Quantum Machine Learning Methods for Malware Detection

Eliška Krátká

Quantum Machine Learning Methods for Malware Detection

QUANTUM

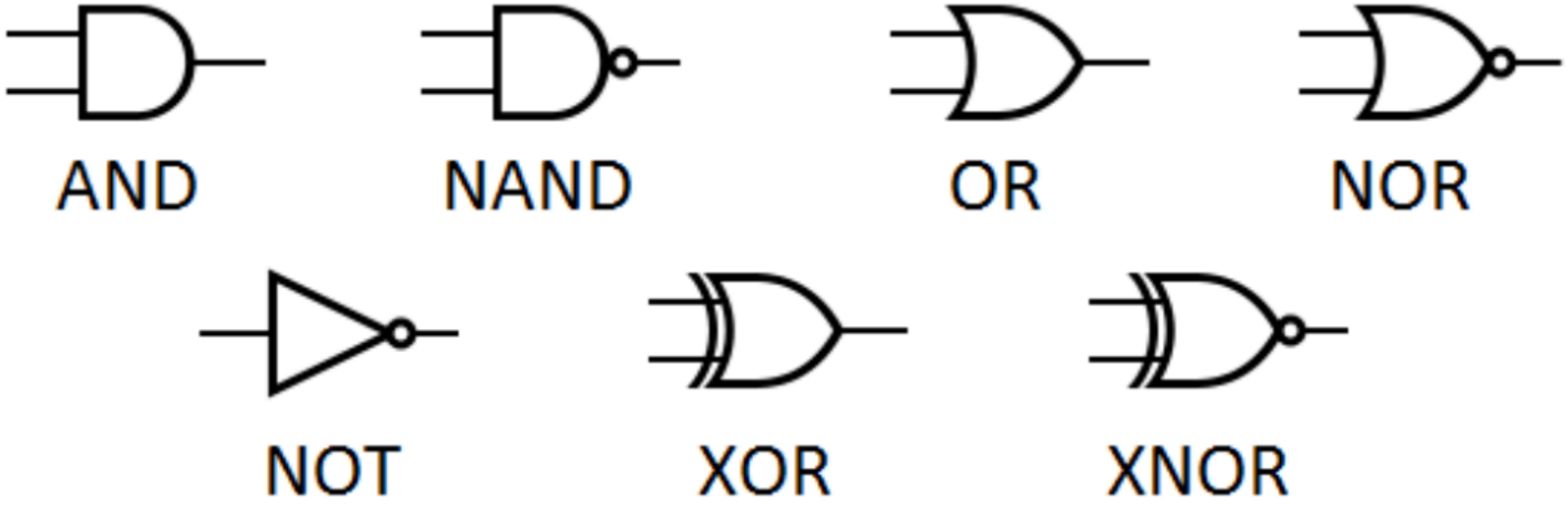
What is a quantum computer?

Just a different type of computer.

What is a classical computer?

A machine for processing information.

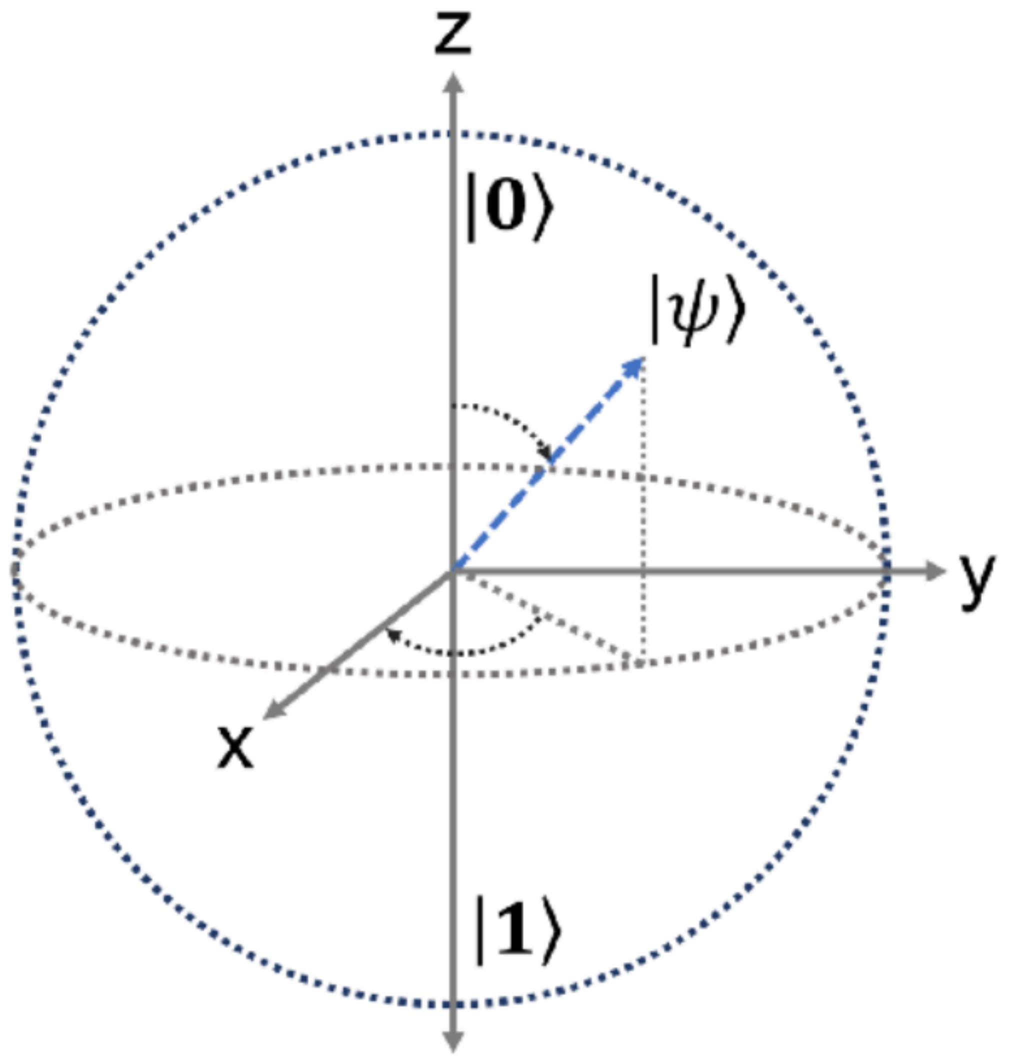
010011101100010001011
001001111010111101000
100010011010101110010
110000010100101010001
011011000011101111011
000100100100011101010
111001011011011010101

