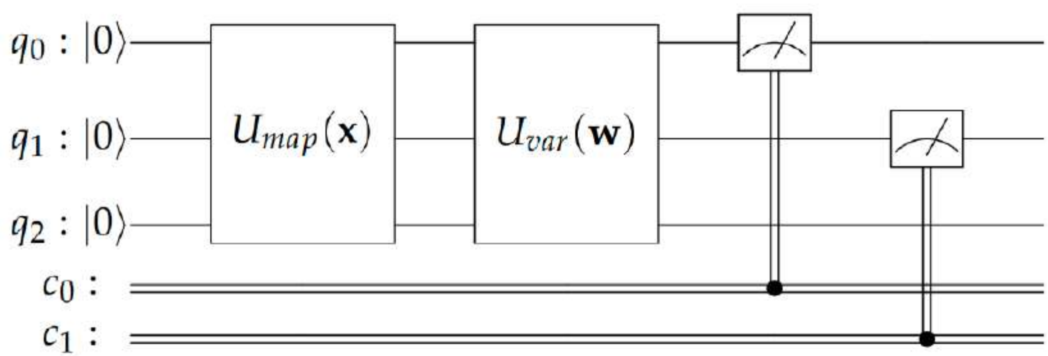
Quantum neural networks as parametrized quantum circuits

- in general: hybrid quantum/classical algorithms



Benedetti et al., Quant. Sci. Tech **4**, 043001 (2019)

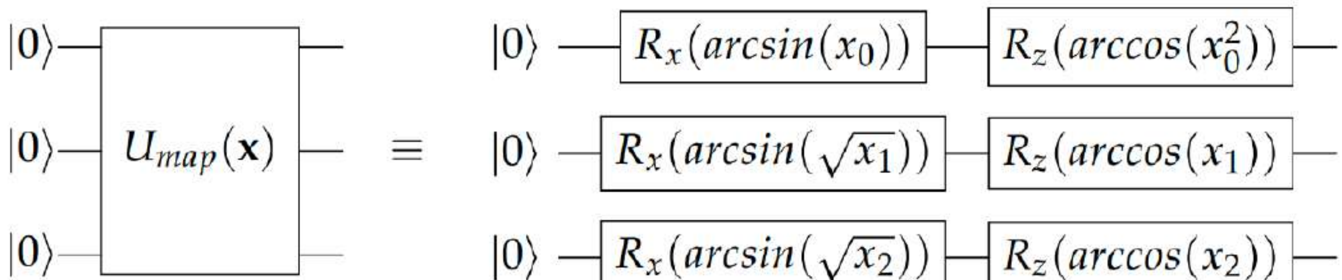
- specifically in this work



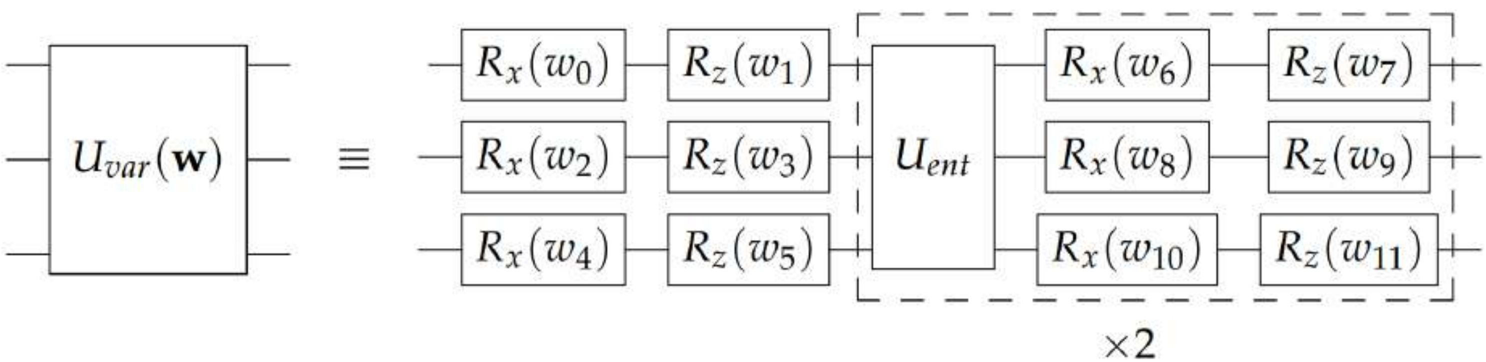
$\mathbf{x}$ → vector of input variables $\mathbf{w}$ → parameters optimized during training

