Developing an Adaptive Energy-Efficient Routing Protocol Using Reinforcement Learning for Wireless Sensor Networks

Tilahun Teferi Weyesa

A Thesis Submitted to the Department of Computer Science and Engineering, College of Electrical Engineering and Computing

Presented in Partial Fulfillment of the Requirement for the Degree of Master’s in Computer Science and Engineering

Office of Graduate Studies
Adama Science and Technology University

June, 2025
Adama, Ethiopia

Developing an Adaptive Energy-Efficient Routing Protocol Using Reinforcement Learning for Wireless Sensor Networks

Tilahun Teferi Weyesa

Advisor: Nune Sreenivas (Ph.D. Associate Professor)

A Thesis Submitted to the Department of Computer Science and Engineering,

College of Electrical Engineering and Computing

Presented in Partial Fulfilment of the Requirement for the Degree of Master's in Computer Science and Engineering

Office of Graduate Studies

Adama Science and Technology University

June, 2025

Adama, Ethiopia

DECLARATION

I hereby declare that this Master thesis entitled **“Developing an Adaptive Energy-Efficient Routing Protocol Using Reinforcement Learning for Wireless Sensor Networks”** is my work original work. That is, it has not been submitted for the award of any academic degree, diploma or certificate in any other university. All sources of materials that are used for this thesis have been dualy acknowledge through citation.

Tilahun Teferi Weyesa   ______________   _______________

Name of student   Signature   Date

RECOMMENDATION OF ADVISORS

I, the Major advisor of this thesis, hereby certify that I have closely advised the student while developing this thesis and read the draft thesis entitled "**Developing an Adaptive Energy-Efficient Routing Protocols Using Reinforcement Learning for Wireless Sensor Networks**" prepared under my guidance by Tilahun Teferi Weyesa submitted in partial fulfillment of the requirement for the degree of Master's of Science in Computer Science and Engineering. Therefore, I recommend the submission of the revised version of the thesis to the department following the applicable procedures.

Nune Sreenivas (Ph.D. Associate Prof.) ____________ ____________________

Major Advisor Signature Date

____________________________________ ____________ ____________________

Co- Advisor Signature Date

APPROVAL PAGE OF M.Sc. THESIS

I/We, the advisors of the thesis entitled "**Developing an Adaptive Energy-Efficient Routing Protocols Using Reinforcement Learning for WSNs**" developed by Tilahun Teferi Weyesa, hereby certify that the recommendation and suggestions made by the board examiners are appropriately incorporated into the final version of the thesis.

Nune Sreenivas (Ph.D. Associate Prof.)	___________	________________
Major Advisor	Signature	Date

We, the undersigned, members of the Board of Examiners of the thesis by Tilahun Teferi Weyesa have read and evaluated the thesis entitled "**Developing an Adaptive-Energy Efficient Routing Protocols Using Reinforcement Learning for WSNs**" and examined the candidate during the open defense. This is, therefore, to certify that the thesis is accepted for the partial fulfillment of the requirement of the degree of Master of Science in Computer Science and Engineering.

Chairperson	Signature	Date
Internal Examiner	Signature	Date
External Examiner	Signature	Date

Final approval and acceptance of the thesis is contingent upon submission of its final copy to the Office of Postgraduate Studies (OPGS) through the Department Graduate Council (DGC) and College Graduate Committee.

Department Head	Signature	Date
College Dean	Signature	Date
Office of Postgraduate Studies	Signature	Date

ACKNOWLEDGMENT

First, I would like to express my gratitude to Almighty God for His unwavering blessings and guidance through this journey. His presence has been a tangible source of strength, inspiration, and motivation, especially during the challenging stages of my study.

I would like to extend my deepest gratitude to my advisor, Dr. Nune Sreeenivas, for his unwavering support, insightful guidance, and invaluable expertise throughout the study of this research. He is available whenever I need his support and guidance regarding my thesis. I am particularly appreciative of his continued availability and dedication through the process.

I extend my deepest and heartfelt thanks also to my instructor, Dr. Ketema Adere (Ph.D., Associate Professor), for his invaluable contributions to this research. However, he is not my formal advisor; his unwavering support, specialized guidance, and deep expertise were instrumental in shaping this work. His appointment, thoughtful feedback, and advice greatly encouraged every stage and step of my study. I am greatly thankful to my respected family for their continuous support and wise counsel from the beginning to the accomplishment of my academic and personal endeavors. Their constant motivation has been vital to my progress. Lastly, I wish to thank all the members and students of the Network Science and Security SIG for their meaningful contributions and encouragement while conducting this study. Our involvement and guidance have been sincerely appreciated.

Thank you all for playing a part in the realization of this research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

BIM	Building Information Modeling
BS	Base Station
DQN	Deep Queue Network
HEED	Hybrid Energy-Efficient Distributed Clustering
IoT	Internet of Things
LEACH	Low-Energy Adaptive Clustering Hierarchy
MAC	Medium Access Control
MDP	Markov Decision Process
NF	Next-Forwarder
PCA	Principal Component Analysis
PDR	Packet Delivery Ratio
RFID	Radio Frequency Identification
RL	Reinforcement Learning
RLBEEP	RL- Based Energy-Efficient Control and Routing Protocol
SN	Sensor Node
TCP	Transmission Control Protocol
TEEN	Threshold Sensitive Energy Efficient Sensor Network
WSN	Wireless Sensor Network

ABSTRACT

One of the core themes in wireless sensor networks is the creation of methods to enhance network lifetime. Their operation indeed relies on energy-efficient routing protocols that use minimal energy. Energy-efficient routing is a problem of utmost importance in WSNs to ensure sustainable data communication between nodes, cluster heads, and base stations without data loss. Because they are battery-powered and resource-constrained resource availability and additional battery supply is called for. Thus, Wireless Sensor Networks (WSNs) are confronted with limited energy availability due to the finite power capacity of sensor nodes, and the necessity of employing adaptive routing protocols for effective energy consumption while maintaining network operability. The present research work suggests an Adaptive Energy-Efficient Routing Protocol based on Reinforcement Learning (AEERPRL) to mitigate such challenges through ongoing optimization of routing choices for simulated and controlled network conditions. Evolution of (RL), AEERPRL streamlines nodes to learn ideal data forwarding routes, modulating energy consumption and network reliability. The protocol is compared against the RL-based RLBEEP protocol on major parameters: Alive Nodes (network lifetime), Packet Delivery Ratio (PDR), and Throughput. The experiment results demonstrate that AEERPRL outperforms RLBEEP with 10% additional surviving nodes over time, achieving a 0.8 PDR (compared to 0.6 RLBEEP) and 25% better throughput, indicating increased management and network robustness. The enhancements indicate AEERPRL's capacity for network lifetime prolongation with assured, reliable data delivery even in changing environments. The findings demonstrate the efficacy of reinforcement learning to develop and deploy adaptive routing protocols in wireless sensor networks with practical applications in precision agriculture, environmental monitoring, and Internet of Things networks where energy efficiency and scalability are essential. The research recommends the use of machine learning in wireless sensor network routing protocols and proposes a scalable model for future research in energy-limited networks.

Keywords: *Adaptive Routing Protocols, Energy Efficiency, Network Lifetime, Optimization, Reinforcement Learning, and Wireless Networks.*

CHAPTER ONE

INTRODUCTION

1.1 Background

Wireless Sensor Networks have a great impact in numerous fields such as environment management, healthcare, agriculture, military operations, and smart technologies, where distributed sensing and communication are natural features. Since WSNs are resource and battery-limited, researchers are searching for the current issues of the system. Yet, the resource-limited nature of sensor nodes is a great problem for their long lifespan. The solution to the issue involves the use of energy-efficient routing protocols because they enable the enhancement of various crucial objectives (Nakas et al., 2020).

Power-aware routing for Wireless Sensor Networks (WSNs) is necessary due to their issue with battery-driven nodes. They measure and report environmental information across large or adversarial environments for enabling applications like industrial automation, traffic monitoring, health care, and environmental regulation. However, constrained energy supplies are the biggest barrier to their long-term operation and deployment (Priyadarshi, 2024a). Sensor nodes transmit data but are low power, and thus energy consumption and network lifetime become issues in the overall operations of such nodes. Future research will be critical in deciding the most effective paths from source to destination nodes to conserve energy and prolong network lifetime. Yet, the building of routing algorithms and the choosing of the most appropriate ones remains a challenging task to save energy effectively and extend network lifetime, challenging to adopt energy-efficient Routing(Al Aghbari et al., 2020).

Effective energy routing is crucial in Wireless Sensor Networks (WSNs) as they are deployed with characteristics of a dynamic topology, in addition to constrained resources and battery life. Effective implementation of energy plays a significant role in the network's lifetime and performance(Pantazis et al., 2013). The energy-efficient routing must be implemented in Wireless Sensor Networks (WSNs) since sensor nodes possess scarce energy supplies.

The adoption of energy-efficient routing is essential in Wireless Sensor Networks (WSNs) because sensor nodes have limited energy resources. WSNs consist of a large number of resource-constrained, low-power devices deployed over a large geographical area, and energy efficiency is a self-evident issue in their deployment (Mann & Singh, 2017).

To extend the life of the network, efficient routing protocols must be employed to reduce the usage of energy in communication and information transfer processes. Wireless Sensor Networks (WSNs) are a field of continuing research, whose nature continues to change every second, particularly concerning node interconnection. Nevertheless, data communication remains a fundamental element that must work maximally for energy efficiency. Simulation has enabled us to conclude that the careful choice of a routing algorithm plays a crucial role in contributing significantly to rendering energy efficiency and network lifetime dependent (Prabhu et al., 2023). Wireless Sensor Networks (WSNs) are instrumental in a wide range of applications and fields, such as ecological management, intelligent agriculture techniques, urbanization, and management of healthcare systems (Qiu et al., 2024).

Wireless Sensor Networks (WSNs) play a vital role in current technology, enabling distributed sensors to be networked throughout the space in process monitoring and control. WSNs can be used as they are a cost-efficient, energy-efficient infrastructure, provide ubiquitous connectivity, adaptability to respond to shifting circumstances, and autonomy (Alsalmi et al., 2024a). Wireless Sensor Networks (WSNs), as the ultimate component of the Internet of Things (IoT) systems, rely greatly on the efficiency of quick data routing in enhancing the overall network performance. Yet, current routing protocols tend to face issues related to aspects such as continuous node mobility, enhancing energy conservation, guaranteeing scalability, and supporting adaptability to changing network conditions (Suresh et al., 2024a). Effective management of energy consumption is a significant concern in wireless sensor network engineering. This has a profound impact on the durability and sustainable operational capability of sensor network systems (Abadi et al., 2022).

IoT wireless sensors operate in the realm of high computation demand and a low energy resources environment. Techniques that can be used to extend the life of these nodes are therefore highly sought after without frequently replacing batteries or relocating them. Among these, sensing-related operations, data transmission, are most energy-hungry, followed by route computation, idle listening for incoming traffic, and data processing. In this case, the correct development of routing protocols that can decide the optimal path for securely delivering data while at the same time consuming the least energy possible is required for driving higher performance in wireless sensor networks driven by the Internet of Things (IoT) (Godfrey et al., 2023). A reinforcement learning-based control and routing protocol for a power-saving network control mechanism was proposed by (Yadawad & Joshi, 2024).

The paper introduces a data acquisition method for energy-harvested wireless sensor networks based on cooperative multi-agent reinforcement learning approaches. (Dvir et al., 2024) have suggested a fresh data gathering approach for energy-harvesting wireless sensor networks based on multi-agent reinforcement learning (MARL) within a decentralized autonomous communication network in which sensors are single agents. Here, we suggest the application of multi-agent reinforcement learning to realize a wireless sensor network routing protocol that will dynamically learn and be energy efficient.

1.2 Motivation of the Study

1.2.1 Energy Limitations in WSNs

A significant disadvantage of wireless sensor networks (WSNs) is that they are battery-dependent (H.Wang et al., 2024). A routing protocol's energy consumption plays a significant role in the lifespan of a wireless sensor network (WSN), as the tiny sensor nodes are non-rechargeable easily upon deployment (Yun & Yoo, 2021). Energy-efficient load and resource utilization balancing for energy-limited WSNs is extremely difficult because of the constraint of energy (Ding et al., 2021). While energy is an important component, WSNs need continuous charging of their batteries to operate continuously and achieve the objectives of the system.

1.2.2 Changing Network Conditions

Wireless sensor networks are widely renowned for progress and active engagement with data collection and forwarding of information in an IoT network environment; however, they face several problems in maximizing the performance and scalability of the network under multidimensional environments (Wakili et al., 2024).

Wireless Sensor Networks (WSNs) are familiar with their basic characteristics for the management and acquisition of environmental conditions using spatially distributed equipment of sensor nodes. These networks are susceptible to dynamic issues because of changing network factors that can affect their efficacy and dependability. This dynamic system conditions tackles the biggest issues regarding dynamically changing network conditions in WSNs, e.g., the impact on network control, protocol design, and power efficiency (López-Ardao et al., 2021). These network-changing conditions influence the functionality of WSNs.

1.2.3 Adaptive Routing with Reinforcement Learning

Adaptive routing in Wireless Sensor Networks (WSNs) is defined by node mobility, energy inefficiency, and scalability. Machine learning-based and conventional energy-efficient WSN routing protocol mechanisms struggle to resolve these issues. To mitigate these constraints, a Reinforcement Learning (RL)-based model is selected to enable distributed decision-making, adaptive routing, and optimal use of resources in current WAN application domains. Running Simulations over and over demonstrates that RL outperforms traditional routing methods (Suresh et al., 2024b). Since Reinforcement learning ML is able to adapt to the dynamic nature WSNs, we preferred DQN RL to solve these problems.

1.2.4 Challenges of Scalability

As IoT-based networks are more technologically advanced, routing techniques must evolve, but scaling them presents difficulties, such as data size, power usage, and the need for flexibility. Systems utilizing reinforcement learning, despite their potential, must address these constraints to guarantee efficient operation in large networks. Improved data integrity, increased energy efficiency, and security complicate this integration further (Priyadarshi, 2024b). To overcome these problems, we selected DQN RL that upgrades QN from the previous research conducted.

1.3 Statement of the Problem

Wireless Sensor Networks (WSNs) are faced with an abundance of challenges, including ineffectiveness of communication, complexities in cluster management, and limited battery node capacity. Saving energy in these networks is of very high importance. Energy consumption is quite steady during data receipt and sensing, but data transmission significantly contributes to the depletion of energy resources. One of the finest approaches to solving this problem is the use of Reinforcement Learning (RL) in WSNs to support intelligent decision-making and smarter energy utilization (Choudhary & Barwar, 2024). (Wakili et al., 2024b) Suggested building an adaptive Reinforcement learning-based, efficient routing protocol for WSN. Although Reinforcement Learning (RL) has the potential to improve routing reliability and minimize latency in Wireless Multimedia Sensor Networks (WMSNs) for the Internet-of-Vehicles, finding a balance between Quality of Service (QoS) and the constrained energy budgets of heterogeneous sensor nodes is an open and challenging task (Wang et al., 2020).

Although WSNs are very effective for sensing physical environments, it is a big challenge to design an energy-efficient routing protocol due to energy-constrained nodes and hotspot issues, and thus, intelligent dynamic routing protocols are required to maximize network lifetime and reduce energy consumption (Keerthika & Berlin Hency, 2022).

1.4 Research Questions

In this study, an attempt will be made to answer the following questions.

1. How can DQN reinforcement learning adopt dynamic routing paths?
2. What are the main constraints of existing energy-efficient routing protocols in WSNs?
3. How can we model WSNs using Deep Q-Network (DQN) reinforcement learning?

1.5 Objectives of the Study

1.5.1 General Objective

- ✓ To develop an adaptive energy-efficient routing protocol using reinforcement learning for wireless sensor networks. To design an adaptive energy-efficient routing protocol using reinforcement learning for wireless sensor networks.

1.5.2 Specific Objectives

The particular objectives identified to reach the general objectives of our research are:

- ✓ To design a reinforcement learning-based routing algorithm that adapts dynamically.
- ✓ To generate a valuable dataset that captures the energy consumption.
- ✓ To test the performance of the proposed protocol through simulations.