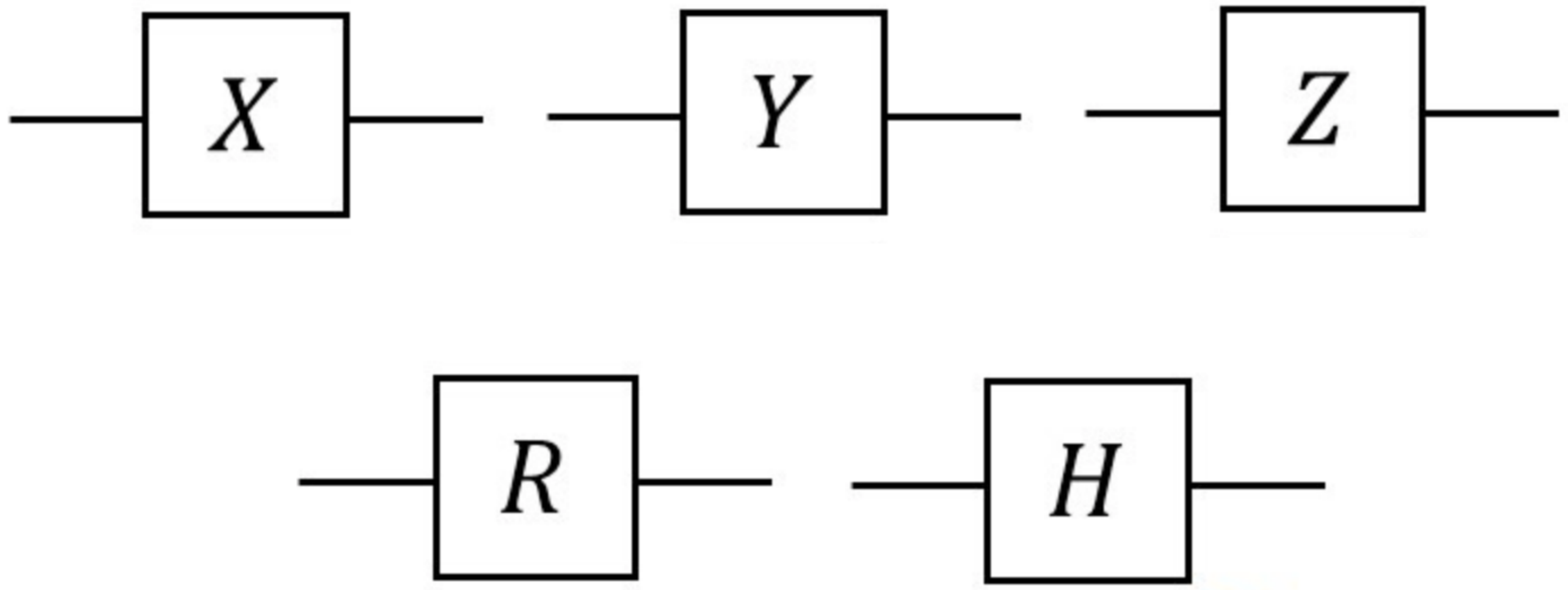


Computation has a predictable deterministic outcome.

Quantum Computer

A machine for processing information. ✅





Computation has a probabilistic outcome.
We only know the result after measurement.

Researchers develop web-tool to estimate when quantum computers outperform classical systems

Bitcoin's Quantum Countdown Has Already Begun, Warns Veteran Hacker

Quantum computers will break RSA!

Quantum computing and the future of cryptography: Understanding the imminent threat

China breaks RSA encryption with a quantum computer, threatening global data security

Quantum communication can be unhackable!

China Telecom Launches Hybrid Quantum-Safe Encryption System, Completes 1,000-Kilometer Secure Call

IBM Launches Its Most Advanced Quantum Computers, Fueling New Scientific Value and Progress towards Quantum Advantage

Quantum countdown has already begun!

