Arshi Jamal (May 2021). Unlocking the Power of Quantum Computing for Big Data Processing and Advanced Analytics
International Journal of Economic Perspectives,*15*(5) 48-60
Retrieved from https://ijeponline.com/index.php/journal

Unlocking the Power of Quantum Computing for Big Data Processing and Advanced Analytics

Arshi Jamal, Asst. Professor
Dept. Computer Science
Govt. First Grade College, Raichur-584101

Abstract
In the era of data-driven decision-making, the exponential growth of data has posed unprecedented challenges for traditional computing methods. Big Data, characterized by vast volumes, high velocity, and diverse formats, demands innovative approaches to processing and analytics. Quantum computing, a disruptive technology harnessing the principles of quantum mechanics, offers a promising solution to tackle these challenges.This research provides a concise overview of the potential of quantum computing in the context of Big Data processing and analytics. We explore how quantum computing's unique properties, such as superposition and entanglement, can revolutionize data handling and analysis. Quantum algorithms, like Grover's and Shor's algorithms, are introduced as powerful tools for searching and factoring large datasets, respectively, with remarkable speedups compared to classical counterparts. Quantum machine learning algorithms, such as quantum support vector machines and quantum neural networks, demonstrate the potential to unlock insights from vast datasets more efficiently than classical counterparts. Quantum computing also promises to enhance data security through quantum-resistant encryption schemes, addressing the vulnerabilities of classical cryptography.

Keywords:-Quantum Computing,Analytics,Quantum Big Data,Data Analytics

Introduction

The explosion of data in today's digital age has ushered in an era where traditional computing methods are struggling to keep up with the demands of Big Data processing and analytics. This abstract delves into the potential of quantum computing as a game-changing technology to address these challenges.Quantum computing, grounded in the principles of quantum mechanics, offers unique capabilities that can revolutionize how we handle and analyze massive datasets. One of the most intriguing features of quantum computing is superposition, which allows quantum bits (qubits) to exist in multiple states simultaneously. This property can exponentially increase computational capacity, enabling the rapid processing of vast volumes of data.Entanglement, another fundamental quantum property, enables qubits to become interconnected, even when physically separated. This interconnectedness can facilitate parallel processing, enhancing the speed and efficiency of data analysis tasks.

Corresponding author **Arshi Jamal**

Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

Quantum algorithms, such as Grover's algorithm for search and Shor's algorithm for factoring large numbers, demonstrate remarkable speedups compared to their classical counterparts. These algorithms hold the promise of significantly reducing the time required for searching through vast datasets or breaking conventional encryption methods. Quantum machine learning algorithms, like quantum support vector machines and quantum neural networks, show great potential for optimizing data analytics tasks. These algorithms can unlock insights from Big Data more efficiently, potentially transforming industries ranging from finance to healthcare.While quantum computing offers exciting possibilities, it also faces challenges, including the need for error mitigation and scalability. Researchers are actively working on developing robust quantum hardware and error-correcting codes to overcome these hurdles. The fusion of quantum computing and Big Data processing and analytics presents a transformative frontier. This abstract provides a glimpse into how quantum computing's unique properties can empower businesses and researchers to process and analyze massive datasets more efficiently, ultimately driving innovation and informed decision-making in the data-driven world.

Quantum computing

Quantum computing is a revolutionary paradigm of computation that leverages the principles of quantum mechanics to perform calculations at speeds and scales unattainable by classical computers. Unlike classical bits, which can represent either a 0 or a 1, quantum bits or qubits can exist in multiple states simultaneously, thanks to a phenomenon known as superposition. Additionally, qubits can be entangled, meaning the state of one qubit is interconnected with the state of another, regardless of the physical distance between them.

These quantum properties enable quantum computers to solve certain problems much more efficiently than classical computers. For instance, quantum algorithms like Shor's algorithm can factor large numbers exponentially faster, which has implications for breaking classical encryption methods, while Grover's algorithm can accelerate database searches.