- ✓ To compare the performance of the new energy-aware routing protocol with the baseline standard routing protocol.
- ✓ To evaluate the proposed routing protocol's performance.

1.6 Significance of the Study

This research has the potential to deliver significant benefits across various fields. Environmental monitoring can optimize data transmission in remote or hard-to-reach areas, such as forests, oceans, or farmlands. In healthcare, it enables energy-efficient data transmission from wearable devices that monitor patients' vital signs. For smart cities, the research supports energy-efficient data exchange in applications like smart grids, traffic monitoring, and infrastructure management. Additionally, in military and defense, it enhances the reliability of sensor networks for surveillance and communication in remote or hazardous environments, ensuring long-term operation.

1.7 Scope and Limitations of the Study

1.7.1 Scope of the Study

The present study aims to develop an adaptive routing protocol tailored for Wireless Sensor Networks (WSNs) based on Deep Q-Network (DQN), one of the reinforcement learning algorithms that are famous for their ability to make informed decisions with the help of feedback provided by the environment. The novel protocol is particularly designed to enhance energy efficiency, a factor of pivotal significance for WSNs due to the latter's reliance on limited power supplies. The research intends to come up with smart routing protocols that are capable of addressing the inherent issues of WSNs, such as time-varying network topologies, changing energy levels among sensor nodes, and time-varying communication conditions.

For evaluating the proposed protocol's efficiency, the research involves detailed simulations with realistic WSN scenarios. These simulations aim to contrast and quantify the performance of the DQN-based protocol with traditional and new developing routing protocols. Total energy consumption, remaining energy, and network lifetime are taken as the key performance indicators as measures. The study also explores how the proposed solution can be integrated in existing WSN frameworks and infrastructure to enable its real-world applicability.

Also, the findings of this research are contrasted against those of existing methodologies to highlight the improvements and innovations introduced by the DQN-based approach. Through such comparative assessment, the research demonstrates the potential of reinforcement learning in increasing routing efficiency significantly, prolonging network lifetime, and optimizing energy consumption, thereby making valuable contributions towards further innovation of energy-efficient WSN technologies.

1.7.2 Limitations of the Study

This research is highly simulation-based and therefore may not fully reflect the complexity of real-world applications within physical landscapes and node mobility. While the scalability of the system has yet to be confirmed within larger-scale or longer-distance networks, such a limitation would most likely restrict its applicability to high-density network environments where real-time responsiveness is less critical. Much of the analysis has centered on specific energy consumption metrics, not on other important factors such as sleep/wake cycles and duty cycling methods that directly affect overall network performance and energy efficiency.

Furthermore, reinforcement learning (RL) algorithms, although promising, unavoidably demand significant amounts of training data and, thus, are not reasonable for real-time utilization in highly dynamic sensor networks. The assumptions that are made regarding static network conditions further fail to reflect the dynamic nature of real environments, leading to a difference between simulated results and real performance. Such assumptions also lead to increased algorithmic complexity and unrealistically optimistic predictions that cannot be replicated once network conditions have shifted. As background information, the chapter introduces Wireless Sensor Networks (WSNs), classifies a number of WSN types, and reviews a number of approaches to energy conservation, of which those with a foundation in reinforcement learning are especially noted for their dynamics and decision assistance.

1.8 Organization of the Thesis

The remaining of our study or thesis is organized as follows to describe the content of the paper as follows:

The first chapter describes the topic for the overall study, the driving power, motivation of the study, issues, goals, and significance, and steps taken to achieve our objectives. The second chapter incorporates a literature review in detail related to WSNs and their topologies, the concepts, and how they work, and related works are discussed through necessary evaluation and examining related works of previous studies. Chapter three discusses the methodology we use and a detailed description of dataset situations, preprocessing of datasets, feature selection mechanisms, the RL model we are going to use, software and hardware tools used in our study, and its evaluation techniques. Under chapter four, we discuss simulation results and validate the proposed method are described. The implementation of the architecture, including the process of model training, is given in Section 5. Discussion of results, giving a detailed description of the results and a summary of the overall research findings, is given in Chapter 6. Lastly, chapter 7 makes suggestions for future research and points out the study's contributions, referencing its significance in field development.

CHAPTER TWO

LITERATURE REVIEWS AND RELATED WORKS

2.1 Overview of WSN

WSNs show a massive step in the information and communication technology field, and it is transforming at a rapid pace in a networked society. WSN is a massive revolution in the field of information technology, and it is growing at a high pace as a network environment. In the next section, we will describe the various types of WSNs with various means of energy optimization, i.e., means based on Deep Queue Network reinforcement learning and ML. This chapter offers an overview of Wireless Sensor. Wireless Sensor Networks (WSNs), WSN types, and energy optimization methods, including reinforcement learning and ML methods. Routing protocols in Wireless Sensor Networks use different mechanisms that revolve around energy optimization. It also involves approaches to improving the energy efficiency of WSNs, categories of energy optimization techniques, and research on adaptive routing protocols based on reinforcement learning (Safiyan et al., 2022).

2.2 Conventional WSN Routing Protocols Vs Adaptive EERP-RL WSNs

WSNs development has led to the creation of low-cost and small-sized sensor nodes that can perceive sets of physical and environmental states, understand information, and operate in various settings. Advances in technology in WSN have enabled the creation of miniature, low-cost sensor nodes with the ability to sense multiple physical and environmental parameters, perform data processing, and wirelessly communicate. However, WSNs are faced with limited transmission range, energy supplies, as well as the processing and storage capabilities of sensor nodes, and therefore require improved data processing and transmission capabilities. Routing Protocols are needed for reliable multi-hop communication in spite of these problems and limitations. (Wei et al., 2011) Reviews WSNs' routing protocols, balancing their advantages and limitations. (Khan et al., 2021) characterize Wireless Sensor Networks (WSNs) as sensors scattered. We study space to measure several physical or environmental phenomena, including temperature, and collect pressure data and send it to a specific point in the network.

Despite advancements, challenges persist, necessitating a comprehensive review of state-of-the-art technologies, functional and performance analyses, and issues in hierarchical routing protocols. Additionally, it identifies challenges and future research on energy-efficient routing protocol design in WSNs. (Shabbir & Hassan, 2017) Wireless Sensor Networks (WSNs) are becoming increasingly significant due to their diverse applications, some of which prioritize rapid data transfer with minimal interruptions, while others focus on maximizing throughput regardless of delays. The selection of key performance metrics relies on the specific application requirements. The choice of key performance parameters depends on the specific application requirements. A thorough understanding of the network structure and the suitability of the routing protocol is crucial for optimal performance.

Conventional routing protocols are categorized into different types of which are used by most researchers. (López-Ardao et al., 2021) proposed conventional routing protocols categories in WSNs into three: flat-based, hierarchical-based, and location-based routing protocols, but Adaptive EERP-RL protocols collaborate with machine learning techniques to optimize routing decisions based on real-time network conditions. These protocols are designed to dynamically adjust their operations based on factors such as channel conditions, energy consumption, and environmental changes. Reinforcement Learning ML is a proposed method to solve these issues.

Some key features can be considered to optimize energy-efficient routing protocols. Key features include: Energy Harvesting Integration, Reinforcement Learning, and Quality of Service (Qos). The issue of energy balancing in Wireless Sensor Networks (WSNs) is critical for their deployment and is addressed through three main approaches: (i) energy conservation by minimizing communication costs, as the radio consumes the most energy; (ii) energy harvesting by converting and storing energy between nodes. Table 2.1 summarizes the key differences between conventional WSN routing protocols and adaptive energy-efficient routing protocols using reinforcement learning.

Table 2.1:Conventional WSN Routing Protocols Vs Adaptive EERP-RL WSNs

Feature	**Conventional WSN Routing Protocols.**	**Adaptive Energy Efficient Routing Protocol Using Reinforcement Learning.**
Definition	Traditional protocols designed for routing in Wireless Sensor Networks (WSNs) often face energy constraints.	A modern approach that utilizes reinforcement learning to adaptively manage energy efficiency in routing for WSNs.
Energy Efficiency	Typically struggles with energy efficiency due to static routing methods and a lack of adaptability to network conditions.	Employs adaptive strategies that optimize energy consumption based on real-time network conditions and node behavior.
Routing Methodology	Often relies on fixed algorithms such as flooding or gossiping, which can lead to inefficiencies like implosion.	Utilizes machine learning techniques to dynamically adjust routing paths and decisions based on learned experiences.
Scalability	May face challenges in scalability as the network size increases, leading to performance degradation.	Designed to scale effectively with network size, adapting routing strategies as new nodes are added or removed.
Data Handling	Data aggregation is often limited; redundancy can occur without efficient management of data flows.	Capable of intelligently aggregating data based on learned patterns, reducing redundancy, and improving overall efficiency.
Network Dynamics	Less effective in handling dynamic changes in the network topology due to node failures or energy depletion.	Proactively adjusts to changes in network topology, optimizing routes based on current node status and energy levels.

Examples of Protocols	LEACH, PEGASIS, SPIN, Directed Diffusion, etc..	Reinforcement learning-based protocols that continuously learn and optimize routing strategies based on feedback from the network.
Implementation Complexity	Generally simpler to implement, but may require manual adjustments for optimization.	More complexity due to the integration of machine learning algorithms, requiring more computational resources at the nodes.

To address such energy efficiency issues in WSNs, scientists have considered numerous types of routing protocols. The advancement from traditional WSN routing protocols to adaptive EERP-RL approaches represents a significant move toward more wise and efficient network management. As applications of WSNs are ubiquitous, from environmental monitoring to smart cities and health care, power-efficient and adaptive solutions become increasingly essential. We also proposed RL DQN model, which has the ability to adapt dynamically to such dynamic environments. Future work will be increasingly focused on optimizing such adaptation processes further, as well as tackling scalability and security concerns in the networks.

2.3 Overview of WSN

Wireless Sensor Networks (WSNs) form the backbone of the majority of applications, and precise localization of sensor nodes is imperative to their efficacy (Ahmad et al., 2024). Wireless Sensor Networks (WSNs) have a vital role in improving the security, reliability, and efficiency of critical infrastructures (CIs). It is very important for many applications, and the proper placement of sensor nodes is a must for them to perform effectively. In the last few years, optimization methods have received much attention as a means to initiate finding WSN nodes. It talks about the worldwide relevance of these kinds of infrastructures, WSN advancement towards industrial use, and categorizes WSNs while speaking on standards and protocols for adversarial CI environments. (Daousis et al., 2024) identifies a significant gap in the literature regarding the quality standards for equipment used in Cis, implying a necessity for more research in this direction. Wireless Sensor Networks (WSNs) are networks of distributed, autonomous sensor nodes that communicate to gather specific information in a range of application domains. (Vassilaras, 2012) identifies facts of issues regarding convergence, scalability, quality of service, size, computational power, energy efficiency, and security of WSNs.

Wireless Sensor Networks (WSNs) are facilities that include sensing, computing, data collecting, and communication elements, which offer the monitoring and interaction with physical environments. WSNs find application in various application areas. Acting as a bridge between the physical and virtual worlds, WSNs have rapidly advanced due to their diverse applications (Buratti et al., 2009). Wireless Sensor Networks (WSNs) are a new generation of real-time embedded systems with unique communication limitations compared to traditional networks (Rajaravivarma et al., 2003).

2.4 WSN Architectures

Wireless Sensor Networks are a diverse component in that they consist of different nodes. The Wireless Sensor Networks framework applies to the organization and configuration of sensor nodes alongside their communication frameworks, facilitating the acquisition, processing, and distribution of environmental information. WSNs are known for their dynamicity nature. Typically, WSNs are characterized by two primary architectural models: Layered Network Clustered Architecture. A more primary architectural model: Layered Networks Architecture and

Clustered Architecture. A more complete description of these architectures is provided in the later sections. Generally, the WSN structure is divided into various layers as illustrated in Fig. 2.2, with each layer serving specific functions (Alkhatib & Baicher, 2012). The following is a general description of the main components and layers of the majority of WSN architectures. The application layer manages traffic flow and offers software utilities that translate data into a meaningful conclusion format. Wireless Sensor Networks find applications in different fields. They are utilized in different areas, including agriculture, military operations, environmental monitoring, and medical care. The Wireless Sensor Network Application is the highest layer in the WSN protocol stack and is crucial for facilitating proper communication between user applications and the network. Utilizing these advancements in WSN, the use of energy-efficient routing is crucial.

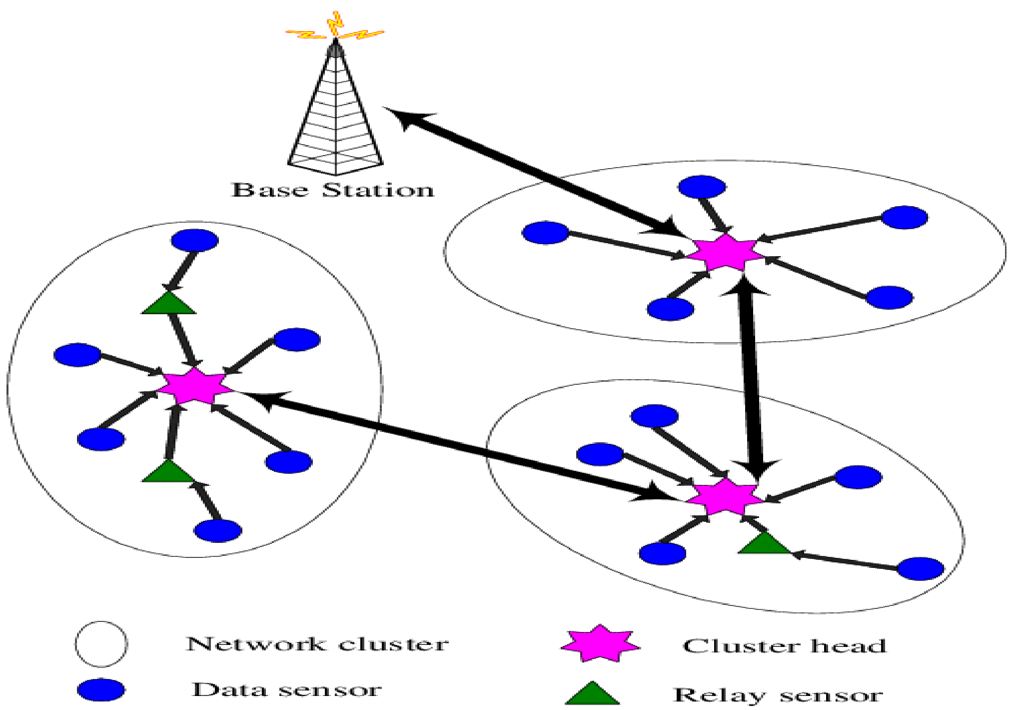

Figure 2.1 : Cluster-based WSN Architecture (Nadeem & Agrawala, 2005)

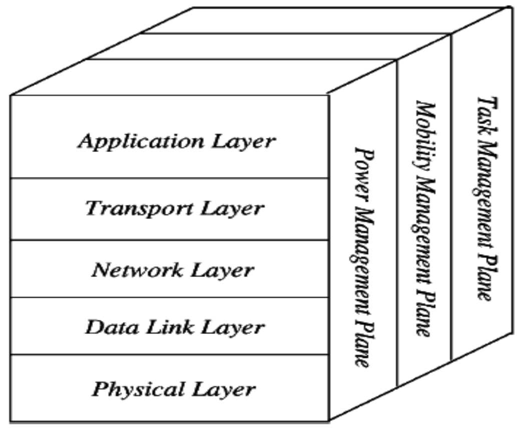

Figure 2.2: Wireless Sensor Network Layers (Sardar & Majumder, 2013)

2.4.1 Application Layer

The application layers of Wireless Sensor Networks allow the realization of applications. The application layer directs traffic and provides software utilities to convert the data into a useful format for acquiring knowledge. Wireless Sensor Networks are utilized across a variety of domains, including agriculture, military, environmental monitoring, and health fields. Wireless Sensor Network application Layer is the topmost layer in the WSN protocol stack and plays a vital role in enabling the user and network to communicate effectively with applications. It is an architecture that supports the deployment of a range of applications with application-specific protocols and services (Mubarak et al., 2024). Wireless Sensor Network application layers facilitate the usage of such applications.