The Quantum Apocalypse Is Coming. Be Very Afraid

Quantum computing will soon crack today's encryption methods. Here are 3 ways businesses can prepare

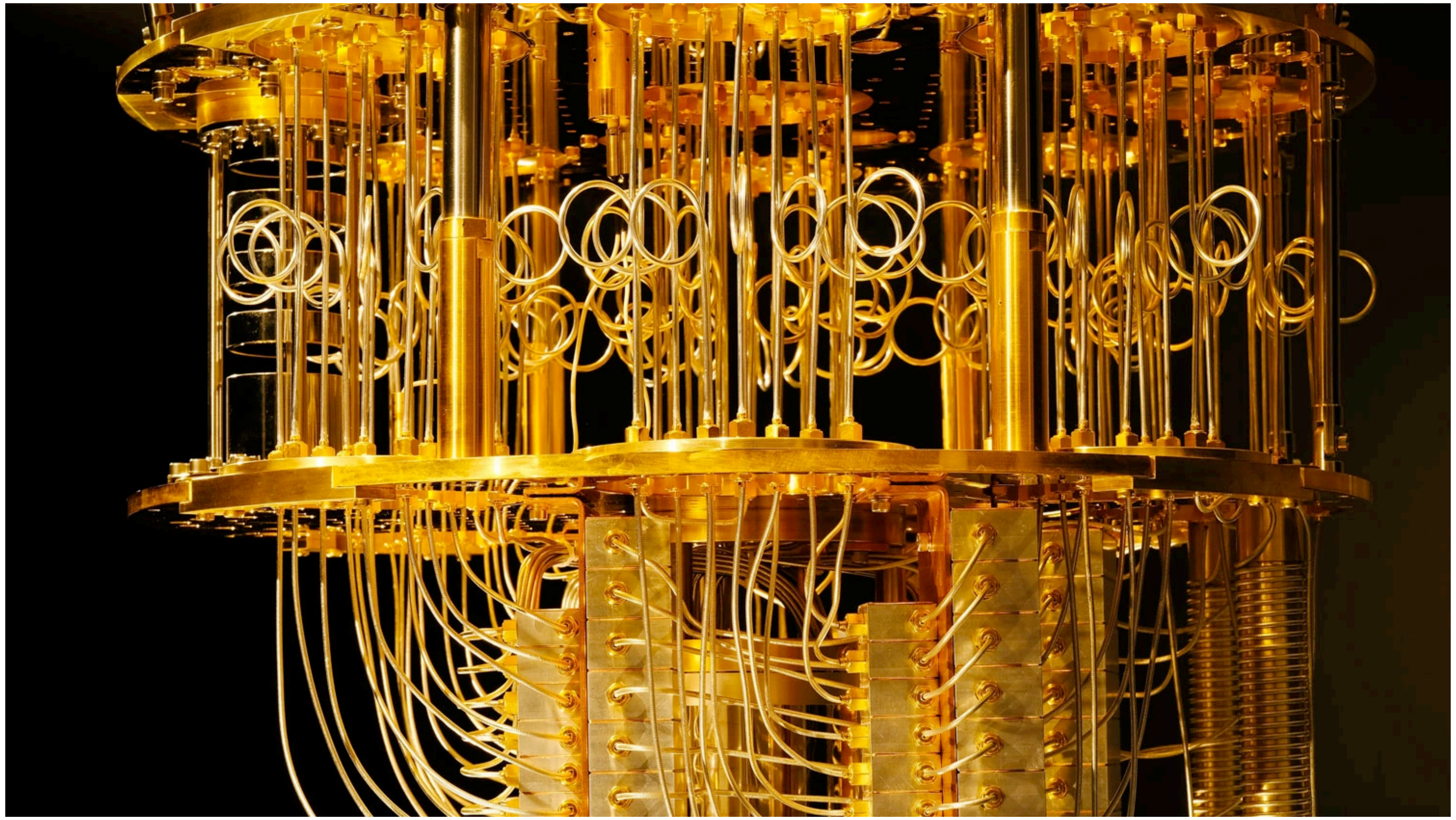
Quantum computing could arrive soon due to Microsoft's new chip

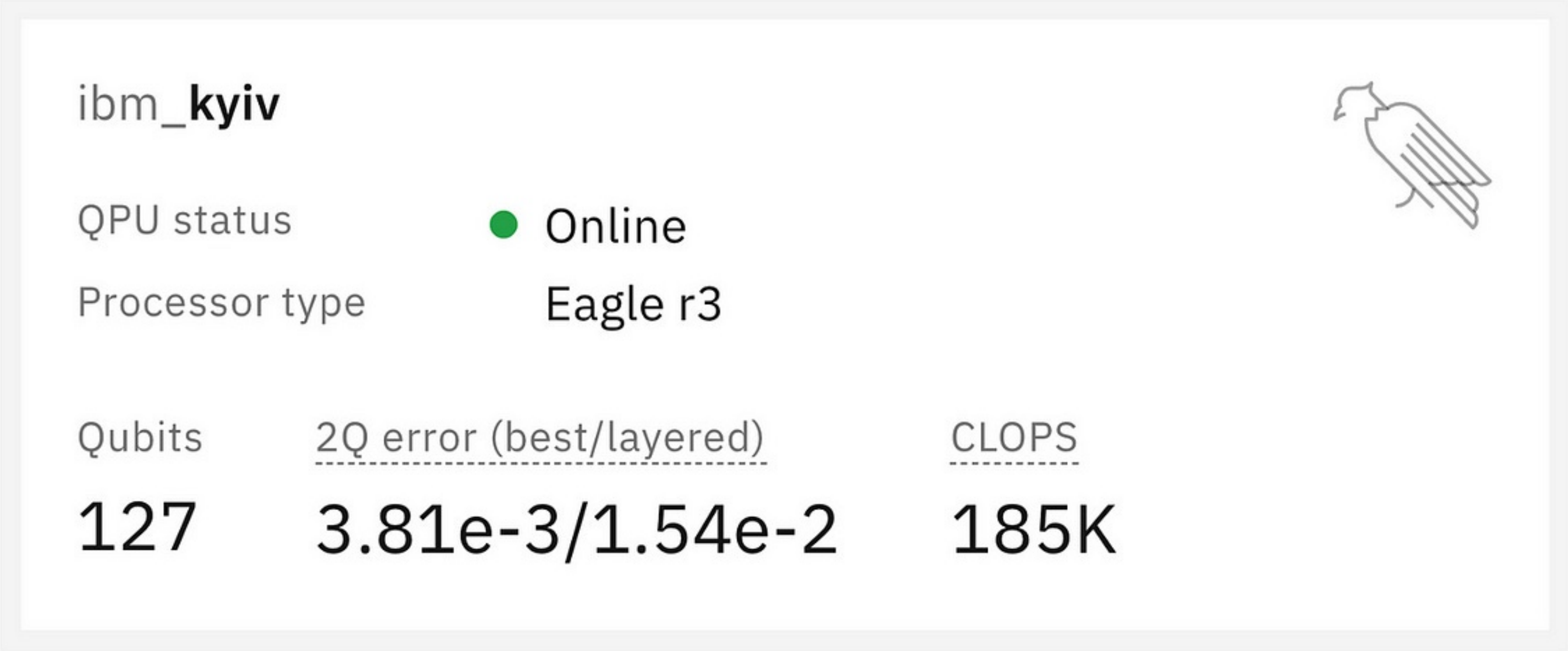
We are not there yet.