Quantum circuits for VBS QML

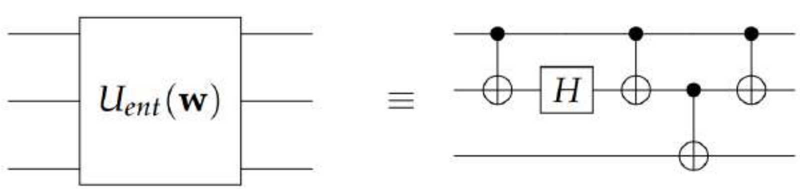
- encoding variables (each variable on a qubit)



- variational quantum operation



with



Deriving and minimizing the Loss f.

- from the qubits measurements to the classification function

e.g. on two qubits

$$p_{signal} \equiv e^{\bar{z}_0} / (e^{\bar{z}_0} + e^{\bar{z}_1})$$
$$p_{background} \equiv e^{\bar{z}_1} / (e^{\bar{z}_0} + e^{\bar{z}_1})$$

where $\bar{z}_0$ $\bar{z}_1$ are averaged over 8192 measurements on σ_z basis

- Loss function is built by the distance between the output function p_{signal} and the ideal classificator (giving results 1 for signal and 0 for background)

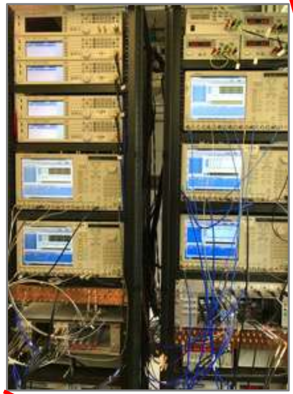
$$\boldsymbol{L}_w(\mathbf{x}, \theta, \mathbf{w}) = \sum_{i=1}^{N} P(x_i) \log \left(\frac{P(x_i)}{Q(x_i)} \right) w_i$$

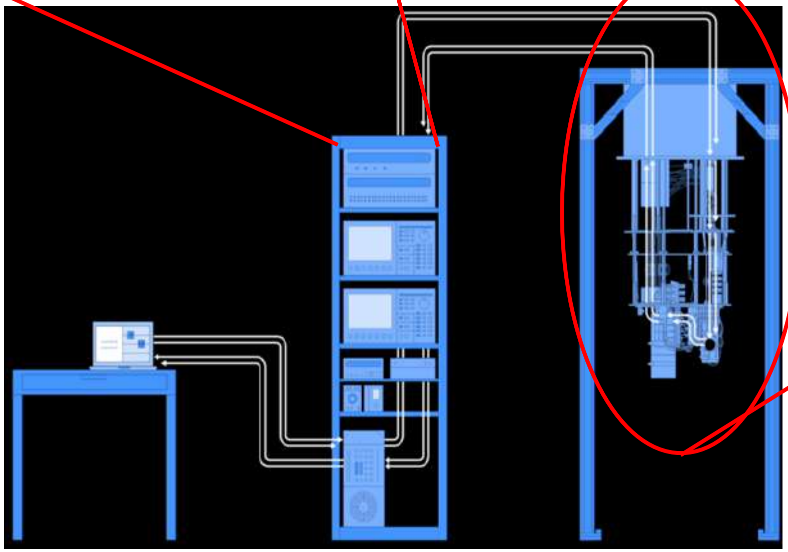
Generalized Kullback-Leibler (KL) divergence for N qubits