2.4.2 Transport Layer

The transport layer of WSN is responsible for the reliability of data delivery, transmission, and sustaining energy efficiency in resource-scarce settings. The Wireless Sensor Network transport layer of a network has a few limitations. The transport layers in WSNs must fulfill certain challenges such as constrained energy, low processing capability, and bandwidth constraints. In WSN, this layer is accountable for end-to-end communication services that promote efficient data transfer and minimize packet loss and congestion (Yu et al., 2024).

2.4.3 Network Layer

The Network layer of WSNs is concerned with routing the data packets among sensor nodes efficiently to the base station, while simultaneously minimizing energy usage and ensuring the trustworthy delivery of information. This layer plays a fundamental role in defining how data is propagated, particularly in resource-scarce environments like WSNs. (Veena et al., 2024) sets the preconditions of Industrial Internet of Things (IIoT), the performance of the Network. The layer, in this case, determines the scalability, reliability, and power efficiency of industrial applications.

2.4.4 Data Link Layer

Wireless Sensor Networks (WSNs) Data Link Layer is a key component architecture of Wireless Sensor Networks that operates at one level above the physical layer of the OSI model, and plays a key role in facilitating efficient and dependable communication among sensor nodes. Its primary functions such as medium access control (MAC), error detection, retransmission, and allowing control over data communication (Jain, 2024).

2.4.5 Physical Layer

The physical layer of Wireless Sensor Networks (WSNs) constitutes the lowermost level in the network hierarchy. Architecture is significant in the direct transport and interpretation of raw bit streams through the wireless medium. It is a key factor that influences the overall performance, reliability, and energy efficiency (Daousis et al., 2024).

2.5 WSN Benefits

(Daousis et al., 2024) elaborates in great detail on integration with Building Information Modeling (BIM) and Radio Frequency Identification (RFID) systems. WSNs offer several unique benefits that power efficiency, data-driven decision-making, and resource optimization. Wireless Sensor Networks (WSNs) provide numerous benefits, particularly in the healthcare sector, whose applications are transformative. It brings revolutionary advantages in water quality monitoring, overcoming significant limitations with conventional lab-based approaches. However, these advantages come with challenges, such as the need for strong and energy-efficient authentication protocols designed for resource-limited sensor nodes (Alate et al., 2024).

These benefits are prohibited by their limitations. These traditional practices have a tendency to struggle with the necessity for instantaneous feedback and prompt reactions to issues. Conversely, WSN-based methods enable ongoing, widespread monitoring over big regions, giving current details to stakeholders (Y. Singh & Walingo, 2024).

2.6 Overview of Cluster-Based Wireless Sensor Networks

In conventional wireless sensor networks (WSNs), the users, sink nodes, and sensor nodes are static, and the network is deployed in a planar single-layer manner, which is not sufficient to provide support for mobile sensor nodes. An optimized hierarchical WSN topology for mobile sensor nodes reduces routing complexity for mobile sensor nodes, improving network performance in terms of reduced energy consumption and data transmission latency. Simulation results confirm the effectiveness of this hierarchical approach for achieving maximum energy efficiency at minimum latency (Chen & Yu, 2010).

But WSNs are greatly challenged in routing because of the constraints of finite battery life, processing capability, and memory in sensor nodes. The Cluster-based method is the one that can be used to boost energy efficiency. This strategy lowers energy usage by minimizing the transmission distance and balancing energy consumption among various nodes. Recent Development in cluster-based protocols has demonstrated potential to enhance energy efficiency and overall network performance in WSNs (Journal, 2024).

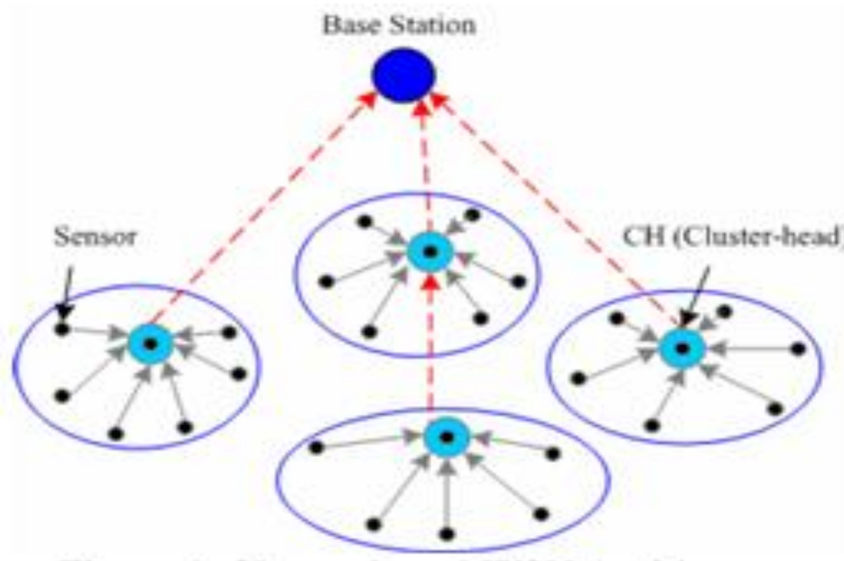

Figure 2.3: WSN Clustering Structure (Thein & Thein, 2009)

2.6.1 Sensor Nodes

Within the described context, sensor nodes are complete systems that are purposely designed to detect and manage environmental conditions, particularly the presence of toxic gases and chemicals. These devices are usually employed where instant reaction to any possible leaks is vital, such as in automated labs in which chemical vapors can present dangers to both human safety and devices (Al-Okby et al., 2024).

2.6.2 Cluster Head

A cluster head (CH) in the Internet of Things (IoT) is a node that is responsible for managing and coordinating the communication among a group of sensor nodes. The CH is the key to energy efficiency and network performance by collecting, gathering, and transmitting data from the nodes in its cluster to the other components of the network. In most cases, the sink conserves energy and expands the network (Somula et al., 2024).

2.6.3 Base Station

In Internet of Things (IoT) capable Wireless Sensor Networks (WSNs), the Sink node plays a central role as the primary nucleus of the collection, accumulation, and exchange of information in the network. The location of the Bs mostly influences the whole performance and energy efficiency of the WSN. Its positioning impacts the efficiency of data transfer, energy consumption, and the network’s scalability. Base stations have different roles in WSNs. In some IoT-assisted WSN implementations, the role of the BS may vary depending on the integration of opportunistic routing protocols (Chaurasia & Kumar, 2024).

2.7 WSN Routing Protocols

Wireless Sensor Networks (WSNs) have brought significant advancements, driven by the development of micro-sensor devices. It consists of heterogeneous devices. These networks consist of hundreds to thousands of sensor nodes capable of collecting data from remote locations and transmitting it to users based on specific applications. However, sensor nodes face limitations in energy, storage, and computational power, making energy-efficient routing a critical aspect; they also benefit from collecting and processing data.

Routing protocols in WSNs enable data transmission between nodes and are designed to explore the unique constraints of these networks. These routing protocols are categorized by some factors. Protocols can be categorized based on key factors such as energy efficiency, scalability, and communication overhead (Bhattacharyya et al., 2010).

2.8 Overview of Energy-Efficient Routing Protocols

Reinforcement Learning-based routing protocols can address the issues of energy-efficient routing protocols. Energy-efficient routing in Wireless Sensor Networks (WSNs) is critical to the limited resources of sensor nodes and their hostile conditions of operation. (Lakshmi et al., 2024). While WSNs are the foundation of IoT applications, they are beset by resource-limited dimensions, e.g., power consumption issues, network lifetime, routing efficiency, throughput, latency, and (Saadallah & Alabady, 2024). These have to proper routing paths to be implemented in WSNs' application areas.

2.8.1 Categories of Energy-Efficient Routing Protocols for WSNs

WSNs use multi-hop communication, i.e., nodes relay data to the base station through other nodes. Routing protocols in WSNs try to save energy, extend the life of the network, and operate effectively with limited resources, but they have some disadvantages. Existing protocols often emphasize either extending network lifetime or implementing security mechanisms, sometimes at the expense of energy efficiency (D et al., 2024). Depending on the energy awareness techniques employed, routing protocols can be classified into the following categories:

2.8.1.1 Data-centric Routing Protocols

The purpose of these routing protocols is to enhance energy efficiency and prolong the working lifetime of Wireless Sensor Networks (WSNs), specializing in data transmission, aggregation, and optimization redundancy elimination. In contrast to conventional address-based routing, these protocols favor the data content rather than the actual sensor nodes transmitting it. These problems necessitate an adaptive approach to address the challenges of WSNs. (Batool et al., 2024) Introduced the main goal to reduce the energy usage of sensor nodes and provide efficient data transmission of relevant information to the sink node.

2.8.1.2 Hierarchical Routing Protocols

Routing protocols used in wireless sensor networks (WSNs) play an important role in maximizing energy usage, scalability, and network lifetime, necessary for resource-constrained settings. Reinforcement learning-based WSN routing protocols are a complex operation aimed at saving energy. These operations structure the network into clusters or layers, enabling good data interaction and organization. Both internally and externally, the hierarchical structure has received much attention in tackling these issues. Hierarchical routing has received considerable academic attention in the last few years in recent times. Attention is increasingly focused on non-traditional hierarchical strategies (Liu, 2015).

2.8.1.3 Location-Based Routing Protocols

Dynamic environments, especially in military cyber-physical systems characterized by variability environments require effective and adaptive communication strategies. These routing protocols are an essential component of Wireless Sensor Networks (WSNs), especially in urban cyber-physical systems where dynamic contexts demand extremely adaptive and productive communication strategies. The UALRP (Location-Based Routing Protocol) integrates real-time data analysis and adaptive machine learning models in its algorithmic framework, which enables dynamic optimization of making routing choices in reaction to the constantly developing urban environment (Akinola, 2024).

2.8.1.4 Multipath Routing Protocols

They are routing protocols that offer an effective solution for WSNs to the issues that are occurring as a consequence of environmental management activities under climate change and anthropogenic processes. The networks play a crucial role in sensing and monitoring incidents such as wildfires, floods, and landslides, where ongoing and dependable operation is required to reduce hazardous events. The inherent characteristics of WSNs low-power sensors, self-maintaining power sources, wireless communication, and myriad sensors, make them ideal for these missions, especially in remote or hostile regions (Üveges & Oláh, 2024).

2.8.1.5 QoS-Aware Routing Protocols

For IoT, which feeds into central applications such as intelligent infrastructure, traffic management, and security systems, enabling extended lifetimes of networks and extensive sensor coverage is an important aspect. These are routing protocols that take a prominent place in Wireless Sensor networks. Wireless Sensor Networks (WSNs) within the IoT paradigm, and more specifically in applications like smart cities. (Karunkuzhali et al., 2024) suggested protocols aimed at enhancing communication and transfer in IoT networks according to their requirements, which include factors such as energy efficiency, latency, reliability, and network lifetime. Adaptive energy-efficient reinforcement learning-based routing protocols play an important role in WSNs.

2.8.1.6 Sleep-Awake (Duty-Cycle) Based Routing Protocols

It focuses on minimizing energy consumption without compromising network connectivity and performance. (Jeyakarthic & Selvakumar, 2024) presented protocols that operate based on cycles of "sleep" and "awake" to reduce energy consumption, which is essential in WSNs where sensors traditionally operate on battery power.

2.9 Machine Learning for Energy-Efficient Routing Protocols in WSN

(Ullah Khan et al., 2024) Provide details about a revolutionary technology that has come to combat the energy-efficient routing protocol (E-ER-P) challenges. WSNs components are of The properties are heterogeneous. Among the critical components in WSNs, routing protocols are essential for optimizing energy consumption during data transmission. While traditional routing protocols offer fundamental solutions, they cannot often adapt to the changing and unpredictable nature of WSN environments.

Wireless Sensor Network needs an advanced ML-based routing protocol to address dynamic changing environments. WSNs need an adaptive method to overcome these issues. (V. Singh & Rastogi, 2019) proposed ML-based approaches that can address the limitations of traditional routing methods by leveraging the ability to learn from adapt to changes in real time. Reinforcement Learning based routing protocols can address these issues by adapting dynamically.

2.10 Reinforcement Learning for Energy-Efficient Routing in WSNs

Reinforcement Learning (RL) can be described as a subdivision of Machine Learning (ML) that is concerned with training through interaction within the working space. This process is undertaken through the assimilation of new information to execute tasks and or make changes as required by the system or the given environment. The latter may provide varying feedback, including positive rewards, negative rewards, or none at all. Such gradual interactions with the environment enable an agent to learn the best possible behavior to accumulate rewards over time. Reiterating on the approach, it uses rewards and focuses on the last result, thus focusing on the problem at hand instead of theoretical models, hence making it appropriate for the problem at hand (Kumar et al., 2024).

The incorporation of reinforcement learning (RL) for energy-efficient routing in wireless sensor networks (WSNs) addresses some of the most pressing challenges faced by IoT-enabled applications, including battery longevity and energy efficiency. The dynamic properties of WSNs can not adapt to the energy-constrained environments. In IoT applications, which drive critical applications such as smart infrastructure, traffic management, and security systems, ensuring prolonged network lifetimes and extensive sensor coverage is crucial. However, traditional routing solutions often struggle with real-world applicability, due to their complexity and lack of scalability, and they can not achieve energy-efficient routing.

RL-based routing, which is a dynamic and adaptive routing protocol, can be developed to optimize energy usage across sensor nodes and ensure energy-efficient WSNs routing. WSNs, by their nature, are resource-constrained and dynamic. This technique identifies the most energy-efficient transmission path from the source node to the sink node by evaluating the energy status of every intermediary node. These RL algorithm in WSNs learns and improve over time using a reward function designed to adjust energy consumption and data transmission effectiveness, achieving an intelligent decision-making process that prioritizes network longevity and reliability (Bhimshetty & Ikechukwu, 2024).

Because RL only optimizes the system's actions and feedback, it allows all possible outcomes to be achieved by balancing interactable parameters. This reward function can evaluate the performance of each node in the network. The design of RL even allows it to be deployed in edge cases, such as environments that are noisy and constrained. Unsupervised Reinforcement Learning (RL) overcomes sequential decision-making issues by feedback from the environment in an interactive way. In theory, RL performs exceedingly well in complicated video games where many mistakes can be made, but in reality, such mistakes are not practical (Glanois et al., 2024). So, it can be observed that WSNs can benefit greatly from Reinforcement Learning to improve their performance, increase reliability, and enhance their adaptability. However, effective implementation of RL requires addressing issues of resource constraints and timeliness.

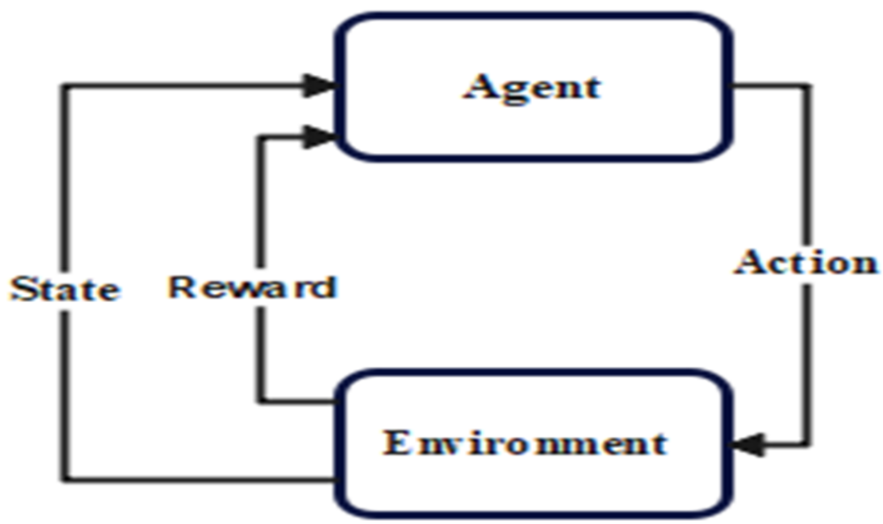

Figure 2.4: Reinforcement Learning Agent Interactions (Georgeon et al., 2015)

2.11 Datasets Scenarios

This study will employ Q-Network (DQN reinforcement) learning mechanisms to address energy-efficient routing protocol challenges in WSNs. Deep Q-Network (DQN) reinforcement algorithms provide a transformative role in solving complex decision-making problems, especially in dynamic and resource-constrained areas such as WSNs. DQNs combine the rules of reinforcement learning with deep learning, enabling intelligent agents can learn optimal actions directly from high-dimensional input states without the need for pre-defined rules or complete knowledge of the environment.