What do we have?

NISQ devices

(**N**oisy **I**ntermediate-**S**cale **Q**uantum)





ibm_**kyiv**

QPU status ● Online

Processor type Eagle r3

Qubits	2Q error (best/layered)	CLOPS
127	3.81e-3/1.54e-2	185K

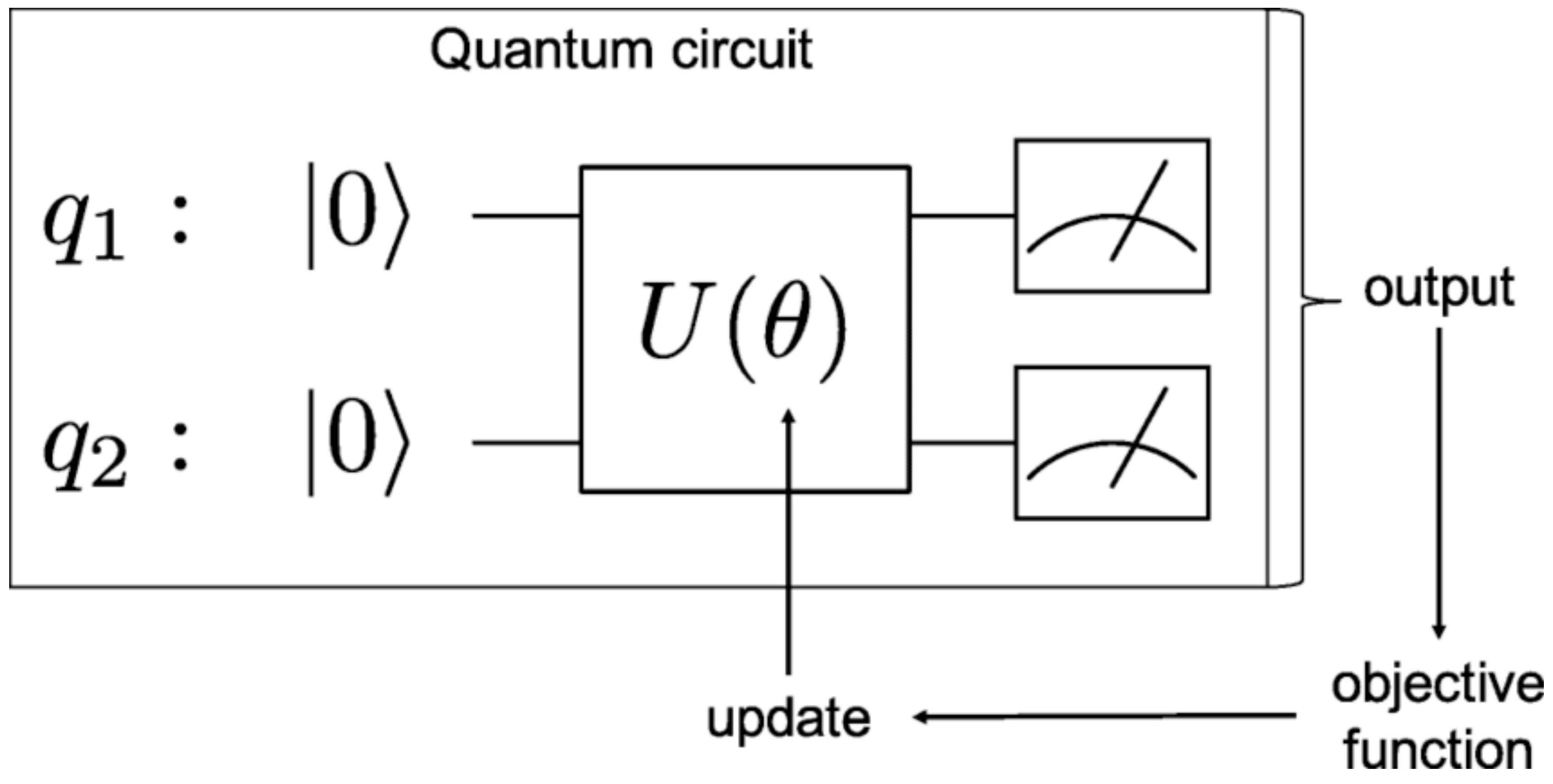
Quantum Machine Learning Methods for Malware Detection