Quantum computing has the potential to transform various fields, including cryptography, materials science, drug discovery, optimization, and artificial intelligence. However, building and maintaining stable quantum computers is an enormous scientific and engineering challenge, as qubits are highly susceptible to noise and environmental interference.

Researchers and companies worldwide are actively working on developing practical quantum computers and exploring quantum algorithms' applications. As the field advances, quantum computing holds the promise of ushering in a new era of computing that could revolutionize industries and solve complex problems currently beyond the reach of classical computers.

Corresponding author **Arshi Jamal**

Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

Type of Quantum computing

Quantum computing encompasses several different approaches and technologies. The primary types of quantum computing are:

1. **Gate-Based Quantum Computing**: This is the most well-known and widely explored type of quantum computing. It uses quantum gates, which are analogous to classical logic gates, to manipulate qubits and perform calculations. Examples of gate-based quantum computing technologies include superconducting qubits (used by companies like IBM and Google), trapped-ion qubits (employed by companies like IonQ), and topological qubits (an emerging technology).

2. **Quantum Annealing**: Quantum annealers, such as those developed by D-Wave Systems, use a different approach to solve optimization problems. Instead of using gates and circuits, they rely on quantum annealing processes to find the lowest energy state of a system, which corresponds to the solution of the optimization problem.

3. **Adiabatic Quantum Computing**: Adiabatic quantum computing is a variation of quantum annealing that relies on slowly changing the Hamiltonian (a mathematical operator representing the system's energy) to find the ground state, which contains the solution to the problem.

4. **Quantum Simulators**: Quantum simulators are specialized quantum computers designed to simulate physical systems at the quantum level. They are used for research in fields like chemistry, materials science, and quantum physics.

5. **Quantum-inspired Computing**: While not true quantum computers, quantum-inspired computers use classical hardware and algorithms inspired by quantum principles to perform certain calculations more efficiently than traditional classical computers. They offer some advantages for specific problem domains.

6. **Photonic Quantum Computing**: This approach uses photons as qubits. Photonic quantum computers are known for their potential in quantum communication and quantum cryptography due to the ability of photons to travel long distances without significant interference.

7. **NMR Quantum Computing**: Nuclear Magnetic Resonance (NMR) quantum computing utilizes the principles of quantum mechanics to manipulate the nuclear spins of molecules. It has been used for small-scale quantum computing experiments.

Corresponding author Arshi Jamal
Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

Each type of quantum computing has its own set of advantages, challenges, and applications. Researchers and companies are actively exploring these various approaches to develop practical quantum computers for a wide range of computational tasks and scientific advancements.

Need of the Study

The study on "Quantum Computing for Big Data Processing and Analytics" is indispensable in our data-driven age, primarily due to the exponential growth in data generation and the limitations of classical computing methods. With vast volumes of data being generated daily, traditional computers struggle to efficiently process and analyze this wealth of information. This study becomes crucial for organizations seeking a competitive advantage, as those capable of harnessing Big Data for strategic insights are better positioned in the market. Additionally, the security of data is paramount, and quantum-resistant encryption techniques, an outcome of quantum computing, are essential in safeguarding sensitive information. Beyond business applications, quantum computing has far-reaching implications in scientific research, materials science, drug discovery, and environmental sustainability. Furthermore, its potential to optimize computational resource utilization and reduce environmental impact makes it an attractive avenue for exploration. As the quantum computing field continues to evolve, understanding its potential in Big Data analytics is essential for future-proofing technological strategies and staying ahead in the rapidly changing landscape of information technology.

Literature Review

Pandey, A et al (2015) The integration of quantum computing into the domain of Big Data analysis marks a pivotal advancement in the era of data-driven decision-making. Quantum computing harnesses the extraordinary properties of quantum bits (qubits) to offer unprecedented computational power. This paper highlights the transformative potential of quantum computing in Big Data analysis. Quantum computers leverage superposition and entanglement, allowing them to explore multiple solutions simultaneously and process vast datasets with remarkable efficiency. Quantum algorithms like Grover's and Shor's enable accelerated data searching and factorization, making complex tasks faster and more accessible.Quantum machine learning algorithms further enhance data analysis by unveiling hidden patterns within large datasets. Additionally, quantum computing addresses data security concerns with the development of quantum-resistant encryption methods.