Our suggested protocol is simulated under the consideration of some hyperparameters. These hyper-parameters involve parameters such as the number of network nodes in total, the number of clusters, the permitted time to transmit data packets from a point to another, the initial energy of nodes, earning rate co-efficient (alpha), the energy consumption in various scenarios, the maximum distance in length and width between nodes, simulation length, and the number of times that the simulation executes (epoch). Reinforcement Learning for Wireless Sensor Networks will use the wsn-indfeat-dataset Public datasets for simulation purposes. These datasets are synthesized by using simulation tools like NS3 or MATLAB.

2.12 Related Works

Traditional wireless sensor network routing protocols have provided fixed techniques. However, in many applications involving a large number of sensor nodes, assigning unique global identifiers to each node is often impractical. Wireless Sensor Networks are embedded in their environments. Furthermore, sensor nodes usually determine their locations randomly within such environments, and it is hard to identify particular nodes for particular communication or commands. To overcome these challenges, routing protocols utilize data aggregation and guide traffic based on the aggregated results to enhance network effectiveness and conserve energy. (Bhimshetty & Ikechukwu, 2024) suggested a technique that enhances the lifespan and strengthens the scalability of IoT networks by balancing energy consumption evenly and distributing energy in devices. (Alsalmi et al., 2024b) proposed algorithm that utilizes a Reinforcement Learning (RL) framework for finding the best transmission paths by promoting nodes to relay data of local energy data of local energy is still a key consideration factor. The proposed algorithms are executed with the goal of facilitating energy-aware routing as well as throughput maximization, thus addressing the above-stated problems founded on RL-based solutions. It determines the best transmission paths by motivating nodes to share information of local relevance, while accessible energy is a key consideration. Suggested reference algorithms rely on the packet delivery ratio (PDR), packet delivery time, and energy efficiency. (Arafat et al., 2024) proposed a way of measuring the time for which a path will remain operational, termed link lifetime estimation (LLE), to improve packet routing in VANETs. Benchmark algorithms based on packet delivery ratio (PDR), packet delivery time, and energy efficiency were proposed by (Okine et al., 2024). (Yun & Yoo, 2021) Proposed RL-based energy-efficient routing protocols that are specific to finding the global optimum route, reducing overall energy consumption, and extending the WSN's lifetime. The approach involves the degree of possible data collection from adjacent nodes to determine the routing path. (Ghamry & Shukry, 2024) suggests a multi-objective deep reinforcement learning-based clustering and routing strategy, the proposed approach dynamically divides the network into unequal clusters based on sensor node data loads, reducing premature network failures.

(Suresh et al., 2024c) proposed Federated Deep Reinforcement Learning, which allows rapid and decentralized decision-making and is the concept of adaptive routing for dynamic network conditions. (Abadi et al., 2022) proposed the utilization of reinforcement learning for optimizing energy management via the network. (Keerthika & Berlin Hency, 2022) proposed Reinforcement-Learning-Based Energy Efficient Optimized Routing (RLER). A framework consisting of three basic algorithms—coverage gap identification by using Delaunay triangulation, dynamic node placement with optimized cluster positioning, and an adaptive tree-based routing protocols that use a Minimum Spanning Tree to enhance network coverage and energy efficiency. Efficiency and performance in multi-access edge environments were proposed by(Sivadasan & Nagarajan, 2025).

Table 2.2: Related Works summary of this study

Author	Methods	Sleeping Scheduling	Q-Learning	Periodic Battery Monitoring	Limitation
(Suresh et al., 2024)	FDRL	No	Yes	No	Improves energy efficiency with only a 15% improvement compared to existing methods.
(Arafat et al., 2024)	ARMR	No	Yes	No	The lack of explicit energy-saving mechanisms, such as clustering or duty-cycling.
(Abadi et al., 2022)	RLBEEP	Yes	Yes	No	Does not explicitly address the trade-off between energy efficiency and other parameters like average reward.
(Bhimshetty & Ikechukwu, 2024)	DQN	No	Yes	No	Limited discussion about the trade-offs between energy consumption and other network metrics, such as latency or packet loss.
(Alsalmi et al., 2024)	DRL	No	Yes	No	Does not thoroughly explore trade-offs between energy efficiency and other network performance metrics like latency or packet delivery ratio.
(Okine et al., 2024)	MADRL	No	No	No	Limited adaptability in their effectiveness.
(Yun & Yoo, 2021)	Q-DAEER	No	Yes	No	Lack of coordination between routing decisions and aggregation potential reduces its effectiveness.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Overview

This section outlines a step-by-step procedure and methodologies employed in the development, training, and evaluation of an optimum routing policy directed towards achieving maximum energy efficiency and overall network functionality. It represents a systematic method for designing, training, and assessing a guide routing aimed at advancing energy efficiency and the whole network performance and limitations of WSN routing protocols. It starts by initializing a simulation-based WSN configuration using the NS-3 simulator, simulating different network conditions to best reflect real-world operations. These simulations produce a large dataset that includes important parameters along with energy consumption, throughput, packet delivery ratio (PDR), latency, and inter-node distances.

When preprocessed, this data serves as the foundation for calculating the reinforcement learning environment in which the routing problem is considered as a Markov Decision Process (MDP). A Deep Q-Network (DQN) agent is trained to learn energy-aware routing decisions by assessing network status and incentive signals for reducing energy consumption and enhancing packet transmission rates. Finally, our proposed protocol is analyzed for performance and compared against the baseline protocols, such as RL-based routing protocol methods, e.g., RLBEEP, to verify its stability under various circumstances.

Implications of the present study are vast in numerous areas that utilize WSNs. These application areas are smart automation systems, healthcare monitoring, industrial wireless sensor networks, remote telemetry, wearable technologies, energy management, and smart agriculture. This section presents the research methodology adopted to attain the described objectives. It outlines the selected research design and framework, data collection activities, model selection process, and the evaluation metrics employed. To address the exploratory nature of the research questions, the chapter also assesses the specific methods and models in developing the proposed routing protocols for WSNs.

3.2 System Model Architecture

Our model aims to simulate a practical, cluster-based wireless sensor network where sensor nodes are randomly distributed over a specific area. Each node operates with a limited energy supply, limited processing power, and wireless communication capability. These nodes are organized into clusters, each overseen by a cluster head that gathers data from its members and forwards it to the base station or sink. It employs a multi-hop communication strategy, allowing data to be conveyed through several intermediate nodes, thereby conserving transmission energy and expanding network coverage.

The model incorporates dynamic factors such as node movement, fluctuating channel quality, and evolving network topology. Essential parameters like residual energy, node distance, link quality, queue status, and throughput are constantly followed to support intelligent routing. RL agent integrated within each node or cluster head uses these parameters as part of its state input to determine energy-efficient routing paths that enhance both network lifespan and data delivery efficiency.

3.3 Dataset Generation and Description

The nodes are systematically arranged into groups, and each group is led by a cluster head that collects information from its members and forwards it to the base station or sink. It follows a multi-hop communication strategy, where the data transmission is enabled and delivered through several intermediate nodes, thereby conserving transmission energy and expanding network coverage. To perform a simulation, each node transmitted data packets at regular intervals while key network metrics were continuously recorded. These different attributes provide the reinforcement learning model with a robust view of the network, enhancing it to make intelligent and dynamic decisions. This dataset forms the foundation for training the Deep Q-Network (DQN) and benchmarking its effectiveness against traditional routing approaches.

The dataset used in this study was generated using the NS-3 simulator and exported in CSV format. To comprises approximately 49.99 MB in size and contains over 1.5 million records, making it computationally intensive and challenging to process on standard computing systems with minimized processing capabilities. To achieve the objectives of this study on energy-efficient routing protocols, a selected subset of the dataset comprising 500,000 samples was selected for preprocessing. As described in Figure 3.1, this subset implies simulations involving 1,000 node instances with a traffic rate of 100 packets per second. Below is a sample of a Python script executed in the NS-3 simulation environment.

```
   Node_energy_levels  Residual_energy   Latency  Network_lifetime  Reward
0            0.654111         0.318327  0.704747          0.040160   0.485
1            0.979667         0.323583  0.840606          0.626227   0.730
2            0.021111         0.080210  0.597980          0.976438   0.420
3            0.966889         0.933958  0.779495          0.042978   0.040
4            0.851667         0.192070  0.696970          0.946055   0.280

   Hop_count_to_sink  Throughput    PDR  Next_hop_selection  Node_ID  \
0           1.000000    0.350544  0.125               0.877    0.000
1           0.578947    0.806733  0.525               0.068    0.001
2           1.000000    0.085144  0.700               0.487    0.002
3           0.000000    0.230267  0.850               0.224    0.003
4           0.105263    0.823356  0.400               0.376    0.005

   Node_distance  Timestamp  Queue_length  Neighboring_node_states  \
0       0.926104        0.0          0.62                 0.555556
1       0.047390        0.0          0.04                 0.777778
2       0.958835        0.0          0.44                 1.000000
3       0.570000        0.0          0.04                 0.111111
4       0.758273        0.0          0.74                 0.222222
```

Figure 3.1: Dataset preprocessing sample

The following table outlines each feature, its data type, and a brief description of the dataset we have generated using the NS-3 simulator.

Table 3.1: Summary of feature description

Sno.	Features name	Description	Data types
1	Timestamp	Records the exact time at which a data entry or sensor activity was logged during simulation monitoring.	float64
2	Node_ID	Unique identification is assigned to each node in the WSN to distinguish and monitor individual sensor activities.	float64
3	Node_energy_levels	The current energy level of a node, often expressed in Joules or Percent,	float64

		represents the power available before depletion.	
4	Residual_energy	The remaining energy after performing routing, communication, or sensing tasks.	float64
5	Energy_consumption	Total energy used by a node in transmitting, receiving, and processing packets within a given time frame or routing session.	float64
6	Link_quality	A metric representing the quality of the communication link between two nodes.	float64
7	Channel_condition	Indicates the state of the communication channel.	object
8	Node_distance	Physical or logical distance between nodes or from a node to the sink.	float64
9	Queue_length	Number of packets waiting the node's buffer for transmission.	float64
10	Hop_count_to_sink	The number of intermediate nodes a data packet traverses before reaching the sink node.	float64
11	Neighboring_node_states	The operational status of adjacent nodes, crucial for making informed next-hop decisions.	float64
12	Next_hop_selection	Indicates the chosen neighboring node for forwarding the packet.	object

13	Action	The decision taken by the reinforcement learning agent, based on the current state observation.	object
14	Reward	Feedback signal given to the RL agent based on the outcome of an action.	float64
15	Successful_packet_delivery	Binary or count value representing whether the transmitted packet reached its destination successfully.	float64
16	PDR	Ratio of successfully delivered packets to the total sent.	float64
17	Throughput	The amount of data successfully transmitted over the network in a given period.	float64
18	Latency	Time taken for a packet to travel from source to destination.	float64
19	Network_lifetime	Duration until the first or last node dies due to energy depletion.	float64

3.3.1 Simulation Environment

To realistically model the operational dynamics of Wireless Sensor Nodes under varying network conditions, the simulation environment should be designed properly. For our study, the simulation setup environment is based on NS-3 (Network Simulator 3), a discrete-event simulator modified for studying networking protocols and communication systems. A comprehensive platform to model and simulate wireless sensor networks (WSNs) with various protocols and configurations we selected for our study is the NS-3 network simulator. The model includes dynamic variables like node mobility, time-varying channel quality, and developing network topology. Crucial factors such as residual energy, node distance, link quality, queue status, and throughput are constantly monitored to facilitate intelligent routing. RL agent incorporated into every node or cluster head, utilizes these parameters as components of its state input to determine energy-efficient paths that achieve maximum network lifetime and data delivery efficiency.

For enhancing the training and evaluation of the suggested adaptive energy-efficient routing protocol, an extensive dataset was created using the NS-3.36.1 simulator. The simulation environment depicted a wireless sensor network composed of various nodes systematically organized in a clustered implementation with AODV as the base routing protocol. A variety of network conditions, such as different traffic rates, number of nodes, and changing energy levels were modeled to represent dynamic operating conditions. Various tools, package libraries, and ML structures, used for various purposes in the simulation setup, are mentioned below. This includes various tools, package libraries, and ML structures used for various purposes as mentioned below.

Ubuntu 22.04: Ubuntu 22.04 serves as the primary working environment for implementing the model architecture in the development of an adaptive energy-efficient routing protocol using reinforcement learning for wireless sensor networks. The Linux-based operating system is chosen for its stability, extensive support for networking tools, and compatibility with essential machine learning and simulation frameworks.

NS-3: As a discrete-event network simulator, NS-3 (version 3.36.1) provides a highly flexible environment for simulating WSNs, allowing researchers to design, test, and optimize the proposed RL-based routing protocol before real-world deployment. It facilitates the modeling of sensor nodes, energy consumption dynamics, communication protocols, and network topologies, enabling precise evaluation of energy efficiency and adaptive behavior. NS-3's component-based design supports the integration of RL algorithms, where nodes can learn optimal routing strategies based on energy consumption, network traffic jams, and link reliability. Additionally, NS-3 enables performance analysis based on various performance measures.

Kaggle: Kaggle environments supply GPU and TPU access and are thus suitable for managing the computational demands of Deep Q-Network (DQN) and Hierarchical DQN models in this study. It provides an instant working space for model training, execution, and dataset preprocessing evaluation. It also supports Python and significant libraries such as Tensor-Flow, PyTorch, Numpy, and Pandas, to create an efficient workflow for adopting and experimenting with the adaptive routing protocol.

Python: Python is also a highly effective and stable programming language, backed by an extensive ecosystem of machine learning tools. The framework utilizes libraries like TensorFlow and PyTorch to implement the reinforcement learning (RL) agent, which employs Deep Q-Networks (DQNs) or policy gradient approaches to find energy-efficient routing approaches that minimize energy consumption while maintaining network stability. The flexibility of the language simplifies the implementation of neural networks for estimating Q-values or policy-based actions, enabling real-time modification of dynamic network parameters such as sensor battery levels, connection reliability, and structure configurations. Such ability maintains the system in an adaptable state to evolving operational demands in wireless sensor settings.

NumPy: NumPy is a fundamental package for numerical data and streaming data processing structures. It's an ideally designed array-processing that offers efficient processing of the reinforcement learning (RL) paradigm. By incorporating NumPy into the RL framework, the protocol wisely adjusts routing pathways, remarkable a balance between exploring new routes and via established energy-saving paths. This combination guarantees prolonged energy conservation with high network reliability.

Jupyter Notebook: Jupyter Notebook was an agile environment conducive to iterative testing and visual inspection. Being dynamic and interactive, it was easy to incorporate the application of reinforcement learning (RL) methodologies, including Deep Q-Network (DQN), in simulated network environments to sustain collaborative simulations that integrate algorithmic design and system implementation modeling.

Scikit-Learn: Scikit-Learn is crucial in streamlining the processes of reinforcement learning. (RL) framework through the solution of data preprocessing and additional machine learning Functions. Although the primary RL algorithms, such as Q-learning, might be developed independently or incorporated through dedicated libraries, Scikit-Learn facilitates streamlined feature engineering and data normalization processes. This ensures high-quality training data inputs by refining raw network parameters (e.g., residual energy, transmission distances) into standardized, meaningful state representations, which are vital for robust RL model performance.

Matplotlib: Matplotlib plays a pivotal role in the visualization and validation of the model's behavior and performance metrics. Throughout the training phase, the library is utilized to create visual representations of learning curves, reward progression trends, and convergence behaviors of the reinforcement learning agent (such as Q-learning or Conservative Q-Learnings), facilitating simulation observation of the training dynamics.