QUANTUM ✅

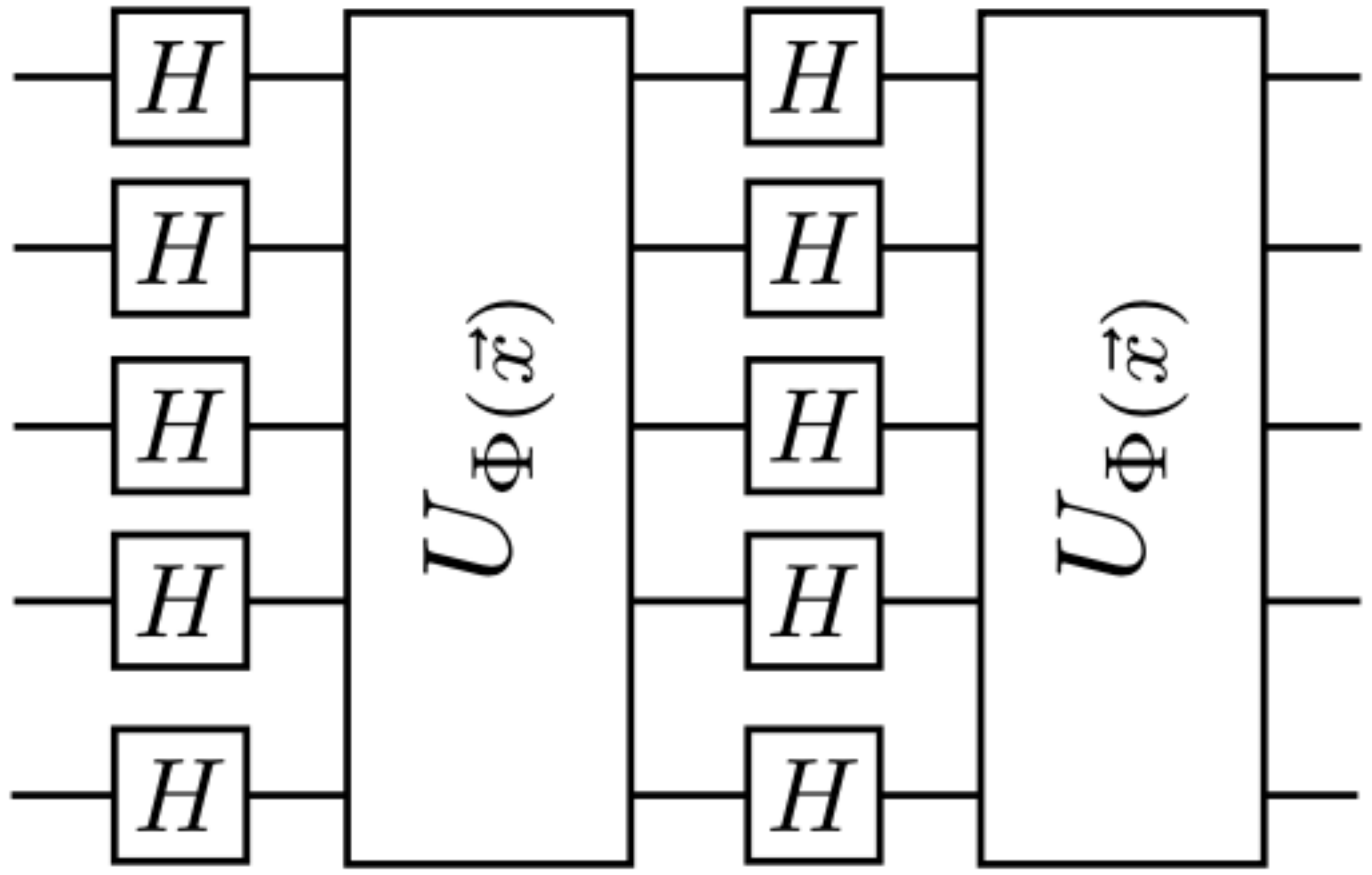
MACHINE LEARNING

Quantum Machine Learning

One of the most active research areas on NISQ hardware.