Shaikh, T. A., & Ali, R. (2016). This survey paper explores the burgeoning intersection of quantum computing and big data analytics. In the era of data proliferation, traditional computing methods face formidable challenges in processing and extracting insights from

Corresponding author **Arshi Jamal**
Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

massive datasets. Quantum computing, harnessing the principles of quantum mechanics, offers a novel approach to tackle these challenges.This paperprovides a concise overview of the survey's scope, emphasizing the relevance of quantum computing in the realm of big data analytics. It highlights the unique properties of quantum bits (qubits), including superposition and entanglement, which enable quantum computers to process vast datasets exponentially faster than classical counterparts.

Heger, D.et al (2014) The field of big data analytics has made significant strides, but it stands at a pivotal juncture, prompting the question, "Where to go from here?" This abstract provides a succinct overview of the key challenges and future directions in the realm of big data analytics.Big data has ushered in a data-rich but insight-poor era, necessitating advanced analytics techniques. The future lies in the convergence of artificial intelligence, machine learning, and big data, enabling deeper and more accurate insights. Augmented analytics, with its focus on automating data preparation and insight generation, will play a pivotal role.Data privacy and ethics concerns also demand attention, leading to the development of responsible AI and data governance practices.

Wiebe, N., Braun, D., & Lloyd, S. (2012). The development of a quantum algorithm for data fitting represents a significant advancement at the intersection of quantum computing and data analysis. Traditional data fitting methods often face challenges when dealing with large and complex datasets. Quantum computing, leveraging the power of quantum bits, offers a novel approach to enhance data fitting processes.This research provides a concise overview of the quantum algorithm's potential in data fitting. Quantum computers exploit superposition and quantum parallelism to process data in multiple states simultaneously, enabling more efficient curve fitting and parameter optimization. Quantum algorithms can potentially identify underlying patterns and relationships in data, even in the presence of noise.The application of quantum computing to data fitting has the potential to revolutionize fields such as finance, healthcare, and scientific research by accelerating the discovery of meaningful insights from vast and intricate datasets.

Wu, X.,et al (2013) Data mining with big data represents a pivotal frontier in contemporary data analytics. The proliferation of massive datasets from diverse sources has brought both opportunities and challenges to the field. With data volumes surpassing traditional processing capabilities, scalable solutions have become imperative. Distributed computing frameworks like Hadoop and Spark have emerged to tackle this scalability issue. Moreover, the real-time nature of big data demands adaptive data mining techniques capable of processing streaming data for timely insights and decision-making.

Corresponding author **Arshi Jamal**

Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

Scope of Data Processing

As depicted in Figure 1, there exist four distinct approaches for managing data through current computational methods. These data processing methods, in conjunction with computational techniques, serve diverse applications tailored to various types of data. Within this context, "classical data" encompasses the various data types encountered in the world, encompassing even big data. In the context of our thesis, our primary emphasis lies on the Classical Data and Quantum Processing (CDQP) category. Our research within this category primarily concentrates on numerical and spatial big data processing.

Data \ Processing	Classical	Quantum
Classical	Classical Data (CD) Classical Processing (CP)	**Classical Data (CD) Quantum Processing (QP)**
Quantum	Quantum Data (QD) Classical Processing (CP)	Quantum Data (QD) Quantum Processing (QP)

Figure 1: Data processing and computational techniques.

Quantum Machine Learning Framework

The Quantum machine learning approach to big data analytics involves a series of steps, as illustrated in Figure 2. In our research, we assume that the collection and organization of big data have been accomplished using established techniques, and this data is readily accessible online. Subsequently, we employ a quantum machine learning framework to conduct the processing of this substantial dataset.