In our study, we employed the creation of a simulated or lab-based Wireless Sensor Network (WSN) environment where sensor data was generated under specific, pre-defined conditions. These conditions included: Varying routing protocols, Consistent traffic patterns, a fixed number of nodes, and uniform data size. This approach allowed us to conduct replacement experiments, making it easier to test and validate the performance of different routing protocols under identical conditions. It allows us to avoid the risks and costs associated with deploying and experimenting in real-world WSNs, which can often be located in harsh or inaccessible areas.

Furthermore, this simulation enabled us to generate large, labeled datasets suitable for training Reinforcement Learning (RL) models. However, we acknowledge that simulated data might not perfectly capture the complexities and uncertainties of real-world WSNs, such as noise, interference, dynamic environmental changes, and unpredictable node behavior. A sample CSV file for NS-3 data generation and collection is shown in Figure 3.2 below.

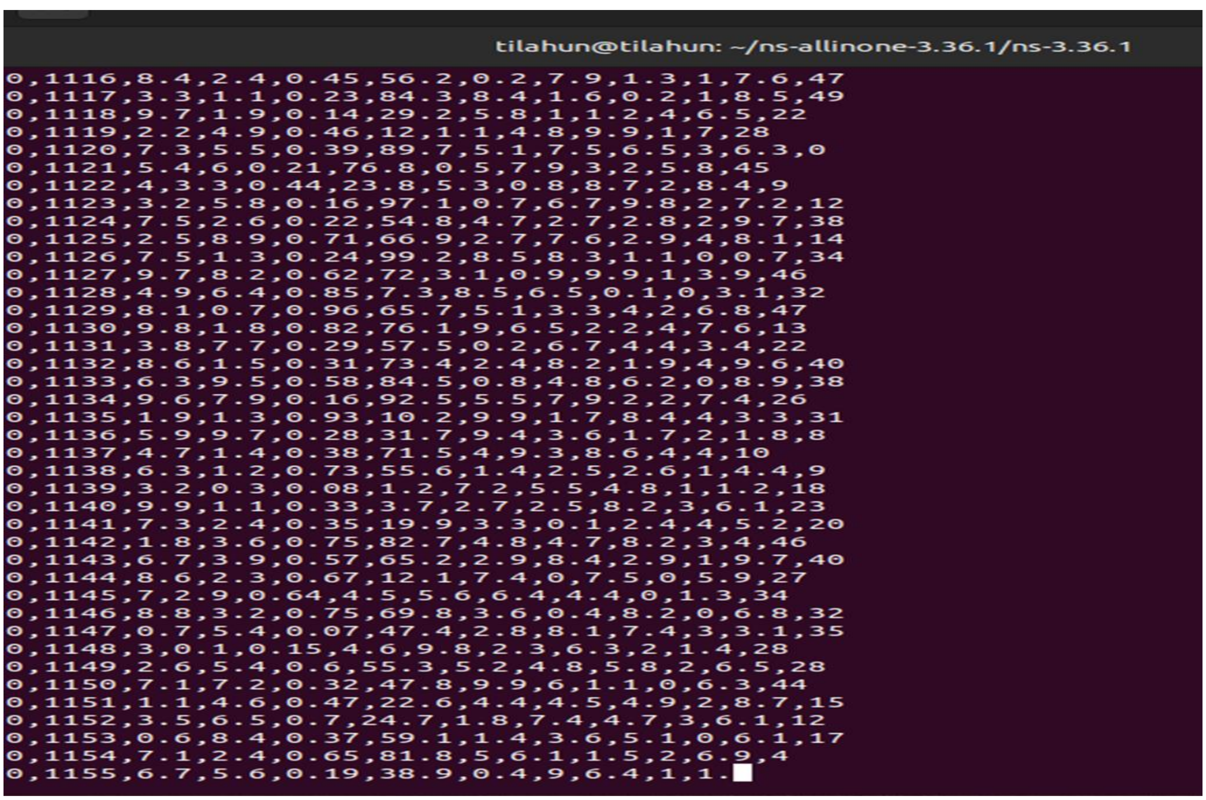

Figure 3.2: A sample CSV file dataset generated by NS-3

After mounting Google Drive and loading the dataset, we initiated the preprocessing steps essential for implementing a reinforcement learning and machine learning-based adaptive energy-efficient routing protocol for wireless sensor networks (WSNs). We first examined the dataset by displaying its column names and the first few records to understand its structure. Next, we handled missing values in the numeric columns by filling them with their respective means. Lastly, we ensured the dataset's integrity by detecting and eliminating duplicate rows, preparing it for reliable training of the RL model, as illustrated in Figure 3.3 below:

```python
# Step 1: Import Libraries
import numpy as np
import pandas as pd
import tensorflow as tf
from tensorflow.keras import layers, Input, Model
from collections import deque
import random
import matplotlib.pyplot as plt
from sklearn.metrics import classification_report, confusion_matrix
import seaborn as sns

# Step 2: Load and Prepare Data
df = pd.read_csv('/kaggle/input/aeerprl-normalized-ds/Aeerprl_normalized_dataset.csv')

# Updated features
state_features = [
    'Node_energy_levels', 'Residual_energy', 'Link_quality', 'Channel_condition',
    'Node_distance', 'Queue_length', 'Hop_count_to_sink', 'Neighboring_node_states',
    'PDR', 'Throughput', 'Channel_condition_numeric'
]
action_column = 'Next_hop_selection'
reward_column = 'Reward'
energy_consumption_column = 'Energy_consumption'
```

Figure 3.3: RL Model Implementation

3.3.2 Network Parameters and Topology

In the course of the present study, the network topology structure is usually achieved through a cluster-based model, with the nodes positioned in a two-dimensional area and communicating wirelessly. In this scenario, specific nodes are utilized as cluster heads with the responsibility of collecting and relaying data to the base station, thereby minimizing energy consumption and enhancing the network scalability. The key parameters characterizing the network involve the number of nodes per unit area, along with node mobility (i.e., whether nodes are stationary or mobile), node energy states, and routing configurations. Additionally, factors such as channel conditions (signal quality), latency, packet delivery ratio (PDR), link quality, and node distance from sink nodes are optimized. These parameters are carefully configured in the simulation environment NS-3 to model realistic network behavior and test the robustness of the proposed protocol under varying conditions.

3.3.3 Features and Data Attributes

The characteristics and data attributes selected in our research have been identified with caution to capture the dynamic performance and behavior characteristics of energy-operating wireless sensor networks limitations. All these features together form the global state of the network and play a crucial role in facilitating the reinforcement learning agent's ability to make well-informed and effective routing decisions. This full range of capabilities allows the learning model to respond properly to the changing network conditions and stimulates energy throughout the system.

3.4. Data Cleaning and Relevant Feature Selection

Data cleaning and the selection of suitable features are necessary to enhance the efficiency and correctness of the reinforcement learning-based routing protocol. In our study, raw dataset values derived from NS-3 simulations are initially enhanced by completing missing data, repairing contradictions, and duplicates. Numerical variables are normalized by scaling strategies to maintain consistent value ranges, thus fostering greater convergence of the model.

All the categorical variables are converted to numerical representations using appropriate encoding methods. To identify the most important factors that impact routing decisions, such as residual energy, link quality, node distance, queue length, and throughput feature selection methods like mutual information and correlation analysis are used. This filtering aids in data reduction complexity, speeding up the training process, and improving the learning model's ability to Best routing choices depend on focusing on crucial predictive attributes.

3.4.1 Handling Missing Values

In this study, lost data, usually caused by sensor node failures, lost packets, or unstable connectivity, is resolved by proper allocation methods like mean, median, or forward/backward fill based on time-series trends. Noisy entries, which can be caused by the device Anomalies or external interferences, are detected through statistical approaches, for example, the z-score or interquartile range analysis, and corrected by employing a smoothing technique or removing outliers. This first phase ensures that the reinforcement learning model is trained on data that is free from inaccurate information, thereby making it more effective at making the best routing decisions in dynamic network environments.

3.4.2 Feature Selection

The major feature selection procedure guarantees that just the most effective and significant data features are employed, reducing computational complexity while enhancing the efficacy of the reinforcement learning model. The most significant features typically are energy consumption, channel condition, next hop selection, and remaining energy, since these directly influence routing decisions and energy efficiency. Picking these traits carefully, our model can commit to learning the correct Routing mechanisms aimed at reducing energy consumption without compromising reliable communication. Handling missing values or irrelevant attributes avoids overfitting and maximizes the protocol's scalability to large-scale WSN environments. Effective attribute selection mechanisms are essential to enhance the reinforcement learning process and minimize computational complexity. Some popular attribute selection methods:

Filter Method: This approach focuses on filtering simulated data to choose and utilize only the most important features that significantly impact routing performance and energy consumption. SDFM utilizes statistical and heuristic approaches to eliminate redundant or irrelevant data attributes and thereby optimize the decision-making for the reinforcement learning model. By prioritizing features like node energy levels, network topology dynamics, link quality, and traffic load, the method enhances the protocol's adaptability to dynamic network conditions efficiently. Additionally, SDFM minimizes computational complexity to allow the reinforcement learning agent to converge faster while maintaining high accuracy in selecting energy-efficient routing paths. This approach can also optimize the protocol's overall energy efficiency, but also prolongs the network lifetime, which is optimally applicable to energy-limited wireless sensor networks. Some of the well-known filtering techniques include:

Correlation-based filtering: Seeks to select and utilize only the best characteristics by examining the relationships between simulation data variables. An extremely high correlation between features often indicates redundancy, thus making it easy to eliminate less important variables to minimize the dataset size. This data flow reduces computational overhead while keeping vital parameters, i.e., residual energy, node connectivity, and next hop determination, intact for effective decision-making.

Mutual Information: A method used to measure the dependency between variables, selecting attributes with the best predictive power for target variables while minimizing duplication. By analyzing the important information between routing features, such as residual energy, channel condition, and next hop selection, and the desired routing objectives, such as energy efficiency and reliability, MI filtering attains the selection of attributes that significantly influence decision-making. These techniques not only simplify the model complexity but also enhance the computational efficiency of reinforcement learning algorithms, making them suitable for resource-limited WSNs. Leveraging MI selects aligns to create adaptive protocols that dynamically adjust routing decisions based on the most vital and energy-efficient features, ultimately prolonging network lifetime and maximizing performance.

Variance Thresholds: It validates that only valuable and important attributes are utilized for training our model, RL. Threshold-based attribute selection techniques, e.g., treatment of attributes with low Variance, aid in eliminating redundant or undesirable data, thereby reducing computational demands and improving model performance.

Wrapper Method: It provides a disciplined mechanism for generating encapsulation and simulation data analysis to replicate real WSN situations as closely as possible. It promotes the integration of diverse data sources, reveals differing environmental factors and network topologies, and node behaviors, into the reinforcement learning framework. By encapsulating simulation data with metadata and contextual data, the approach enables more significant assessments of residual, network lifetime, and decision-making steps. Classic examples of wrapper approaches include:

Forward Selection: This method is utilized for the purpose of enhancing the capabilities of machine learning models by filtering the most relevant attributes for training models. It can be applied during the steps of data simulation to choose the significant parameters that affect energy consumption and performance of networks. This method starts with an empty set of attributes and incrementally adds the most significant feature, determined based on a pre-determined criterion such as precision or minimization of error. This goes on until a subset of features is achieved that can then be used to train the reinforcement learning model.

Backward Elimination: This method involves an iterative procedure where traits are initially included within the framework and subsequently eliminated depending on whether they contribute to the prediction accuracy. In energy-efficient routing, it is used to find the most useful factors-such as node energy level, channel condition, and next hop selection, which influence routing decisions.

Recursive Feature Elimination: It is an efficient attribute filtering technique widely utilized in machine learning to reduce model overfitting and improve performance by iteratively addressing the less important features. The process entails fitting a model, then ranking the features in terms of importance, and recursively removing the least important ones until the enhanced subset is achieved. By decreasing the dimensionality of the data, RFE assists in attaining more efficient and Reinforcement learning-based computationally viable routing protocols, guaranteeing the protocol and accommodating varying network conditions using little energy.

Embedded Method. This method collaborates with feature filtering right in the learning process steps, enhancing both the model's performance and its computational efficiency. They work by including attribute filtering methods in the training of a machine learning algorithm, like reinforcement learning. It enables the simultaneous optimization of feature subsets and model parameters. Typical examples of embedded methods for the feature selection process are:

3.4.3 Encoding the Labeled Data

This method includes the creation of a simulation platform that precisely represents the network's modifying behavior, working with annotated datasets to cover a range of network situations, such as residual energy, node energy metrics, and energy usage. These categorized datasets enable the reinforcement learning agent is tasked with identify the optimal and non-optimal routing paths, encouraging its conduct to facilitate energy-conserving choices. By incorporating the categorized data during the simulation, the protocol can learn step by step from its environment, adapting in real-time to changes intended to reduce energy consumption.

3.4.4 Feature Scaling and Normalization

These scaling and normalization procedures are essential to data structuring for Reinforcement learning, especially in problems with attributes that have different units or sizes. In wireless sensor networks (WSNs), parameters may encompass large variability in magnitude. If this difference is not considered, features with high numeric values can dominate others, resulting in unbalanced learning results. To amend this, techniques such as Min-Max Scaling normalize all feature values to a uniform range, often 0 to 1, so that they can contribute equally in model training.

3.4.5 Dimensionality Reduction

This method is used to improve the efficiency and performance of the reinforcement learning model. Since the dataset is feature-rich, there is a need to decrease the input variables, decrease the computational expenses, and reduce the overfitting risk. To achieve Principal Component Analysis (PCA) is employed to choose the most significant features while addressing aspects that are redundant or of lesser strength. By focusing attention on the most relevant state parameters, a reinforcement learning model can converge more quickly and make more precise determinations in the variable and resource-constrained environment of Wireless Sensor Networks.

3.5. Reinforcement Learning Framework

The decision-making process in routing is formulated as a Markov Decision Process (MDP), in which the sensor node is a learning agent that communicates with the surrounding network setting in which to develop effective routing mechanisms over time. Such an environment demonstrates the changing nature of wireless sensor networks, triggered by dynamic energy levels, channel conditions, and network traffic. Each agent depends on its present state-characterized by relevant network attributes, and takes an action based on a policy derived from a Deep Q-Network (DQN). The agent is rewarded depending on the success of the chosen action and consequently receives positive feedback for reducing energy usage, promoting packet delivery, and prolonging network lifespan. This reward policy enables the protocol to learn dynamic network circumstances.

3.6 Research Design

Our research design will try to follow an orderly approach so as to enable the formulation and Assessment of an effective protocol. To begin with, the research will encompass the examination and analysis of fundamental routing and energy efficiency issues in WSNs, such as limited battery life and dynamic network topology. The reinforcement learning system will then be built, where nodes act as active learners who acquire best practices in routing by engaging with their environments. The proposed design will prioritize the integration of energy-efficient parameters, including residual energy and channel condition, and hop selection into the reward function to promote energy-efficient operation. The protocol will be validated using analytical modeling and simulation runs with small-scale test bed data to ascertain both durability and flexibility. Key performance indicators, such as energy consumption, network lifetime, and packet delivery ratio, are to be measured to compare the proposed protocol compared to conventional methods, ensuring its utility in various scenarios.