variational circuits

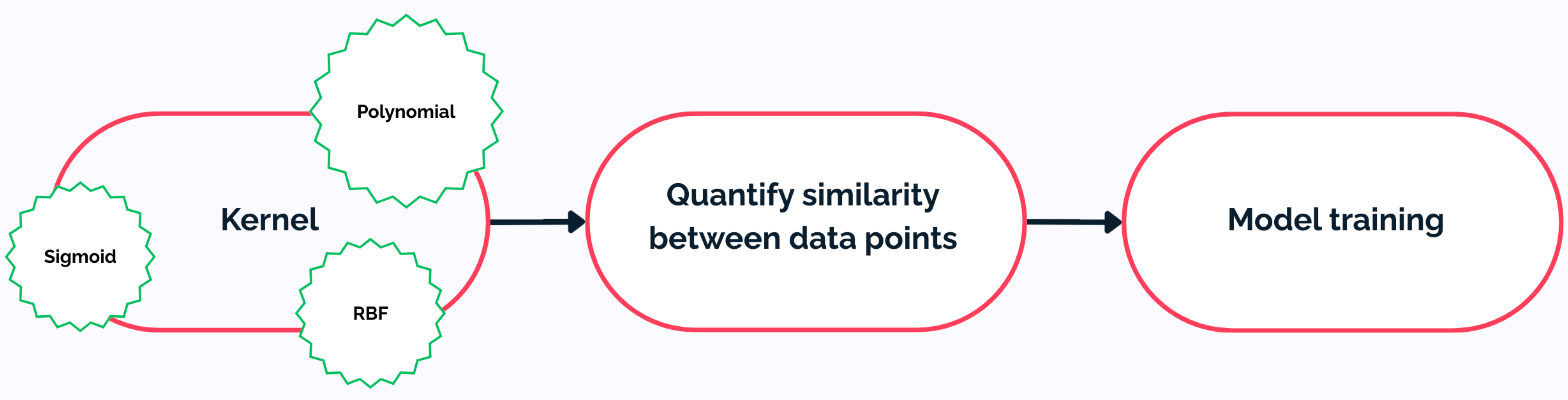


kernel methods

Usually **hybrid**: combines quantum and classical approaches.

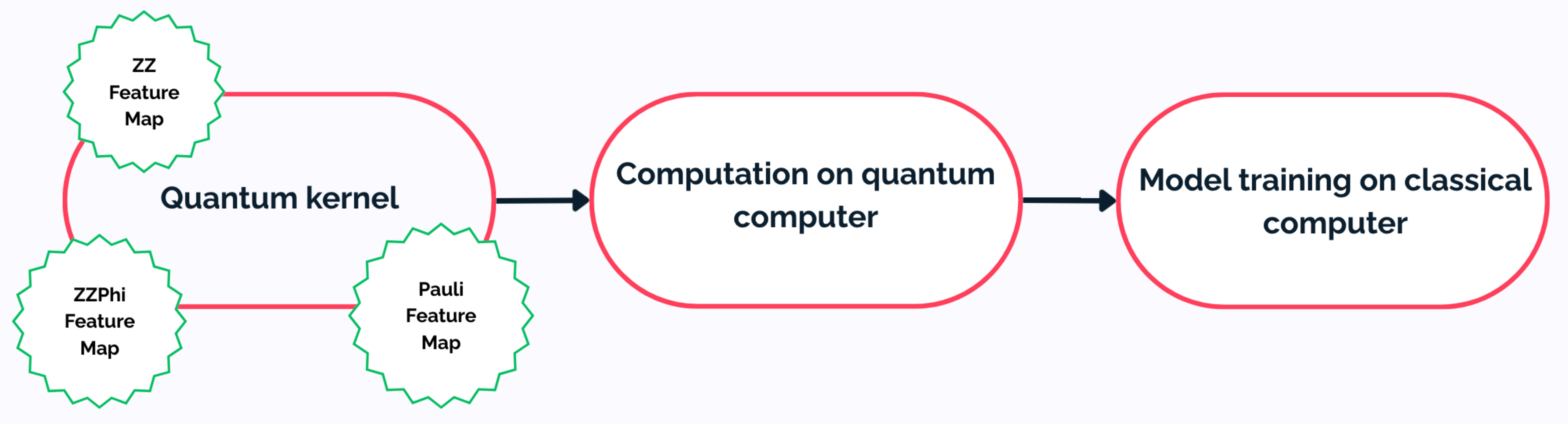
Support Vector Machine (SVM)

Kernel is a function which **measures similarity** between the data points.



Quantum SVM

Replace the classical kernel function with quantum computation.
Same SVM, just a new kind of similarity.



ZZ Feature Map
Quantum kernel
ZZPhi Feature Map
Pauli Feature Map
Computation on quantum computer
Model training on classical computer