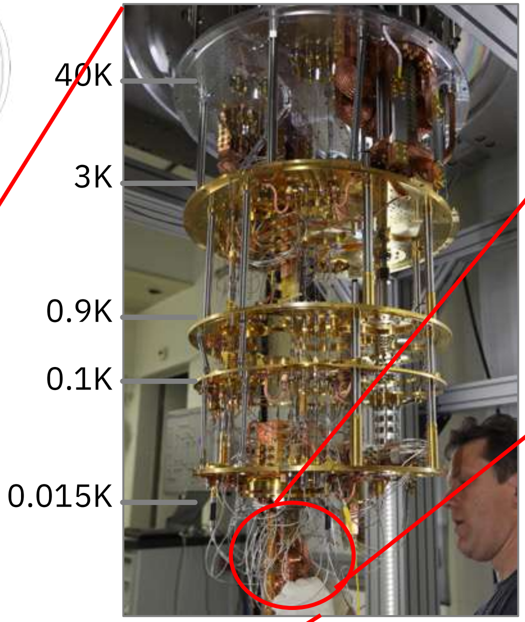
- Minimization allows to find $\mathbf{w}_{opt}$

IBM Quantum processors

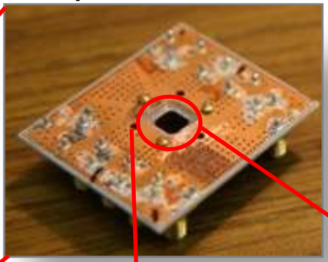
Microwave electronics

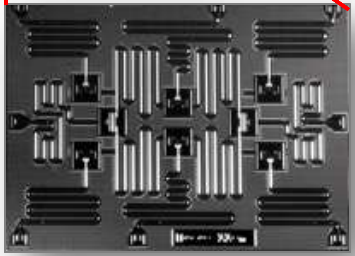




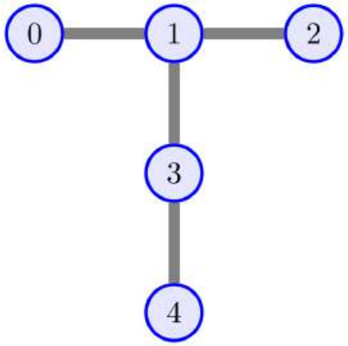


Chip Board

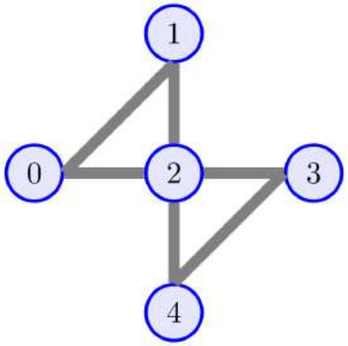




5-qubit NISQ devices →



