Quantum Machine Learning for Big Data Analytics

A thesis submitted

in partial fulfillment for the award of the degree of

Doctor of Philosophy

by

Avinash Chalumuri

Department of Avionics
Indian Institute of Space Science and Technology
Thiruvananthapuram, India

October 2022

Abstract

Enormous amount of data is generated globally at an exponential rate in various sectors such as government, agriculture, finance, defense, engineering, and medicine. Due to the recent advancements in technology, a rapid growth rate in data generation and collection of different varieties of data is observed. The exponentially growing data that can be analyzed computationally to reveal patterns, trends, and associations are considered as *big data*. Numerous benefits can be obtained from analyzing such *big data* that helps in decision-making and finding solutions for various real-time problems. Over the years, Artificial Intelligence (AI)/Machine learning (ML) algorithms revolutionized the way of analyzing vast amounts of data for valuable insights and applications. Machine learning emphasizes developing computer programs that can access and learn from data to build a generalized model. Supervised learning is an efficient learning technique to build ML models using "labeled" training data to predict the output. The model parameters such as weights are optimized to produce the desired result during the training process. Also, hyperparameters are tuned during the training phase to improve the accuracy of the model.

Due to the rapid speeds at which the data grows, big data processing using ML algorithms is an area of concern. ML algorithms face many challenges in dealing with big data, including computational resources, model selection, optimizing the parameters, and increasing the algorithm accuracy. Deep learning algorithms such as convolutional neural networks (CNNs) require huge computational facilities to process very-large datasets for supervised learning. Also, difficulty in training such deep learning models increases with the large and complex datasets. In this context, Quantum Machine Learning (QML) is emerging as a field of interest in computer science with the intersection of quantum computing and machine learning. Quantum computers are fundamentally different from classical computers as principles of quantum mechanics are used for information processing. Hence, quantum computing techniques can solve specific computational problems difficult for a classical computer. Also, quantum computing can enhance classical machine learning techniques as powerful quantum tools exist for linear algebra. As linear algebra is the basis for machine learning, quantum computing offers practical performance advantages over classical approaches. Hence, there is a necessity to explore the area of quantum machine learning, to advance the existing machine learning techniques.

In this thesis, we propose and study quantum machine learning techniques to enhance supervised learning of classical data. The ability of quantum computational approaches to improve classical machine learning algorithms is explored. In particular, the proposal and study of hybrid quantum-classical machine learning methods to solve supervised learning problems are addressed in the thesis. As a first step towards quantum machine learning, an artificial neural network (ANN) using quantum bits (qubits) as artificial neurons is proposed for binary classification. The quantum computing approach for ANN (QC ANN) aims to develop a clear understanding of the impact of qubits in training an artificial neural network for binary classification of numerical data. Further, the work is extended to design a quantum multi-class classifier (QMCC) for multi-class classification. QMCC is intended to be a quantum circuit with parameterized quantum layers for machine learning. For QMCC, an encoding method for state preparation to input the data into qubits is also proposed. QMCC, in total, is a parameterized quantum circuit with multiple trainable layers for the multi-class classification of numerical data. The experimental results show that the proposed techniques perform binary and multi-class classification with good accuracy.

In the second phase of our research, we proposed QML frameworks for processing spatial big data analytics tasks. At first, a three-layered hybrid quantum-classical (HybridQC) architecture is proposed for satellite image scene classification. The proposed model consists of the following steps: (i) a classical preprocessing step, (ii) a quantum processing step to extract image representations, and (iii) a deep neural network built with the extracted image representations. Our experimental results show that the total parameters for training a deep neural network reduced with the proposed approach. Next, a data augmentation technique is proposed using a quantum circuit that can be used to enhance datasets during training deep neural networks. Also, a hybrid model with a combination of vanilla convolutional neural network (CNN) and quantum processing is proposed for image scene classification. Finally, we propose quantum processing techniques to process synthetic aperture radar (SAR) images for deep learning.

We also discuss the performance advantages of the hybrid quantum-classical approach over classical computation on both numerical data and spatial data. Our work show that quantum computational techniques can enhance classical machine learning by reducing the trainable parameters of the models and, as a result, improvement in the classification accuracy can also be observed. Finally, the thesis is concluded with the future scope of QML algorithms in solving complex machine learning problems.