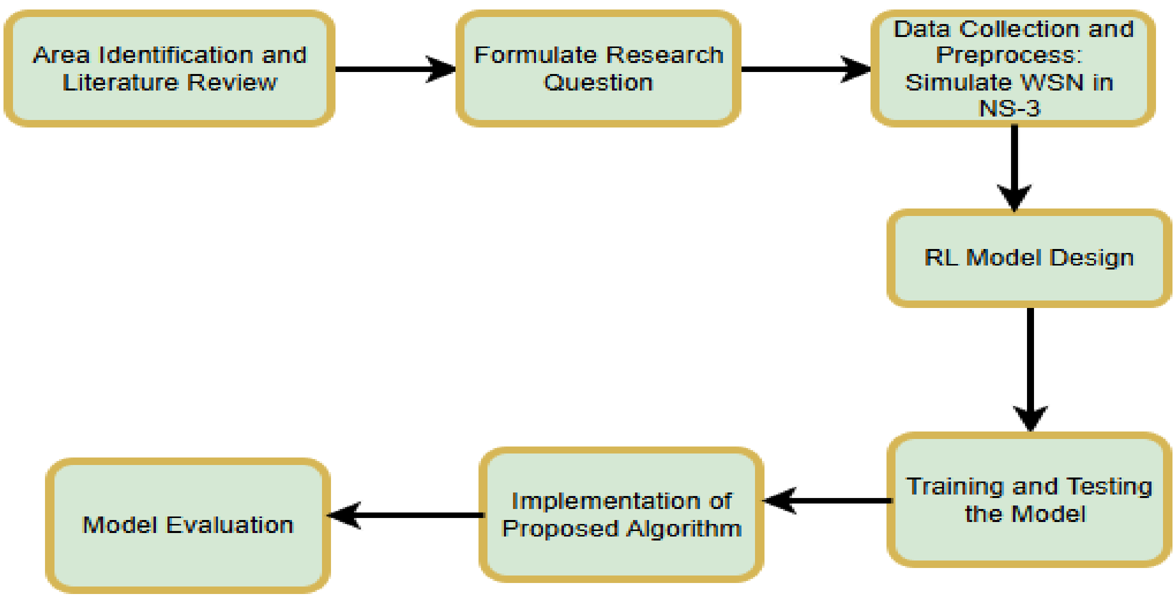

Figure 3.4: Proposed methodology of the Study

3.7 Parameter Tuning

It serves as a way of tuning diverse parameters of reinforcement learning algorithms to enhance their performance in managing energy consumption and enhancing network lifespan. Effective parameter tuning can significantly affect routing choices at sensor nodes, enabling them to cope with evolving environmental circumstances and operational requirements. For instance, intensive function-related variables in RL are crucial; they guide the agent in assessing the efficacy of its interventions on energy consumption, balance of load, and interconnectivity quality. By adjusting these parameters, the protocol can learn the best routing paths more efficiently, which minimizes the energy usage while maximizing the data delivery reliability.

Furthermore, incorporating habits such as sleep planning and changing path selection even more increases energy efficiency, so that nodes conserve energy during idle periods and dedicate traffic uniformly throughout the network. Last but not least, fine-tuning of parameters enhances the immediate operation of WSNs but also engages in long-term sustainability and efficiency in various contexts (Malik, 2023).

3.8 Model Selection

This phase deals with the selection of the most appropriate algorithm in the model, like approaches involving reinforcement learning or cluster-based approaches, to maximize energy efficiency, network longevity, and performance while addressing the problems of sensor nodes (Sharma et al., 2022). The model selection entails the identification of superior reinforcement learning algorithms, including Q-learning and Deep Q-Network (DQN), or Policy Gradient approaches, based on the complexity of the WSN environment. Classical models like Q-learning are simpler and needed for smaller networks with small state-action spaces, whereas DQN and other deep reinforcement learning models show greater efficacy in large and dynamic networks because they can cope with high-dimensional state spaces. The choice is also determined by energy constraints, the computational capacity of sensor nodes, and network scalability requirements.

In addition, hybrid frameworks that blend reinforcement learning with heuristic or optimization mechanisms can be investigated to improve reliability and energy efficiency. Apart from this, the model should also account for both the distributed and dynamic WSN nature, enabling nodes to learn independently and adaptively optimal routing decisions by themselves and minimize communication overhead. Following in pursuit of our objective, we identified and evaluated the following candidate models.

Q-Learning: Deep Q-Learning, a model-free reinforcement learning approach, allows wireless sensor networks (WSNs) to make optimum routing decisions through learning from environmental feedback, as demonstrated in the (Nandyala et al., 2022) proposed Q-learning based on the policy-aware routing (QTAR) protocol, improving energy efficiency and minimizing latency, and prolonging network lifetime by intelligently selecting next-forwarder (NF) nodes.

This is an off-policy reinforcement learning algorithm that enables nodes in a sensor network to identify optimal routing paths by iterative optimization of Q-values from reward feedback of the actions taken within the setting requires a model specifically designed to meet the unique challenges of wireless sensor networks, such as limited energy resources, dynamic topology, and variable data traffic. The most critical of model choice considerations are the action and state design, spaces, which are designed to gather interesting network variables like residual energy, link quality, and hop count.

Deep Q-Network (DQN): DQN model promotes choosing in WSNs by leveraging deep Reinforcement learning to improve routing procedures, making ARCMRP-DDQN change adaptive multicast routes based on network dynamics, and therefore improving energy efficiency, packet delivery, and latency (Selvakumar & Jeyakarthic, 2024).

It collaborates Q-learning with deep neural networks to approximate the Q-value function, allowing for efficient decision-making in high-dimensional state-action spaces. For WSNs, selecting the best DQN model is a trade-off between energy efficiency and computational complexity. Lightweight network structures are chosen to reduce energy consumption while offering enough capacity to receive the information on changes in the network, for example, node energy levels, link quality, and traffic patterns. With a properly calibrated DQN, the routing protocol can intelligently choose energy-saving routes, prolonging network lifetime and optimizing overall reliability.

Policy Gradient Methods: This strategy can enhance energy management in WSNs. Additionally, by optimizing the energy adaptation policy through gradient-based updates and thus improving RLMan seeks to learn about more effective ways of power consumption while adjusting to changing energy supplies, thus enhancing the overall long-term quality of service and network performance longevity (Aoudia et al., 2017).

These approaches target direct policy optimization by estimating the rise in the expected intensity in terms of the policy variables, enabling a proper investigation of high-dimensional and complex environments typical of WSNs. It may be employed to dynamically tune routing decisions depending on real-time conditions of the network, including energy levels, node density, and transmission quality.

3.9 Model Training

DQN combines Q-learning with deep neural networks to estimate the Q-value function, allowing efficient handling of huge state spaces that are typical in wireless sensor networks. DQN approach combines techniques like knowledge replay and target network improvement to stimulate training. Training a simulation model involves creating a virtual network area where sensor nodes can communicate, collect information, and adjust routing decisions based on incentives that encourage energy efficiency. In this context, the DQN model is trained iteratively, with its policy optimized to minimize energy consumption while with network performance and reliability.

Experience replay: It can be utilized in combination with the Deep Q-Network approach for WSNs to further enhance it improves its performance through the storage and reuse of past events during training. This method allows the deep reinforcement learning (DRL) agent to decouple the correlation between consecutive data samples and enhance learning stability. By re-listening to the history of state-action reward transitions, the model can optimize data transmission paths more effectively, improve the defense mechanism against attacks, and change adjustment strategies in changing network environments, providing improved energy efficiency along with security in wireless sensor networks (Hu et al., 2024).

Target Network Updates: In the context of wireless sensor networks (WSNs), the goal network updates during model training assist as a fundamental method to streamline decision-making processes in multi-objective optimization models. These updates enable the educational model can accommodate the changing and typically competing WSN goals very well, design, develop, and operate through balancing the learning process, enhancing meeting points, and stabilizing a balance among conflicting objectives such as energy efficiency, coverage, and reliability. This kind of strategy enhances the flexibility of WSN solutions to different constraints and contexts of use, allowing the promotion of bespoke and effective optimization strategies (Iqbal et al., 2015).

3.10 Performance Evaluation Metrics

Performance evaluation metrics for Wireless Sensor Networks (WSNs), such as average reward, PDR, and energy consumption, are vital for assessing network quality systematically, as highlighted in a presented model that combines the entropy weight method and AHP to address the problems of single subjective or objective evaluations, thereby providing a complete approach for maximizing WSN performance(Iqbal et al., 2015).

Energy Consumption: It represents a vital component in the architecture of Wireless Sensor Networks (WSNs), particularly concerning routing protocols that prioritize energy efficiency.

Here, we will discuss some important formulas and variables associated with energy consumption in WSNs to maximize energy efficiency.

Total Energy Consumption: Used to gauge the total energy consumed in data processing and transmission. Lower energy usage indicates more effective protocols. Minimum energy consumption indicates better protocol efficiency. The total energy consumption of each node includes energy

- ✓ Data transmission,
- ✓ Data reception,
- ✓ Sensing (if applicable),
- ✓ Idle listening and processing.

$$E_{total} = E_{tx} + E_{rx} + E_{sense} + E_{idle} \qquad 3.1$$

Where:

E_{sense} : Energy used by the sensor module during sensing.

E_{idle} **:** Energy utilized in idle mode.

Network Lifetime: It is typically described as the time until the first node depletes its energy or until the network becomes non-functional. A common formula is:

$$L_{network} = \frac{TotalInitialEnergyofNetwork}{TotalEnergyConsumptionperUnitTime} \qquad 3.2$$

Packet Delivery Ratio (PDR): This metric is crucial in wireless sensor networks (WSNs) as it is utilized to quantify the effectiveness of energy-aware routing protocols. It is an important parameter in wireless sensor networks (WSNs) used to analyze the performance of energy-efficient routing protocols. PDR calculates the proportion of packets that are delivered to the arrival point node of the total packets sent from the source node. This kind of node is utilized in assessing the strength and efficacy of interactions within the network, especially in energy-constrained WSNs.

$$PDR = \frac{NumberofPacketsDeliveredSuccessfully}{TotalNumberofPacketsSent} x100 \qquad 3.\ 3$$

Throughput: The throughput in Wireless Sensor Networks (WSNs) is one of the measures of the effectiveness of energy-efficient routing protocols. It calculates the amount of data properly transmitted through the network within an interval. In energy-efficient routing, the goal is to achieve minimum energy consumption with maximum throughput. The overall throughput in a WSN can be defined as:

$$Throughput = \frac{N_{Data}}{T_{Total}} \qquad 3.4$$

Where:

N_{Data} is the amount of data that is transferred (in bits or bytes).

T_{Total} is the total time during which the data was transmitted (in seconds).

3.10.1 Classification Metrics

Based on the preprocessing methods and model training, testing the model measures its effectiveness on unseen or new data. This is accessed through model evaluation methods or classification metrics to ensure model compatibility and measures how well the model can correctly classify data points into their respective categories. We can access our model using a variety of evaluation metrics. Energy Efficiency Metrics, network lifetime metrics, Packet Delivery Metrics, Routing Efficiency Metrics, and Reinforcement Learning-Specific Metrics are common evaluation metrics. Below are relevant metrics and their importance in this context.

Packet Delivery Ratio (PDR): It represents the percentage of successfully delivered packets over the total transmitted packets, serving as an important metric for evaluating the reliability and efficiency of the network (Venkateswara Rao & Malladi, 2021).

Latency: Latency is a critical problem, as simultaneous transmissions to optimize network capacity often result in interference, leading to high delays, which promotes effective mechanisms like interference alignment (IA) to avoid such effects and optimize overall network performance (Zibakalam & HosseinKahaei, 2012).

Hop Count: In WSNs the hop count refers to the number of intermediate nodes a data packet traverse to reach its destination, and in (Institute of Electrical and Electronics Engineers, 2013), it is leveraged to create a novel localization algorithm that minimize mapping errors and energy consumption by locally deriving the mean hop size at position-unaware nodes without depending on anchor broadcasts.

Convergence Time: The duration required for all nodes to achieve a consensus or synchronization, which can be prolonged in large-scale networks due to iterative algorithms; however, the proposed (Shi et al., 2020) MACTS model significantly accelerates this process by leveraging multi-hop communication to motivates algebraic connectivity and adjust efficiency, accuracy, and complexity.

Reward Value: The reward value in Reinforcement (RL) shows the network's property to achieve objectives such as reducing energy consumption, maximizing data throughput, enhancing reliability, and achieving real-time communication, making it a critical factor in optimizing IIoT-based communication protocols (Alem et al., 2023).

Packet Delivery Accuracy: This is related to Packet Delivery Ratio (PDR), a common metric in WSNs.

$$PacketDeliveryAccuracy = \frac{NumberofPacketSuccessfullyDelivered}{TotalNumberofPacketTrasmitted} \qquad 3.\ 5$$

Reward-Based Accuracy (Reinforcement Learning Context): In RL-based routing protocols, accuracy can also evaluate how well the learned policy maximizes rewards over time. This is often expressed as the ratio of actual reward obtained to the maximum possible reward:

$$Reward - BasedAccuracy = \frac{\sum R_{Obtained}}{\sum R_{Optimal}} \qquad 3.\ 7$$

Where:

$R_{obtained}$: Rewards attained for selecting specific routing paths (e.g., paths with minimum energy consumption or shortest delay).

$R_{Optimal}$: Maximum achievable rewards under optimal conditions.

3.10.2. Design and Development Tools

In the design and development of energy-efficient RL-based a strategic selection of tools and platforms to ensure efficient implementation, testing, and evaluation is a basic requirement. These requirements are listed as follows:

3.10.2.1 Design Tools

Design tools are essential for efficient implementation, simulation, and analysis. The choice of the right design tools is essential to the successful implementation, simulation, and evaluation. One of the tools for designing that we have used is EdrawMax for developing and building the proposed solution model.

EdrawMax: It is a flexible and accommodating design tool that has the potential to significantly improve the process of the ideas are visualizing and communicating ideas using a broad array of templates, shapes, and symbols central to the development of complicated network diagrams, flowcharts, and conceptual maps that are crucial in demonstrating the design and operation flow of the suggested protocol. Through its drag with drag-and-drop functionality and the scope for customization, we methodically plan routes, energy flow mechanisms, and reinforcement learning models, with clarity in their representation.

3.10.2.2 Development Tools

When designing and developing adaptive energy-efficient routing protocols using reinforcement for Wireless Sensor Networks, development tools are a very important aspect. These play a crucial role in adaptive energy-efficient routing protocol design and implementation by utilizing reinforcement learning for Wireless Sensor Networks (WSNs). Integrated Development Environments (IDEs) such as Visual Studio Code or Eclipse can facilitate seamless code development and debugging. For the application and testing of reinforcement learning algorithms, due to their extensive libraries, programming tools such as TensorFlow or PyTorch, which are backed by machine learning tasks, are ideal.

Simulation network simulators such as NS-3 and OMNET++ offer support for modeling and analysis of WSNs by providing insightful information regarding network behavior under varying conditions. Furthermore, hardware-in-the-loop devices, i.e., Arduino or Raspberry Pi, can be utilized to verify protocol performance in real-use scenarios. Performance-measuring tools are utilized to monitor Data transmission in combination with energy profiling tools, is important to determine the energy efficiency of the protocol. In general, the development, simulation, and testing of our designed routing protocol is made possible through these tools.

Software Tools: These are core Machine learning architectures for the use of RL algorithms to facilitate efficient training and testing of our model. The software listed is as follows:

Network Simulator (NS-3): It is an open-source general-purpose discrete-event network simulator utilized for research and educational activities. It presents an adaptive and scalable model for network protocols and simulates complex situations. Its ability to simulate low-power wireless communication protocols such as IEEE 802.15.4 and support for pluggable energy models make it the best tool for evaluating the energy efficiency of routing protocols. With its comprehensive logging and tracing mechanisms, we can analyze protocol performance, validate energy-saving mechanisms, and analyze their proposed solutions with respect to contemporary methodology, thereby ensuring the practicality of their design before implementation in real-world WSNs.

Python and C++: It is used because of its simplicity, large libraries, and frameworks that include TensorFlow and PyTorch, which are necessary to train and deploy RL models efficiently. It aids quick modeling and testing of algorithms for quicker interactions during the development phase. In another example, C++ is ideally used in delivering high performance and advanced implementation, thereby rendering it suitable for the construction of RL models in resource-scarce WSN environments. We can develop using Python and implement using C++. We can integrate ease of development with enhanced performance to achieve a strong and scalable energy-efficient routing protocol.

Jupyter Notebooks: It produces a more suitable environment that supports Python and other programming languages, enables continuous integration of code, visualization, and text description. By using Jupyter Notebooks, we can simulate WSN scenarios, implement RL methods, and analyze their performance in real time. The ability to visualize energy consumption, routing efficiency, and network reliability through dynamic plots and graphs promotes understanding and facilitates debugging.

Microsoft Office: It provides us with the necessary tools for documentation, data analysis, presentation, and project management to support the research process effectively.