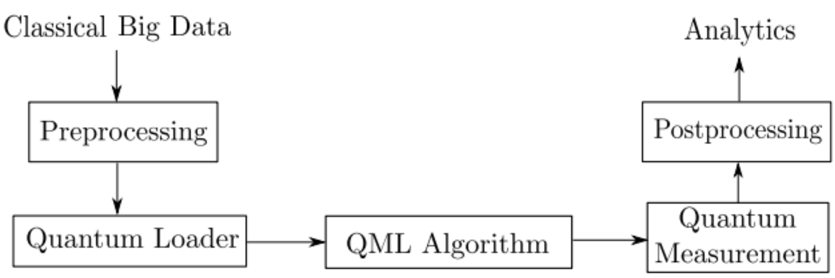

Figure 2: Block diagram of QML approach.

Corresponding author **Arshi Jamal**
Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

Preprocessing

A substantial volume of classical data originates from various sources within the realm of big data. This data undergoes a cleaning process to eliminate any instances of missing values or inconsistencies. Additionally, data preprocessing methods are employed to modify the data, making it compatible for encoding into qubits. Consequently, preprocessing procedures, including data transformation and data reduction, are applied to the classical dataset.

Quantum Loader

A "Quantum Loader" is a term commonly associated with the realm of quantum computing, a cutting-edge field with transformative potential. In this context, it typically refers to software or hardware components integral to the quantum computing process. One of its primary functions is to facilitate the execution of quantum algorithms or circuits on a quantum computer. Quantum computers require specialized software tools to prepare and transmit quantum instructions to the underlying quantum hardware.

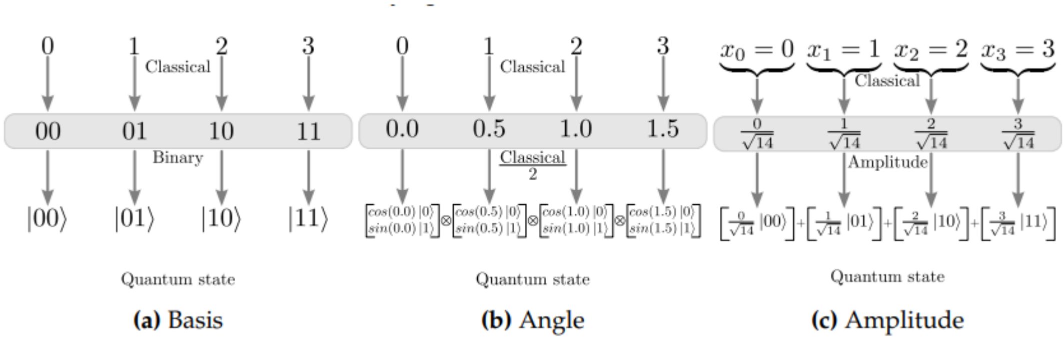

Figure 3: Three embedding techniques

Furthermore, a Quantum Loader may also encompass physical components within a quantum computing system that play a crucial role in initializing qubits or quantum states, ensuring the quantum computer is prepared for specific computations. It serves as an interface between classical instructions and the quantum realm, making quantum computing accessible to researchers, programmers, and organizations seeking to harness its capabilities.

Basis Embedding

"Basis embedding" is a term commonly used in the context of quantum computing. It refers to the process of representing or encoding quantum information from one quantum computing basis into another. Quantum computing relies on qubits, which can exist in various quantum states. Different quantum computing platforms may have different natural or default bases for representing qubit states.

Corresponding author **Arshi Jamal**

Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

$$|\psi\rangle = 0\,|00\rangle + 0\,|01\rangle + 1\,|10\rangle + 0\,|10\rangle\,.$$

The concept of basis embedding becomes relevant when you want to perform quantum operations or calculations in a basis that is different from the native or default basis of the quantum hardware. This process involves transforming the quantum states of qubits to align with the desired basis for a specific computation.

Angle Embedding

"Angle embedding" is a technique commonly employed in quantum computing to represent classical data in a quantum format, making it suitable for quantum computations. This technique is particularly relevant in quantum machine learning and quantum algorithms for optimization.

In angle embedding, classical data, often numerical values, are mapped onto the amplitudes of quantum states. This is achieved by encoding the information into the angles or phases of qubits within a quantum circuit. By doing so, classical data is transformed into a quantum superposition state, which can simultaneously represent multiple classical data points.

