UK Quantum Hackathon

Organised by the National Quantum Computing Centre

University of Edinburgh, 21st – 23rd July 2025

Challenge provider: Applied Quantum Computing

Quantum providers:

Emulator	Primary device	Backup device
Infleqtion	Infleqtion	Rigetti (via AWS)

Title: Quantum Machine Learning for Diabetes Identification

Background

Applied Quantum Computing Ltd is an independent quantum software company, aiming to provide quantum (or hybrid quantum/classical) software and associated consultancy advice specialising in healthcare.

AQC currently focuses on the application of quantum algorithms on current and near-term quantum devices, seeking to build (quantum) software solutions for healthcare-related problems.

Overview

Diabetes is relatively easy to diagnose, but a diagnosis is usually only carried out once the patient displays symptoms. This can cause delays in diagnosis since Type 1 diabetes may appear quickly, and early-stage type 2 diabetes can remain asymptomatic for some time.

Early diagnosis of diabetes can help prevent the risk or severity of long-term complications such as eye, kidney or nerve damage, or cardiovascular disease. Early treatment or lifestyle changes can help prevent prediabetes developing into type 2 diabetes, or reduce the likelihood of sudden hospitalisations in type 2 diabetes. Furthermore an early diagnosis can minimise the likelihood of undiagnosed patients presenting symptoms in a very advanced stage. The ability to detect the presence of diabetes before symptoms appear would improve health outcomes for many.

From an economic viewpoint, it is more cost-effective to manage diabetes preventatively, rather than reactively. Treating diabetes and its related complications is expensive, and early management could reduce costly hospital admissions, dialysis treatment and surgical procedures.

Software which can perform an early diagnosis, using indicators which can be taken as part of routine blood tests and medical examinations, would be of significant interest to the healthcare sector.

Technical description

This is a supervised learning / classification problem [1,2]. The team will be provided with a dataset of 1,000 patients together with some medical information about each patient as outlined below.

Each patient is either not diabetic, prediabetic or diabetic. The dataset contains this information.

The overall task is to be able to predict whether a patient is not diabetic, prediabetic or diabetic based only upon the 11 medical indicators. A quantum support vector machine [3] or quantum neural network [4] could be used to attempt the prediction.

Multi-class classification problems are not so straightforward for quantum algorithms. A more tractable problem may be to perform a related binary classification, for instance non-diabetic compared to either prediabetic or diabetic.

The team should isolate a set of training data and test data from the data provided. Note that limitations of quantum hardware or simulators may mean that the team may have to use a smaller subset of the data. Success can be measured based upon how well the prediction works on the test data, using a metric such as the weighted f1 score.

It would be interesting to compare how the quantum machine learning approach performs compared to an equivalent classical machine learning algorithm. The team should note, however, that we do not necessarily expect the quantum machine learning algorithm to outperform the classical one.

Two stretch targets for the team:
• it would be interesting to investigate whether a subset of the medical information can be used to perform the machine learning task, and whether this provides different relative performance for the quantum or classical predictions.
• It would be interesting to see whether there is a difference in relative performance between the quantum and classical classifiers if different sized training sets are used (for example, compare the performance of the classifiers if the training set consists of 700 patients, 500 patients or 300 patients, with the same test set).

Data

AQC will provide an anonymised data set of around 1,000 patients each providing 11 medical figures:
1. Gender
2. Age
3. Urea level in the blood
4. Creatin level in the blood
5. Glycated haemoglobin, a measure of blood-sugar levels
6. Cholesterol level in the blood
7. Triglycerides (a type of fat) level in the blood
8. High-density lipoprotein cholesterol level (“good” cholesterol)
9. Low-density lipoprotein cholesterol level (“bad” cholesterol)
10. Very low-density lipoprotein cholesterol level
11. Body Mass Index

Each patient is marked as being either non-diabetic, prediabetic or diabetic.

The data is publicly available and anonymous.

Anticipated impact

Early identification of diabetes, at the pre-symptomatic stage, can produce significantly better health outcomes for many people. In the UK it is estimated that over 5 million people in the UK have diabetes, and this number is expected to increase because of the ageing population, sedentary

lifestyles and obesity. A large proportion (perhaps one in three [5]) is expected to remain undiagnosed, so early diagnosis through blood screening could improve many lives.

Furthermore, the NHS and other governments could generate significant cost savings by preventing hospitalisations and treatment, and instead prevent the disease from developing which is more cost-effective.

References

[1] https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained
[2] Qiskit Machine Learning tutorials: https://qiskit-community.github.io/qiskit-machine-learning/tutorials/index.html
[3] Tutorial on quantum support vector machines from Medium: https://medium.com/mit-6-s089-intro-to-quantum-computing/quantum-support-vector-machine-qsvm-134eff6c9d3b
[4] A Abbas, D Sutter et al, “The Power Of Quantum Neural Networks” https://arxiv.org/pdf/2011.00027
[5] Office For National Statistics: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthinequalities/bulletins/riskfactorsforprediabetesandundiagnosedtype2diabetesinengland/2013to2019