Hardware Tools: The hardware tools used in this study are listed in the table below.

Table3.2: Hardware Tools

Hardware	Component Description	Role in Research
RAM	Memory for storing simulation data, algorithms, and results.	Supports large-scale simulations and real-time processing of data.
Processor	Central unit for executing reinforcement learning and network protocols.	Executes complex RL computations and sensor interactions efficiently.
Hard Disk	Storage for simulation data models and logs.	Stores extensive data from simulations for post-execution analysis.

Hardware tools for developing an adaptive energy-efficient routing protocol using reinforcement learning in wireless sensor networks may include low-power sensor nodes, microcontrollers, wireless transceivers, and energy-harvesting models to simulate real-world deployment scenarios.

CHAPTER FOUR

PROPOSED ARCHITECTURE FOR EFFICIENT ROUTING PROTOCOLS

4.1 Chapter Overview

In the previous sections, we assessed a detailed literature review, described the research methodology, and identified the limitations in the existing knowledge, as detailed in the statement of the problem. Additionally, we formulated key functional research questions to conduct our study. In our study, the introduction underscores the challenges related to energy-efficient routing in Wireless Sensor Networks and the importance of energy-efficient routing protocols. The paper starts with an introduction that underscores the challenges associated with energy consumption in wireless sensor networks (WSNs) and the need for efficient routing protocols. This section describes the details of the limitations of traditional routing mechanism, basically in cluster-based routing, where energy efficiency is a key due to the limited battery capacity of sensor nodes. It explains that the problems of current methods, especially in cluster-based routing, are necessary because sensors possess limited battery power and must conserve energy.

From the earlier research of (Abadi et al., 2022), (Rodoshi et al., 2021) & (Pantazis et al., 2013), the scientists used various routing protocols to reduce energy consumption and find the best routing in WSN contexts. The design of the proposed solution that is supposed to optimize energy efficiency for WSNs bridges the mentioned gap and resolves the research problems while optimizing network lifetime. Finally, we discussed the rationale of the algorithm in detail, comprising its model selection process. The primary purpose of this chapter is to demonstrate a complete this is description of a model that minimizes energy consumption and determines the optimal route to allow nodes to send data to the base station quickly.

4.2 Proposed Model Architecture

Deep Q-Network (DQN) is a reinforcement learning algorithm that uses deep neural networks to predict optimum Q-values for decision-making in complicated, high-dimensional environments, thereby being suitable for dynamic WSNs.

(Guo et al., 2023) suggested Deep Q-Network (DQN) reinforcement learning techniques that utilize deep neural networks to help estimate Q-values, which improves decision-making in complex and high-dimensional environments, making it appropriate for changeable WSNs. Traditional WSN Routing protocols typically focus on static energy-conserving methods that apply to some applications. While WSNs rely on dynamic features that need sufficient energy sources, our research, An adaptive EERP-RL WSN, utilized reinforcement learning for adaptive adjustment of routing decisions, optimizing the way we use energy, and reducing real-time communication latencies environment (Al-Karaki & Kamal, 2004). Adaptive-RL with a feature selection mechanism is proposed as a scheme to create a dynamic energy-conserving routing method, Reinforcement learning in wireless sensor networks (WSNs).

From our baseline work (Abadi et al., 2022), (Suresh et al., 2024a), and (Okine et al., 2024), the researcher used all types of features in the datasets. Using all types of dataset features that potentially impact the system's overall energy efficiency and reliability. WSNs are characterized by their dynamic nature. To overcome such problems, dataset preprocessing techniques are a vital step to be followed in the proposed routing protocol implementations.

The second gap from related work is additional energy usage due to frequent battery checks, potential communication delays in transmitting alerts, and the challenge of balancing energy efficiency with the accuracy of controlling, all of which could hinder the protocol's performance in achieving network longevity. Consequently, another problem lies in the inability to adapt to dynamic energy consumption patterns, which can lead to inefficient battery utilization and failure to prioritize nodes with critically low energy levels for optimized routing decisions. In the proposed model, we use an appropriate sensor-type-specific data aggregation reward for our RL model to reduce complexity. This dataset should capture metrics such as energy consumption, residual energy types to facilitate the optimization of routing decisions. This improves the accuracy of the adaptive energy-efficient routing protocol in energy consumption and training time. The attributes of the energy-efficient routing protocol are employed after the feature selection stage for model training. This study uses class misbalancing handling techniques, such as Synthetic Minority Oversampling Technique (SMOTE), to prohibit biased models from favoring majority classes. To address this, techniques such as data resampling, cost-sensitive learning, and synthetic data generation can be employed. Resampling involves either oversampling the minority class or undersampling the majority class to balance the dataset.

(A. Khan & Sampalli, 2012), proposed a static clustering-based routing protocol for Wireless Sensor Networks (WSNs) that achieves cluster head selection by dividing large sensor fields into rectangular clusters and grouping them into zones, improving energy efficiency, network stability, throughput, and lifetime, as demonstrated through MATLAB simulations, outperforming LEACH, MH-LEACH, and SEP protocols in large areas. Modification of the threshold value and incorporating bio-inspired algorithms across diverse applications are one of the optimal routing protocols in WSNs.

(Shah & Yashwanth, 2022) proposed modification of the threshold value and incorporating bio-inspired algorithms, considering real-time factors like remaining energy, cluster count, node functionality, and distance to improve network lifetime, scalability, and packet delivery ratio across diverse applications. Reinforcement learning ML is one of the adaptive, energy-efficient routing protocols in WSNs. In this paper, we proposed an adaptive energy-aware routing protocol based on reinforcement learning (AEERP-RL) for Wireless Sensor Networks (WSNs), which relies on optimizing energy consumption and network lifetime.

Our suggested Adaptive Energy-Efficient Routing Protocol Using Reinforcement Learning (AEERP-RL) for Wireless Sensor Networks (WSNs) is designed to optimize energy use and network lifetime using reinforcement learning for dynamically changing routing policies, and the dynamic nature of WSNs. Since WSNs are resIn this paper, we proposed an adaptive energy-aware routing protocol based on reinforcement learning (AEERP-RL) for Wireless Sensor Networks (WSNs) which relies on optimizing energy consumption and network lifetime. Our suggested Adaptive Energy-Efficient Routing Protocol Using Reinforcement Learning (AEERP-RL) for Wireless Sensor Networks (WSNs) is designed to optimize energy use and network lifetime using reinforcement learning for dynamically changing routing policies and the dynamic nature of WSNs. Since WSNs are resource-constrained, they depend on RL algorithms. (Guo et al., 2023) suggested sleep scheduling, and control data transmission depending on variations in the received.

The Adaptive Energy-Efficient Routing Protocol Based on Reinforcement-Learning-Based Energy Power-Efficient Protocol (AEERP-RL) is designed to utilize less power and reduce network lifetime by reinforcement learning in order to modify routing policies as necessary, and depends on making WSN routing more energy-efficient. Reinforcement learning is a WSN routing method that can solve these problems. Deep Q-Network (DQN) is a method of reinforcement learning that uses deep neural networks to approximate Q-values for the best decision in complicated, high-dimensional environments, and therefore is an appropriate choice for WSN modification. (Guo et al., 2023) Suggested Deep Q Network (DQN) reinforcement learning algorithms that utilize deep neural networks to approximate Q-values to make the best decisions in complex, high-dimensional environments, making it highly suitable for dynamic WSNs. Traditional WSN routing protocols generally tend to focus on static power-saving practices appropriate to particular applications. WSNs rely on varying characteristics that need to be supported by suitable power supplies. Our research, an adaptive EERP-RL WSN with reinforcement learning for dynamic routing decision making and reducing energy consumption and communication overhead in a real-time (Al-Karaki & Kamal, 2004) they depends on RL algorithms.

Adaptive-RL with a feature selection mechanism is the proposed design solution for effectively developing an adaptive energy-efficient routing protocol using reinforcement learning for wireless sensor networks (WSNs). From our baseline work (Abadi et al., 2022), (Suresh et al., 2024a), and (Okine et al., 2024), the researcher used all types of features in the datasets. Using all types of dataset features that potentially impact the system's overall energy efficiency and reliability. WSNs are characterized by their dynamic nature. To overcome such problems, dataset preprocessing techniques are a vital step to be followed in the proposed routing protocol implementations.

The second gap from related work is additional energy usage due to frequent battery checks, potential communication delays in transmitting alerts, and the challenge of balancing energy efficiency with the accuracy of controlling, all of which could hinder the protocol’s performance in achieving network longevity. Consequently, another problem lies in the inability to adapt to dynamic energy consumption patterns, which can lead to inefficient battery utilization and failure to prioritize nodes with critically low energy levels for optimized routing decisions. In the proposed model, we use an appropriate sensor-type-specific data aggregation reward for our RL model to reduce complexity. This dataset should capture metrics such as energy consumption, residual energy types to facilitate the optimization of routing decisions.

This improves the accuracy of the adaptive energy-efficient routing protocol in energy consumption and training time. The attributes of the energy-efficient routing protocol are employed after the feature selection stage for model training. This study uses class misbalancing handling techniques, such as Synthetic Minority Oversampling Technique (SMOTE), to prohibit biased models from favoring majority classes. To address this, techniques such as data resampling, cost-sensitive learning, and synthetic data generation can be employed. Resampling involves either oversampling the minority class or undersampling the majority class to balance the dataset.

(A. Khan & Sampalli, 2012) proposed a static clustering-based routing protocol for Wireless Sensor Networks (WSNs) that achieves cluster head selection by dividing large sensor fields into rectangular clusters and grouping them into zones, improving energy efficiency, network stability, throughput, and lifetime, as demonstrated through MATLAB simulations, outperforming LEACH, MH-LEACH, and SEP protocols in large areas. Modification of the threshold value and incorporating bio-inspired algorithms across diverse applications are one of the optimal routing protocols in WSNs

(Shah & Yashwanth, 2022), proposed modification of the threshold value and incorporating bio-inspired algorithms, considering real-time factors like remaining energy, cluster count, node functionality, and distance to improve network lifetime, scalability, and packet delivery ratio across diverse applications. Reinforcement learning ML is one of the adaptive, energy-efficient routing protocols in WSNs. In our study, we proposed an adaptive energy-efficient routing protocol using reinforcement learning (AEERP-RL) for Wireless Sensor Networks (WSNs) that relies on optimizing energy consumption and network lifetime. Our proposed Adaptive Energy-Efficient Routing Protocol Using Reinforcement Learning (AEERP-RL) for Wireless Sensor Networks (WSNs) aims to optimize energy consumption and network lifetime by utilizing reinforcement learning to dynamically adjust routing policies and the dynamic nature of WSNs. Since WSNs are resource-constrained, they basically depend on RL algorithms.

(Guo et al., 2023) proposed sleep scheduling, and control data transmission based on changes in the received. The Adaptive Energy-Efficient Routing Protocol Using Reinforcement-Learning-

Based Energy Efficient Protocol (AEERP-RL) aims to optimize energy consumption and network lifetime by utilizing reinforcement learning to dynamically adjust routing policies and relies on optimizing energy-efficient WSN routing. A reinforcement learning WSN routing algorithm can solve these issues.

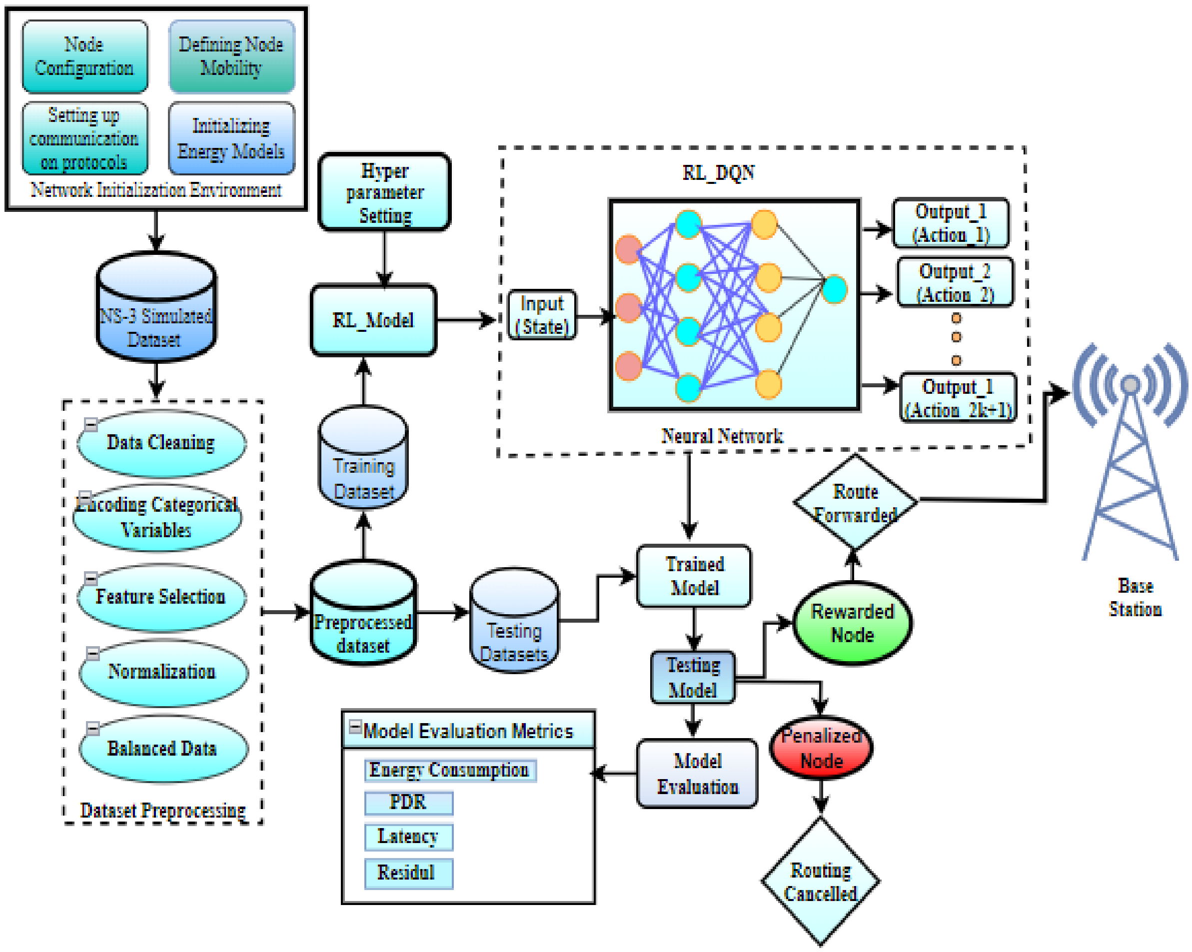

Figure 4.1: Proposed architecture

Figure 4.1 illustrates the end-to-end architecture details workflow for our proposed system, aligning closely with the stated objectives. The process begins with the simulation of a wireless Sensor Network using NS-3, where essential configurations like node behavior, communication protocols, mobility, and energy models are defined to generate a realistic and valuable dataset-specially capturing energy patterns. The core of the system is the RL-based Deep Q-Network (RL_DQN), which dynamically learns optimal routing decisions by interpreting state inputs and selecting the most energy-efficient actions from multiple outputs.

The trained model evaluates routing decisions by forwarding data packets via rewarded nodes or cancelling them in case of inefficiencies (penalized nodes). This adaptive mechanism ensures the protocol can respond intelligently to changing network conditions. Finally, performance is assessed using crucial evaluation metrics, energy consumption, packet delivery ratio (PDR), latency, and residual energy, fulfilling the objectives of protocol validation, simulation-based evaluation, and comparison with baseline routing schemes.

4.3 Proposed Model Algorithm Flows

To address algorithmic flow, a proposed DL model integrates its input and output metrics through operations, including sequential steps in the model's structure. Given that our objective is to create an adaptive energy-efficient routing protocol with Reinforcement Learning for WSNs, the model is trained on collated data rather than interacting with the system in real time. We employ Offline Reinforcement Learning with a prepared preprocessed dataset, from which features, including node energy usage, residual energy, throughput, and latency, among others, are extracted and normalized. Network performance metrics and node-related variables define the state space, whereas possible routing decisions define actions.