Angle embedding is instrumental in quantum algorithms such as the Quantum Variational Eigensolver (QVE), where the angles or parameters of quantum circuits are optimized to find solutions to complex problems, such as optimizing molecular structures or solving combinatorial optimization problems.

Amplitude embedding

"Amplitude embedding" is a technique used in quantum computing to encode classical data into the amplitudes of a quantum state, allowing it to be processed and manipulated by quantum algorithms. This method is particularly relevant in quantum machine learning and optimization tasks.In amplitude embedding, classical data points are represented as probability amplitudes of quantum states, typically using a quantum superposition of basis states. The amplitudes are set according to the values of the classical data points, effectively encoding the information into the quantum state.

$$|\psi_{amp}\rangle = R\,(x_i, \beta)\,|q_1 q_2 \ldots\ldots q_{s-2} q_{s-1}\rangle\,|q_s\rangle$$

This quantum state, which now incorporates the classical data as amplitudes, can be subjected to quantum operations and transformations. Quantum algorithms, such as the Quantum Amplitude Estimation algorithm (QAE), leverage this technique to extract information or find solutions to specific problems.Amplitude embedding capitalizes on

Corresponding author **Arshi Jamal**

Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

quantum superposition and interference, where the quantum state can simultaneously represent multiple classical data points. This allows quantum computers to potentially solve certain problems more efficiently or find optimal solutions in optimization tasks.
Hence, an additional '0' needs to be padded to the input vector X = {x1, x2, . . . , xn}. After the rotations, the encoded quantum state represents the input vector X in its amplitudes.

$$|\psi\rangle = \sum_{i=0}^{n-1} x_i |i\rangle$$

In this context, xi represents the normalized value of the i-th classical feature within the input vector X. Consequently, all the values of the input vector X are encoded as the amplitudes of quantum states in a superposition. This property makes amplitude embedding a space-efficient technique, as it provides exponential scaling by encoding values into the quantum state. It leverages the principles of superposition in quantum computing to simultaneously represent and process multiple classical data points, potentially offering a significant advantage in terms of computational efficiency for certain quantum algorithms and problem-solving tasks.

Experiments on Numerical Data for Classification

In this section, we delve into the comprehensive description of our experimental endeavors, focusing on numerical datasets and employing Quantum Circuit Artificial Neural Networks (QC ANN) for binary classification, as well as Quantum Multi-Class Classification (QMCC) for multi-class classification tasks. These experiments were meticulously carried out using the "default.qubit" simulator, thoughtfully provided by PennyLane. Our experimentation encompasses the utilization of the Wisconsin Breast Cancer dataset for binary classification tasks, employing the power of QC ANN. Additionally, we extended our analysis to encompass a suite of benchmark numerical datasets commonly used in machine learning, including Iris banknote authentication (BNA, and wireless indoor localization. These multi-class datasets were subjected to classification using the QMCC methodology.

Description of the Numerical Datasets

In our experimental setup, we have utilized a set of benchmark numerical datasets widely employed in machine learning. Specifically, we have employed the Breast Cancer Wisconsin (Original) dataset (WBCD) to assess the performance of QC ANN. Additionally, for evaluating the effectiveness of Quantum Multi-Class Classification (QMCC), we have employed the Iris Banknote Authentication (BNA), and Wireless Indoor Localization (WIL)

Corresponding author Arshi Jamal
Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

datasets.Let's delve into the specifics of these datasets, starting with WBCD. The WBCD dataset, sourced from the UCI machine learning repository, serves as a binary classification dataset. It comprises attributes that represent measurements related to breast cancer cases. Within this dataset, we encounter two distinct class labels: "benign" and "malignant." The dimensionality of the dataset stands at 32 attributes, encompassing a total of 569 instances. Among these instances, 357 belong to the benign class, while 212 belong to the malignant class.These datasets serve as critical components of our experimental framework, enabling us to thoroughly assess the performance and capabilities of QC ANN and QMCC in various classification scenarios.