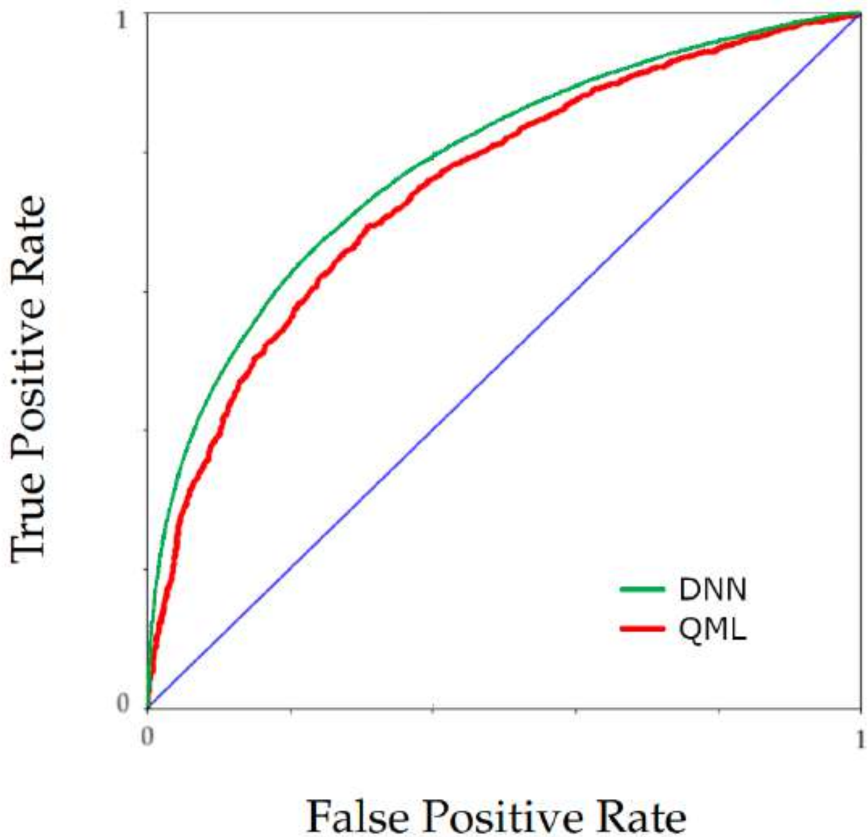
ibmq_belem



ibmq_5_yorktown

Results I: The Receiver Operating Characteristic (ROC) curve

- The classification performance is assessed by ...



$$\mathrm{TPR}=\frac{TP}{TP+FN} \qquad \mathrm{FPR}=\frac{FP}{FP+TN}$$

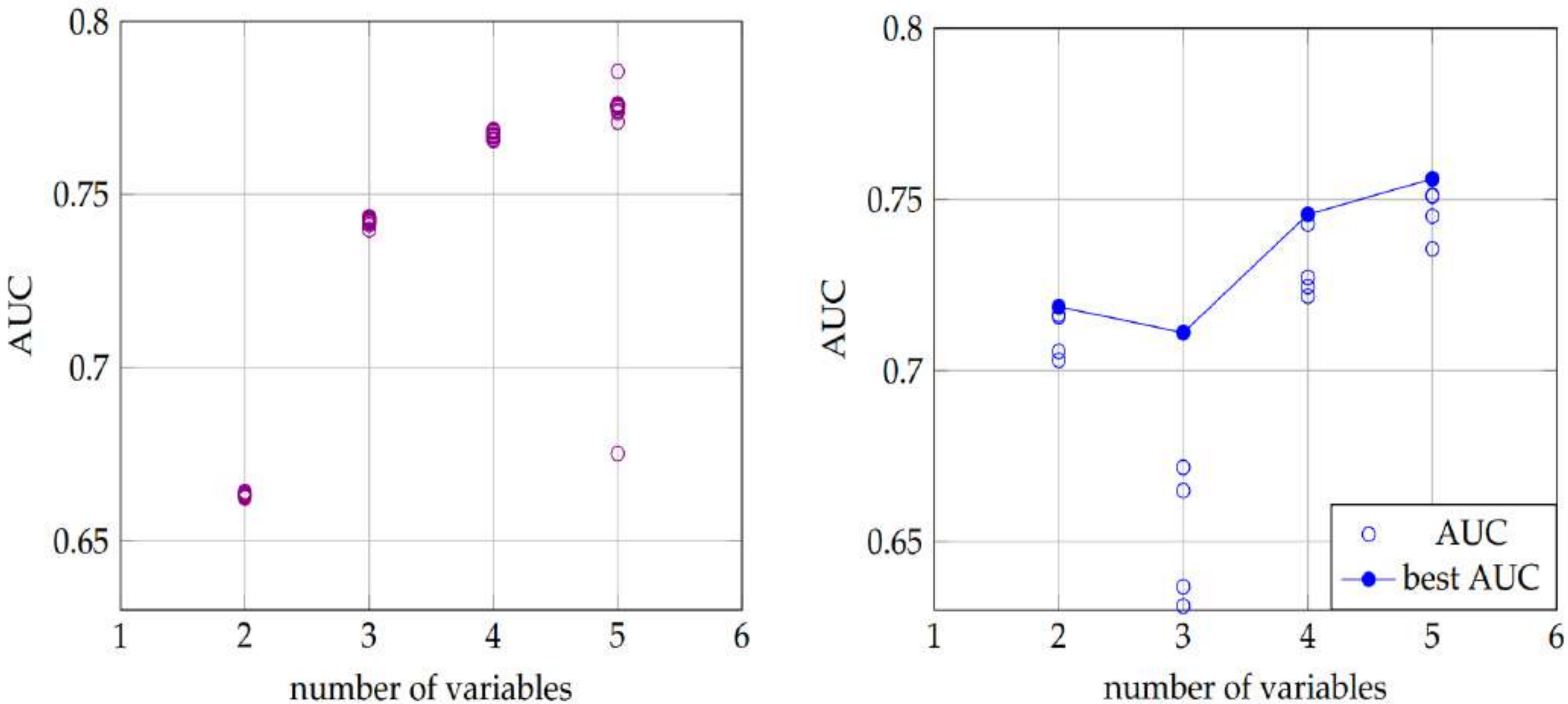
TPR: probability of positive event to be correctly classified