Quantum Machine Learning Methods for Malware Detection

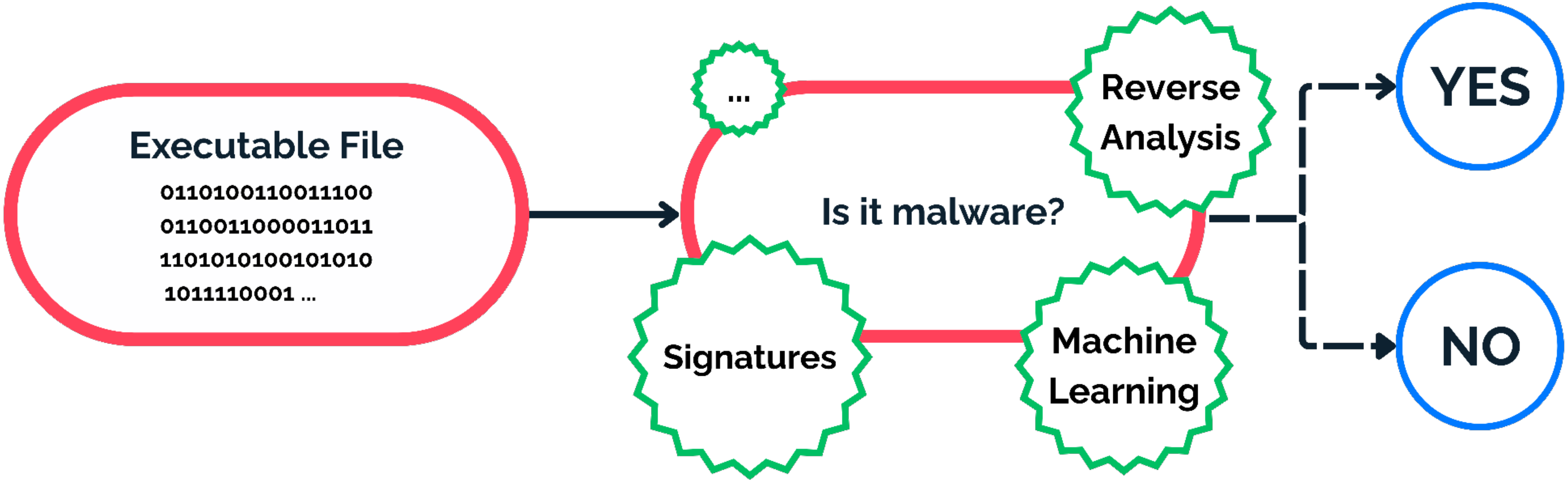
QUANTUM ✅

MACHINE LEARNING ✅

MALWARE

Malware Detection

Identify whether a sample is **malicious or harmless**.



Pipeline

1. Preprocess malware data

2. Compute quantum kernel

3. Train model

4. Evaluate performance

Why bother with quantum?

- Promising for **small**, high-dimensional datasets
- Quantum kernels could capture structures classical kernels can’t
- Useful for rare malware families, novel samples, anomalies?

Quantum Machine Learning Methods for Malware Detection

QUANTUM ✅

MACHINE LEARNING ✅

MALWARE ✅

Hacking quantum computers?

It’s already happening.

It’s not bug, it’s a feature.

It’s the software.

IBM

Qiskit

- Python SDK for quantum computing
- Open-source
- Originally developed and maintained by IBM
- Mainly a community project

Flaws

- Not production ready
- Overengineered but undermaintained
- Written by researchers, not software engineers
- Constantly changing API
- Inconsistent behaviour
- Unpredictable

Transpilation

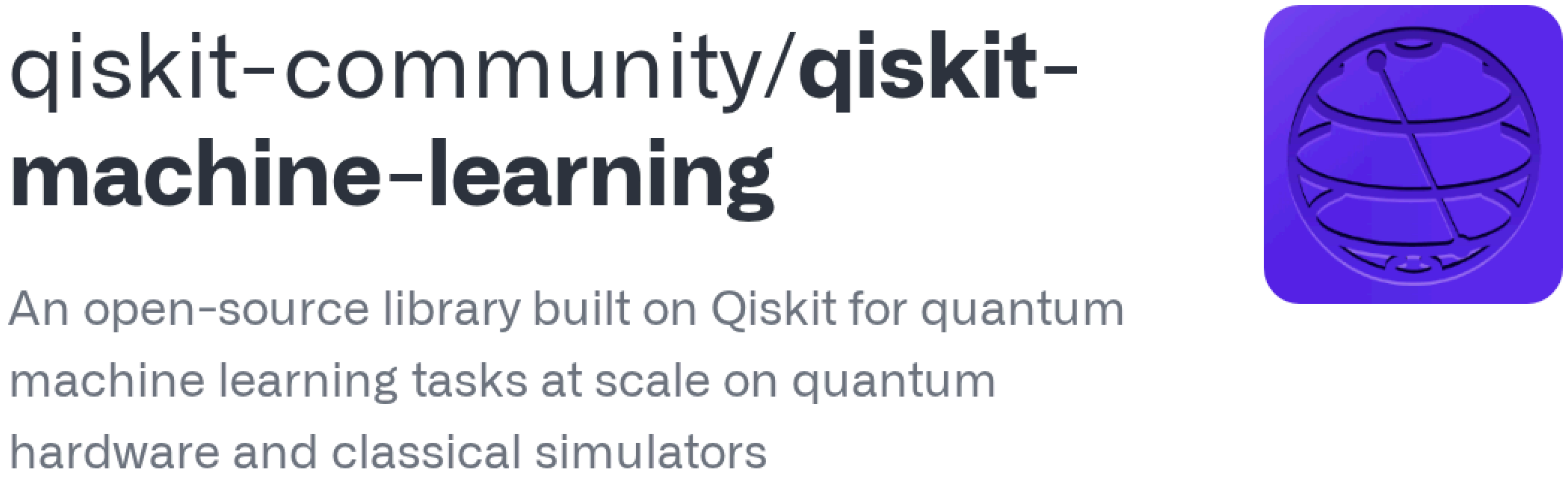
It’s "something like a compilation on a higher level".

Each hardware has unique constraints. Transpilation is needed to adapt circuits to the native gate set.

Problem? Random seed.

Modules maintained by third-parties

Problem? Only for simulators, unusable on real hardware.



qiskit-community/**qiskit-machine-learning**

An open-source library built on Qiskit for quantum machine learning tasks at scale on quantum hardware and classical simulators

IBM Quantum Platform

Provides access to the real hardware, in Qiskit through qiskit-ibm-runtime.

Problem? Many, my favorite are **blocking functions for computational jobs.**

What is malware?

Malware refers to any software ***intentionally*** *designed to cause a harm to a user.*

Malware often infects computer systems ***without the user's knowledge*** *or consent.*

average user

Hacking quantum computers?

Yes. By targeting developers, not qubits.

