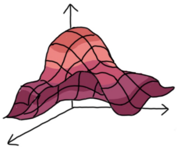
Quantum optimization for K-nearest neighbours and image classification.

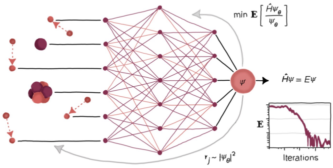
Aleksandra Mulewicz | Downside School

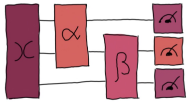
MAIN STUDY

Through subatomic manipulation of particles, we can conduct highly targeted computations, enabling tasks like dataset searches at a pace up to quadratically faster. Moreover, this technology prompts questions about its potential in enhancing machine learning models which is the subject of this study.



Different approaches to increase the efficiency include such techniques as quantum kernels - gaussian and polynomial, or frameworks based on Schrödinger's equation. These frameworks map the problem and then resolve it by finding the ground state of the hamiltonian (the system's total energy) - it's the most stable state as energy is the lowest. This study focuses on oracle function, amplitude estimation and hamming distance.



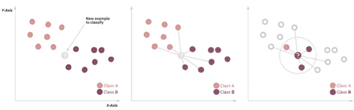


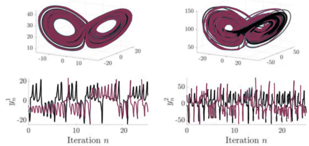
MODEL COMPARISON

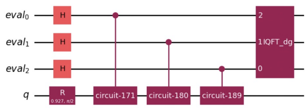
K-nearest neighbours is an algorithm based on the idea of the similarity. The fundamental principle behind this model is to estimate the label of a data point by considering its closest neighbours. It plots information from the dataset provided - where characteristics influence the positions of these points.

The image needs to be preprocessed to represent pixels in a quantum state - in this study Karhunen-Loève transform (PCA) was used. This method is based on manipulating the statistical properties of the image for further compression and feature extraction. This operation results in a high correlation between the pixels and fascilities their further manipulation into multidimensional matrices.The target quantum state is then defined, and the suitability of the approximaion with the target is is assesed using Kullback-Leibler divergence.

The quantum amplitude estimation algorithm is used as a similarity measure to quantify and determine the resemblence between data. It is composed of Quantum Phase Estimation and Grover's search which utilises oracle function to mark the desired state by flipping the sign and amplitude amplification - it maximises the solutions amplitude, minimising the rest.

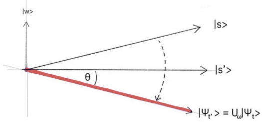


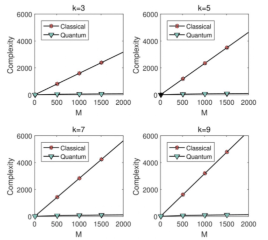


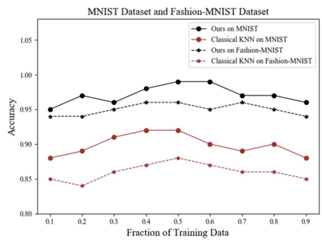


The next step involves finding the value of k with the DHA (Dürr-Hoyer minimum). This process involves finding the theoretical lower bound limit - the minimum number of points (k) required to solve the problem. It repeats a process of finding the smallest index with Grover's search until the probability of this point being the minimum is relatively big.

After this step is completed the similarity is computed with Hamming distance. To manipulate data, it needs to be mapped into binary vectors using a hash function - it takes an input and creates the hash for further operations. This program creates a unique series of numbers to store the pixels as 0s and 1s before comparing them with hamming distance. This method simply outputs the number of bits that are different between two points. For instance if one of them is 0111 and the other 0100, two will be returned.







CONCLUSIONS

The biggest advantage of these algorithms is proved to be manipulating quantum or strictly physical data. However, it can be seen that even for some classical datasets these models are able to improve the efficiency. Even though the quantum computers we actually possess at the moment are not of the power that could contribute to complete changes in the classification we know, the approach itself was proved to work very well.

The main problem with quantum computers is an environment needed for qubits not to lose their quantum state - it cannot be susceptible to light and heat. The entire system - the environment and machine itself, is very expensive and only specific research centres can afford the actual quantum computer. Another problem is the quantum noise which will always be present around the algorithm - however it can be decreased with help of advanced mathematical concepts.

References

Abbas, Amira, and IBM Quantum, IBM Research – Zurich. n.d. "The power of quantum neural networks."

Durr, Christoph. n.d. "A quantum algorithm for finding the minimum."

Gordenstein, Harry. "Quantum KNN. Classical KNN | by Harry Gordenstein."

Zhou, Nanrun. 2021. "Quantum K-Nearest-Neighbour Image Classification Algorithm Based on K-L Transform."

Giurgica-Tiron, Tudor. 2022. "Low depth algorithms for quantum amplitude estimation." arXiv.

Ruan, Yue. 2019. "Quantum Algorithm for K-Nearest Neighbors Classification Based on the Metric of Hamming Distance." Researchgate.

Dang, Yijie. n.d. "Image Classification Based on Quantum KNN Algorithm." arXiv