FPR: probability of negative event to be classified as positive

Cugini, Gerace, Govoni, Perego, Valsecchi, submitted (2022)

Results II: Area under the ROC curve (AUC)

Classical vs Quantum ML as a function of number of variables used in the classification

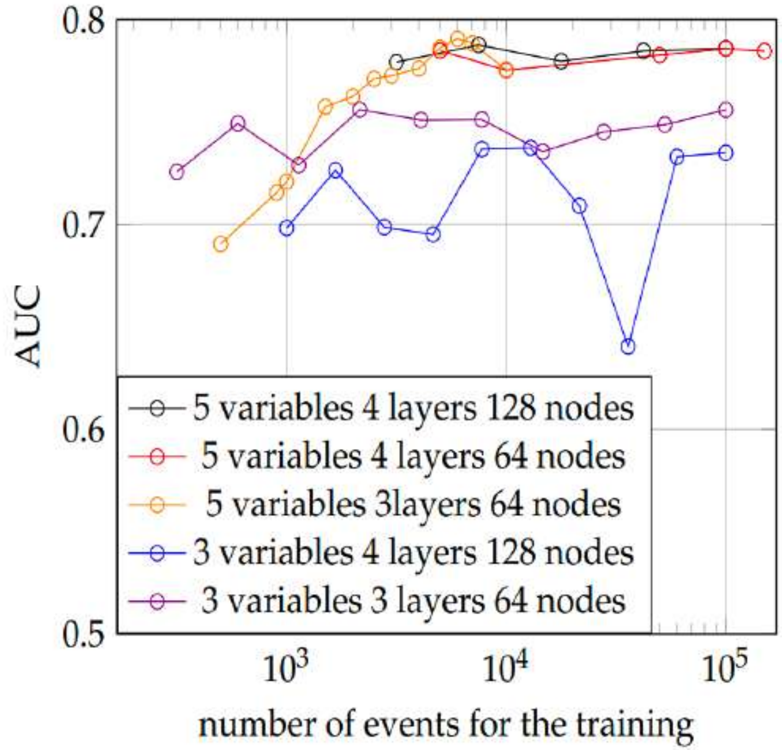


(a) Classical DNN (b) QML

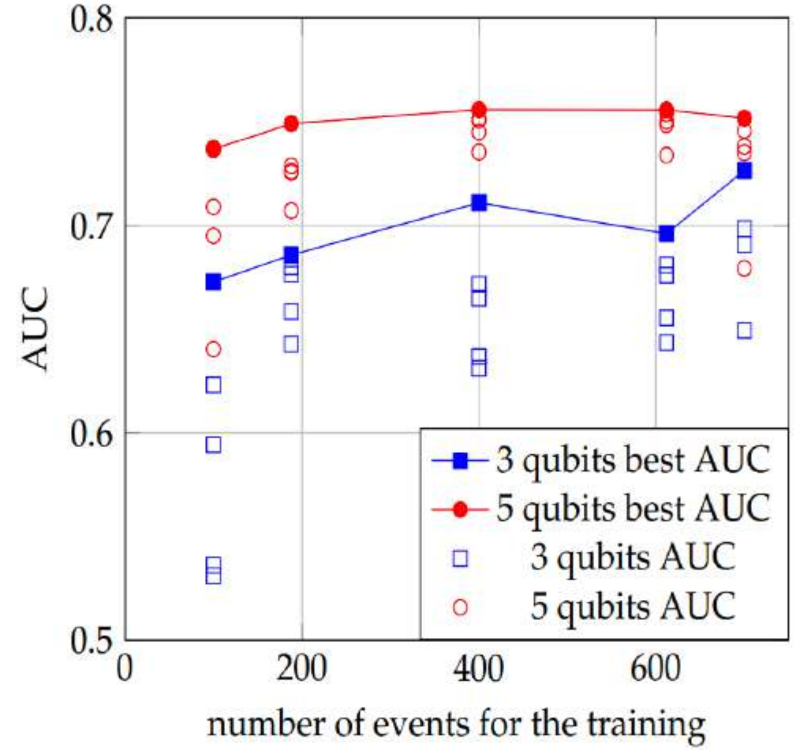
Cugini, Gerace, Govoni, Perego, Valsecchi, submitted (2022)