Weak spot is **people writing fragile Python code on top of billion-dollar machines**.

Average **physicist** will never suspect software being wrong because they already expect everything to behave randomly anyway.

Quantum Machine Learning Methods for Malware Detection

I’m the average user.

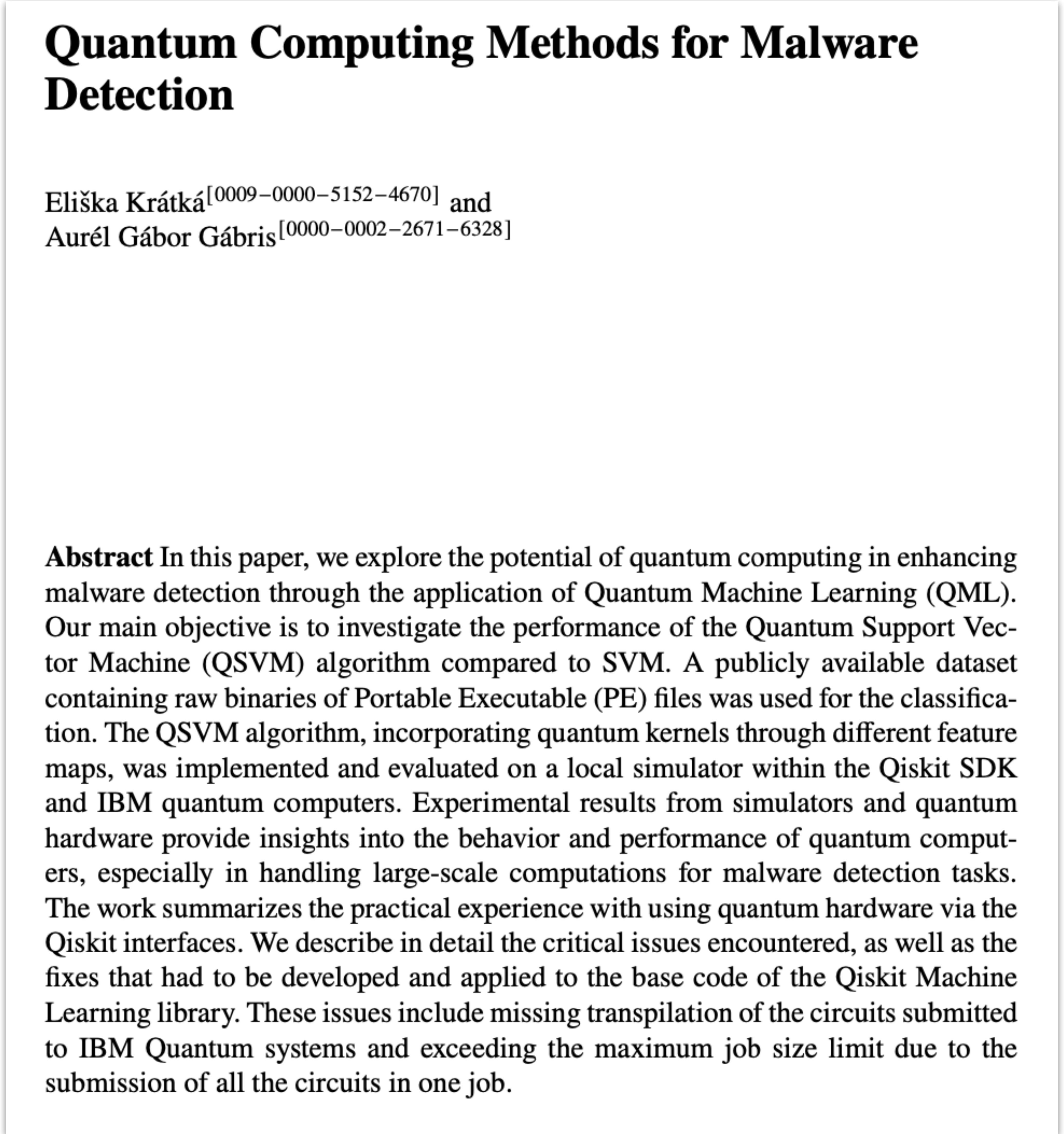
It runs on hardware.

Somehow.

We are limited by a **lot of stuff**.

So **not much** meaning for the practical use.

Yet?



Quantum Computing Methods for Malware Detection

Eliška Krátká[0009−0000−5152−4670] and
Aurél Gábor Gábris[0000−0002−2671−6328]

Abstract In this paper, we explore the potential of quantum computing in enhancing malware detection through the application of Quantum Machine Learning (QML). Our main objective is to investigate the performance of the Quantum Support Vector Machine (QSVM) algorithm compared to SVM. A publicly available dataset containing raw binaries of Portable Executable (PE) files was used for the classification. The QSVM algorithm, incorporating quantum kernels through different feature maps, was implemented and evaluated on a local simulator within the Qiskit SDK and IBM quantum computers. Experimental results from simulators and quantum hardware provide insights into the behavior and performance of quantum computers, especially in handling large-scale computations for malware detection tasks. The work summarizes the practical experience with using quantum hardware via the Qiskit interfaces. We describe in detail the critical issues encountered, as well as the fixes that had to be developed and applied to the base code of the Qiskit Machine Learning library. These issues include missing transpilation of the circuits submitted to IBM Quantum systems and exceeding the maximum job size limit due to the submission of all the circuits in one job.

eliska.kratka@fit.cvut.cz