Performance Analysis of QC ANN for Binary Classification

For the classical Artificial Neural Network (ANN) designed to perform binary classification on the WBCD dataset, the architecture is structured as follows:

1. Input Layer: The input layer is comprised of 30 nodes, reflecting the input dimension, which is 30. Since the dataset features 30 attributes, reducing the number of nodes in the input layer is not feasible.

2. Hidden Layers: Two hidden layers are incorporated into the network. The first hidden layer consists of 9 nodes, while the second hidden layer comprises 3 nodes.

3. Output Layer: The output layer is singular, consisting of 1 node. This configuration aligns with the binary nature of the classification task, where each data record is associated with a binary class label (either 0 or 1).

Consequently, the total number of parameters that require optimization within this ANN architecture amounts to 300. This computation is derived from the sum of the products of nodes between layers, which is calculated as follows: $(30 \times 9) + (9 \times 3) + (3 \times 1) = 300$.

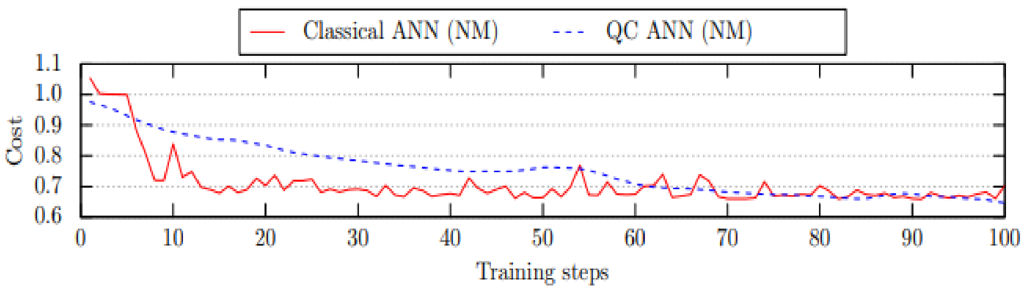

Fig 4 Cost Vs number of training steps of Classical ANN and QC ANN using NM

Corresponding author **Arshi Jamal**

Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

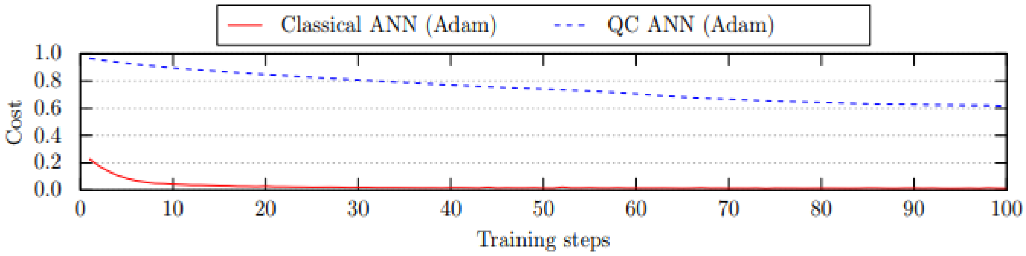

Fig 5 Cost Vs number of training steps of Classical ANN and QC ANN using Adam

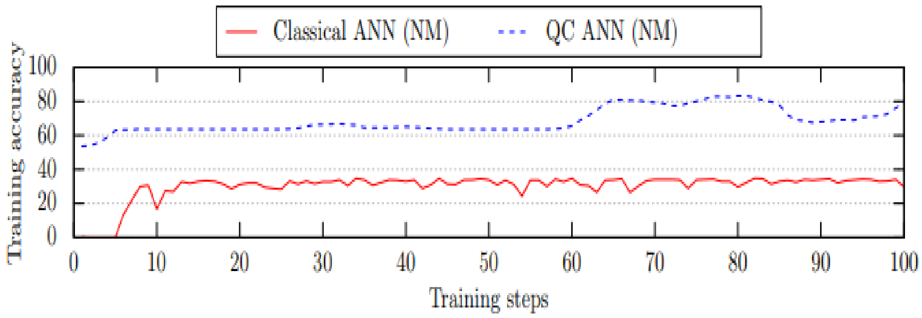

Fig 6 Training accuracy comparison of Classical ANN and QC ANN using NM

The Quantum Circuit Artificial Neural Network (QC ANN) is specifically implemented using 5 qubits in the input layer to accommodate the 30 input values present in the dataset. In this quantum framework, the input values are encoded as amplitudes of quantum states, necessitating a minimum of 5 qubits (as 2^5 = 32) to capture the required quantum superpositions. The amplitude embedding encoding scheme is employed for state preparation, where the 30 input data values are encoded as amplitudes within a quantum state comprising 32 superpositions. To match the size of states in superposition, two additional values for the input are padded with constants. These amplitudes are represented by a state vector with dimensions of 32×1.

Within the hidden layer of the QC ANN, entanglement is established through the use of Controlled NOT (CNOT) gates, and individual qubits undergo rotations via Ry(θ) gates. The Ry(θ) gate signifies a single-qubit rotation around the y-axis, represented within the Bloch sphere, through an angle denoted as θ in radians.

Corresponding author **Arshi Jamal**

Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

Conclusion