Results III: AUC vs number of traning events

Classical vs Quantum ML as a function of the number of events in the training set



(a) Classical DNN



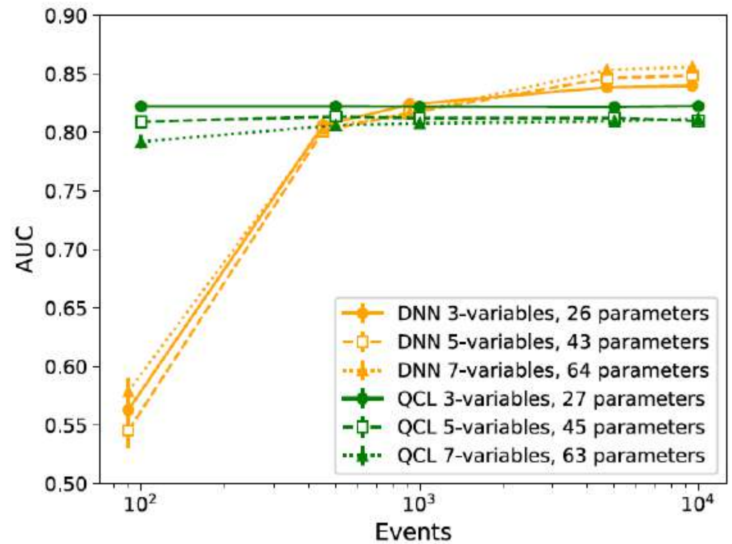
(b) QML

Cugini, Gerace, Govoni, Perego, Valsecchi, submitted (2022)

QML for HEP event reconstruction: Notable examples with similar conclusions

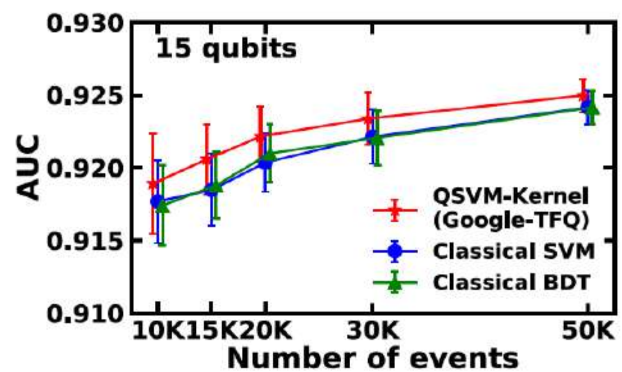
UNIVERSITÀ DI PAVIA

- Event classification of SUSY data set using PQC



Terashi et al.,
Comp. Soft. Big Sci. **5**, 2 (2021)

- QML study of background reduction for $t\bar{t}H$ process at LHC



Sau Lan Wu et al.,
Phys. Rev. Research **3**, 033221 (2021)

Summary

- QML via parametrized quantum circuits successfully applied to signal vs background classification for VBS simulated data from CMS

- Comparison between classical and quantum ML performances:

1) QML reaches comparable performance to DNN for a much smaller number of events in the training set, for the same number of variables;
2) A limited number of variables is sufficient to reach good classifier performances for QML algorithm

- Relevant for practical applications of NISQ devices in HEP, following recent trends within CERN that is massively interested in Quantum Computing applications

What next?

- Test on new IBM Q devices, as well as on different Quantum Computing platforms

- Training and test on the same quantum hardware

- QML on real CMS (or ATLAS) data sets

- Increasing number of qubits (variables over which training is performed) to compare classical and quantum ML in conditions where classical ML performs very well

Acknowledgments & Collaborations

UNIVERSITÀ DI PAVIA

Francesco Tacchino

IBM Research | Zurich

- IBM Quantum Experience is used through its freely available services

- The CMS collaboration is acknoweledged for data made available