In conclusion, the integration of quantum computing into the realm of Big Data processing and analytics holds immense promise and significance. This study has illuminated several key takeaways that underscore its importance. Quantum computing's unique properties, such as superposition and entanglement, provide a fundamentally new approach to handling Big Data. These properties enable the processing of vast volumes of data with unprecedented efficiency and speed, overcoming the limitations of classical computing systems.The potential of quantum algorithms, including Grover's and Shor's algorithms, to dramatically accelerate data search and factorization tasks has far-reaching implications. Businesses and researchers can harness these algorithms to extract insights from massive datasets in real-time, enabling more informed decision-making and enhancing competitiveness.Quantum machine learning algorithms offer another dimension of advancement in Big Data analytics. These algorithms can unlock patterns and trends hidden within complex datasets, revolutionizing fields such as finance, healthcare, and predictive modeling. Quantum computing addresses the pressing concern of data security. Quantum-resistant encryption methods are critical in safeguarding sensitive information in an era where classical encryption methods are susceptible to quantum attacks.

Corresponding author **Arshi Jamal**

Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021

References

1. Pandey, A., & Ramesh, V. (2015). Quantum computing for big data analysis. Indian Journal of Science, 14(43), 98-104.
2. Shaikh, T. A., & Ali, R. (2016, December). Quantum computing in big data analytics: A survey. In 2016 IEEE international conference on computer and information technology (CIT) (pp. 112-115). IEEE.
3. Agarwal, P., & Alam, M. (2021). Exploring Quantum Computing to Revolutionize Big Data Analytics for Various Industrial Sectors. In Big Data Analytics (pp. 113-130). Auerbach Publications.
4. Heger, D. O. M. I. N. I. Q. U. E. (2014). Big data analytics—'Where to go from here'. International Journal of Developments in Big Data and Analytics, 1(1), 42-58.
5. Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. Journal of big data, 6(1), 1-25.
6. Wiebe, N., Braun, D., & Lloyd, S. (2012). Quantum algorithm for data fitting. Physical review letters, 109(5), 050505.
7. Wittek, P. (2014). Quantum machine learning: what quantum computing means to data mining. Academic Press.
8. Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). Trends in big data analytics. Journal of parallel and distributed computing, 74(7), 2561-2573.
9. Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. Journal of business research, 70, 263-286.
10. Zhou, K., Fu, C., & Yang, S. (2016). Big data driven smart energy management: From big data to big insights. Renewable and sustainable energy reviews, 56, 215-225.
11. Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2013). Data mining with big data. IEEE transactions on knowledge and data engineering, 26(1), 97-107.
12. Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. Mobile networks and applications, 19, 171-209.
13. Preskill, J. (1998). Reliable quantum computers. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *454*(1969), 385-410.

Corresponding author **Arshi Jamal**

Submitted: 12 Jan 2021, Revised: 09 March 2021, Accepted: 18 May 2021, Published 30 May 2021