The reward functions aim at maximizing efficiency while incorporating energy expenditure and packet delivery ratio (PDR). For the algorithmic flow, we use a Deep Q Network (DQN) to update Q values without the need to have the model actively interact with the environment, enabling the model to determine effective routing strategies. The network is subsequently updated with the learned policy, which allows energy-efficient routing decisions to be made depending on the predicted state of the network while minimizing energy use and maximizing throughput. Below is a general common template for the selected algorithmic flow of the RL model.

Algorithm 4. 1 Algorithm 1 Proposed Method AEERP-RL

Initialize Parameters:

1. Q-network (policy network): θ (weights)
2. Target Q-network: θ_target ← θ
3. Replay buffer B (preloaded with preprocessed dataset D)
4. Discount factor γ, batch size N, target update interval C

```
Preprocess Dataset D
States (s): [node_energy_consumption, residual_energy, throughput, latency, ...] (normalized)
Actions (a): Set of possible routing decisions (e.g., next-hop node IDs)
Rewards (r): R(s,a) = α * PDR + β * (-energy_cost)  // Combine PDR and
    energy efficiency
    Transitions: (s, a, r, s', done) extracted from D
      Training Loop (for M epochs)
      for epoch = 1 to M do
    Sample batch of N transitions (s, a, r, s', done) from B
    Compute Q-values for current state-action pairs
        Q_current = Q(s, a; θ)  // Forward pass through Q-network
            Compute target Q-values
                for each transition in batch
                    if done
      target = r
Else
 Q_next = max(Q_target(s'; θ_target))  // Max Q-value of next state
  target = r + γ * Q_next
    Compute loss: L = MSE(Q_current, target)
      Update θ via gradient descent on L
        Every C steps: θ_target ← θ  // Soft update: θ_target = τ*θ + (1-τ)*θ_target
        Deploy Learned Policy
    Use trained Q-network to select actions (routing decisions)
```

28. For current network state s

29. a* = argmax(Q(s, a; θ))

30. Update WSN routing tables with a* to minimize energy use and maximize throughput.

Reinforcement learning (RL) based routing protocols play a vital role in the dynamic properties of WSNs. In our research, the reinforcement learning RL model begins with the definition of important elements to train a Deep Q-Network (DQN). These involve initializing the Q network (θ) for approximating action-value functions and a target Q-network (θ_target), which is renewed at regular intervals to maintain training stability. These frames: A replay buffer (B) is preloaded with a cleaned dataset (D) with normalized state features like node energy level, residual energy, and throughput. The action space represents discrete routing choices, i.e., selecting the next-hop node. The reward function incorporates packet delivery ratio (PDR) and energy efficiency, as $R(s,a) = \alpha.PDR + \beta.(-energy\_cost)$, to achieve high reliability and energy conservation. Training is performed on transitions in the form of (state, action, reward, next state, done) logs.

During each epoch, mini-batches of transitions are sampled from the buffer for estimating the current Q-values and target Q-values from the Bellman equation. WSNs are defined by their performance metrics. The network is trained by minimizing the mean squared error (MSE) between predicted and target Q-values via gradient descent, and the target network is updated periodically using soft updates for better learning consistency. For trained, the Q network enables real-time decision-making through choosing actions with maximum expected rewards, enabling on-the-fly modification of routing tables that reduce energy consumption and boost throughput. This RL-based routing strategy ultimately extends network lifespan and ensures more efficient data transmission in resource-limited wireless sensor networks.

4.4 Optimization of Hyperparameters

This process entails the identification of particular values of parameters that control the learning process of the model tries to perform the best on a specific task. Such configurations, which are external to the model and not deduced from the data, must be set before training. In reinforcement learning (RL), hyperparameter optimization is essential to optimize both performance and energy efficiency. In our RL approach, agents learn from pre-collected data (experience relay) rather than interacting with the environment in real-time. This is due to all features are not important in our study. Selecting the best hyperparameters is a critical step in training, as it significantly affects the model's ability to generalize to new, unseen data. The necessary parameters for the proposed models were defined before training as follows:

4.4.1 Regularization Techniques

Dataset cleaning is an important step in solving our model's complexities. Nevertheless, this problem is dealt with through regularization, which has a key role in managing the model's complexity and from being too specialized to the training data. The following delineates key regularization methods:

L2 Regularization (Weight Decay): It is utilized for fighting the issue of overfitting by penalizing large weights in the neural network. This is particularly needed in Deep Q in hierarchical RL and DQN models; the policy that is learned must be generalizable to various WSN environments in order to optimize the system performance.

Target Network Regularization: It enhances learning by acting from a separate target network that updates across time. This prevents overestimation of Q-values and enhances convergence in offline reinforcement learning, and enhances the reliability of routing decisions under changing network conditions.

4.4.2 Learning Rate

Optimizing the learning rate is crucial to achieving a trade-off between stability and convergence speed. A high learning rate will cause the model to diverge from Q-value updates, and a minimal learning rate leads to slow convergence, where the best routing policies change.

4.4.3 Discount Factor

Maximizing the discount factor (γ) is a fundamental approach to trading off short-term energy efficiency and prolonged network performance in RL-based routing, using Deep Q Network. Maximum γ (close to 1) favors long-term energy conservation, promoting network lifetime, but can lead to suboptimal short-term routing decisions. A minimum γ (close to 0) option immediately rewards, which can optimize short-term packet delivery ratio (PDR) and throughput, yet it can cause rapid energy depletion.

4.4.4 Exploration Rate

In wireless sensor networks (WSNs), increasing the exploration rate (ε in ε-greedy policy) plays a crucial role in trading off exploration and exploitation. Since offline RL depends on existing datasets without online interactions, poor exploration will lead to inferior policy learning or overfitting to historical actions. A decaying exploration rate strategy, such as exponential decay or adaptive ε-tuning, can reduce this by essentially enabling varied action selection and progressively focusing on exploitation as training proceeds.

4.4.5 Batch size

It plays a key role in stability and efficiency training. The batch consists of heterogeneous historical experience samples gathered from WSN simulations, spanning various network Parameters, including node density, energy levels, traffic patterns, and environmental conditions.

4.4.6 Reward Clipping

This technique is important in helping us guarantee the stability of our training procedure and avoiding and surmounting prolonged reward results stemming from managing learning advancement. It ought to articulate how the values are clipped or normalized into a particular range (i.e., [-1, 1] or [0, 1] to produce stable learning, especially in Hierarchical DQN models where abrupt changes in energy efficiency measures can create instability.

CHAPTER FIVE

IMPLEMENTATION OF MODEL ARCHITECTURE

5.1 Introduction

In this chapter, we reviewed the suggested solution aimed at fixing the structural issues of our model, which aids in energy-efficient routing. This section discusses the design and incorporation of a reinforcement learning framework in the WSN context, utilizing the NS-3 simulator to enhance routing decisions upon simulated parameters, thereby encouraging energy efficiency. Deep learning-based functionality integration is emphasized, and approximations, i.e., Deep Q-Network (DQN) or Hierarchical DQN, are used to achieve successful learning of energy-efficient routing methods from experience.

5.2 Environment Setup

The experimental framework utilized in our investigation was established through the NS-3.36.1 simulation tool on Ubuntu 22.04 utilized a cluster-based topology to simulate the wireless sensor network (WSN), with various nodes grouped into clusters based on natural deployment situations. In every cluster, member nodes communicated through a specially designated cluster head, which handles data forwarding and aggregation. The AODV routing protocol was initially used to replicate traditional routing techniques and act as a benchmark.

The simulation tracked important performance metrics, including energy use and residual energy, packet delivery ratio (PDR), queue length, latency, throughput, and link quality information, which were essential for reinforcement learning algorithm creation. Also, the environment includes node mobility and heterogeneous channel conditions to create a heterogeneous and adaptive platform with the capacity to support adaptive learning and optimize routing approaches efficiently.

5.3 State Space

It is a complete representation of the evolving state of the wireless sensor network, which is the input to the reinforcement learning algorithm. Each state has node and network parameters that influence routing decisions. These parameters are the current energy level of the node, residual energy, channel quality, and the delay experienced in recent communications. In addition, the state vector involves key performance measures such as throughput and packet delivery ratio (PDR), and latency, which reflect the reliability and efficiency of the communication link. Spatial data, such as the distance of the node from the base station and the condition of the communication channel, is also embedded, enabling the model to evaluate both proximity and feasibility of transmission. To optimize decision-making in hierarchical topologies, neighboring node states, node mobility, and cluster topology features are addressed.

5.4 Action Space

It defines the set of operations an intelligent node can execute during every time slot in order to maximize overall network efficiency. It is meant to optimize effective and energy-efficient routing approaches that are apt to the dynamic nature of Wireless Sensor Networks (WSNs). The RL agents i.e. WSN nodes has four core options: (1) forward a packet to a designed next-hop node, (2) discard the packet when conditions are unfavorable, (3) enter an idle or sleep mode to preserve energy, and (4) engage in cluster-head re-election procedures within clustered network configurations. These options empower the reinforcement learning agent to arrange routing situations in real time based on current network parameters such as saved energy, queue status, link stability, and traffic load.

By incorporating these actions with feedback from the environment via a well-defined reward mechanism, the model evolves to identify policies that promotes energy usage, ensure packet delivery, the model evolves to identify policies that maximize energy usage, ensure packet delivery, and extend network longevity. This adaptable and context-sensitive action space supports intelligent decision-making in response to ongoing network changes.

5.5 Proposed Model Implementation

5.5.1 Dataset Preprocessing

It is quite essential to perform preprocessing on the dataset to prepare the simulation outputs for the model training by arranging them in a way suitable for an energy-efficient guarantee that the data presented to the DQN is appropriate, consistent, and pertinent for insightful learning. With the aid of Python libraries, the preprocessing steps of normalization, transforming data types, as well as label encoding are done. The MinMaxScaler from Scikit-learn was applied to the regression curves or independent variable values to deal with massive variance in feature scales. This helps to avoid a scenario where a certain feature takes significant prominence over the learning process while leaving the target variable unscaled to keep its distribution intact.

A problem faced when working with the raw dataset is the presence of different data formats, which require all values to be in float or int. The need to convert to avoid compatibility issues during model training is the reason for this. Since machine learning works on numerical data, some columns had to be converted to enable the astype () function to convert other columns to the required form. Additionally, reward indicators are Categorical features that had been encoded to become numerical using the LabelEncoder of Scikit-learn; positive rewards were also labeled as (0, 1) while penalties became (-1, 0). These changes improve model training by enabling it to better interpret input data, thus improving accuracy and robustness.

Figure 5.1 describes a situation where the choices for the next hop in our wireless sensor network aren't exactly balanced. Each bar represents a different option for where a data packet can go next, and they all seem to be hovering around the same height, somewhere between 450 and 550 counts. This initial snapshot, "before SMOTE" as the title says, suggests that when we're training our reinforcement learning agent to figure out the best way to route data for energy efficiency, it's seeing roughly the same number of instances for each potential next step. This relatively even distribution across our "Next_hop_selection" options is a good starting point, meaning our agent isn't being overwhelmed by one particular routing choice in the initial data.

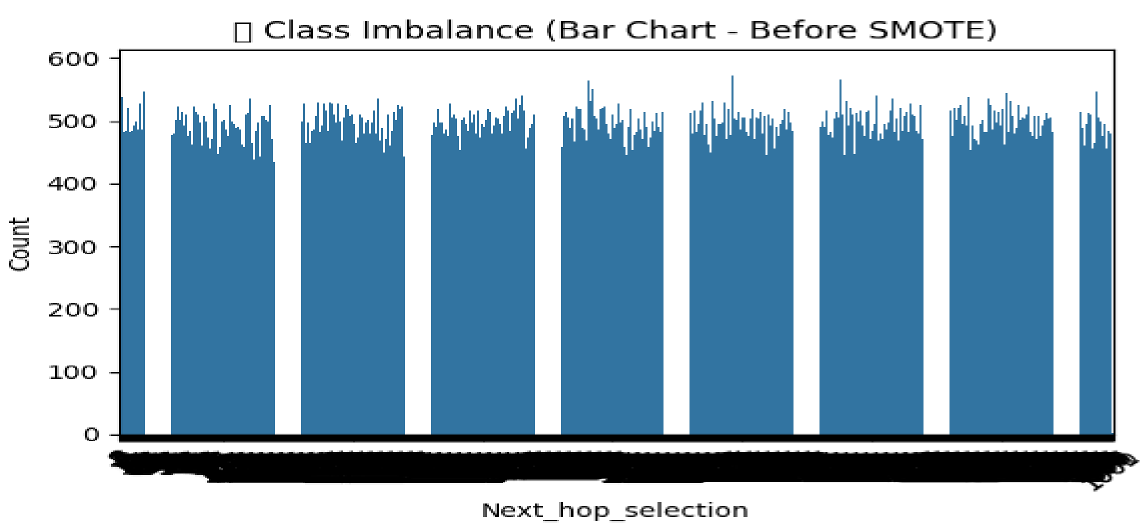

Figure 5.1:Class Imbalance (Bar Chart – Before SMOTE)

Figure 5.2, representing our different "Next hop selection" choices, is almost the same height, sitting right around the 580 mark. It looks like they've evened things out. This balancing act is likely to help our reinforcement learning model train more effectively. By ensuring each potential next hop selection is represented with a similar number, the model should be less biased towards any single routing option and hopefully learn a more nuanced and ultimately more energy-efficient way to send data across our wireless sensor network. The next hop selection needs an optimal path for energy-efficient routing.

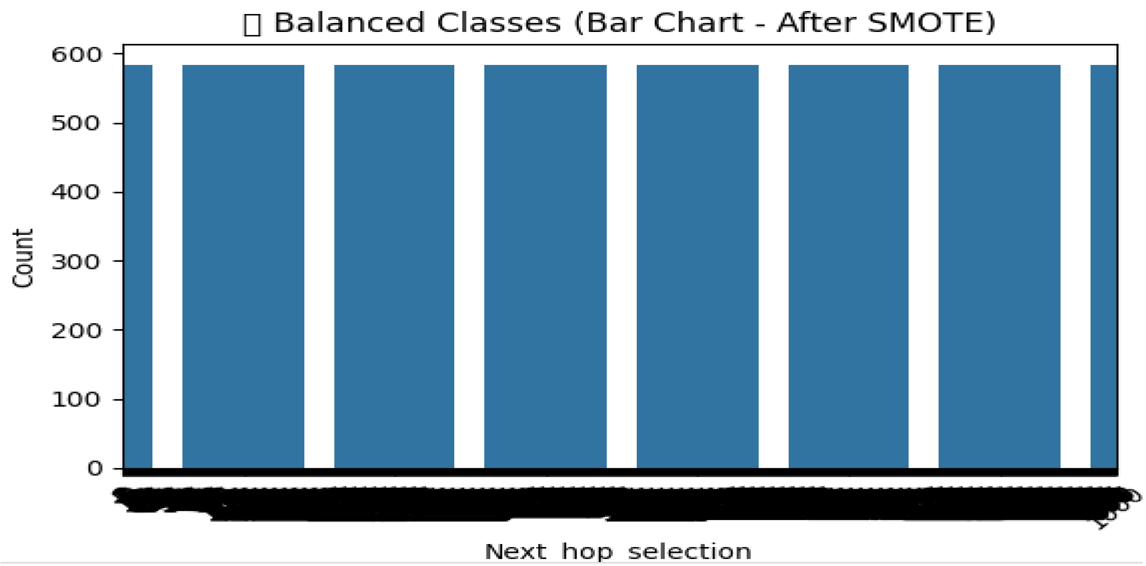

Figure 5.2: Class Imbalance (Bar Chart – Before SMOTE)

In Figure 5.3, Class Distribution of the original breakdown of our reward classes, and it's a best tight split – almost a perfect half-and-half between positive and negative rewards, with just a minor more on the positive side (50.2% green versus 49.8% red). It's nearly 50% for both positive and negative rewards. This kind of balancing is likely a deliberate step to make sure our reinforcement learning model gets a fair share of both reward and penalty feedback during training. It prevents the model from getting too used to one type of outcome and hopefully leads to a more robust and well-rounded routing strategy for our energy-efficient RL routing protocol wireless sensor network.

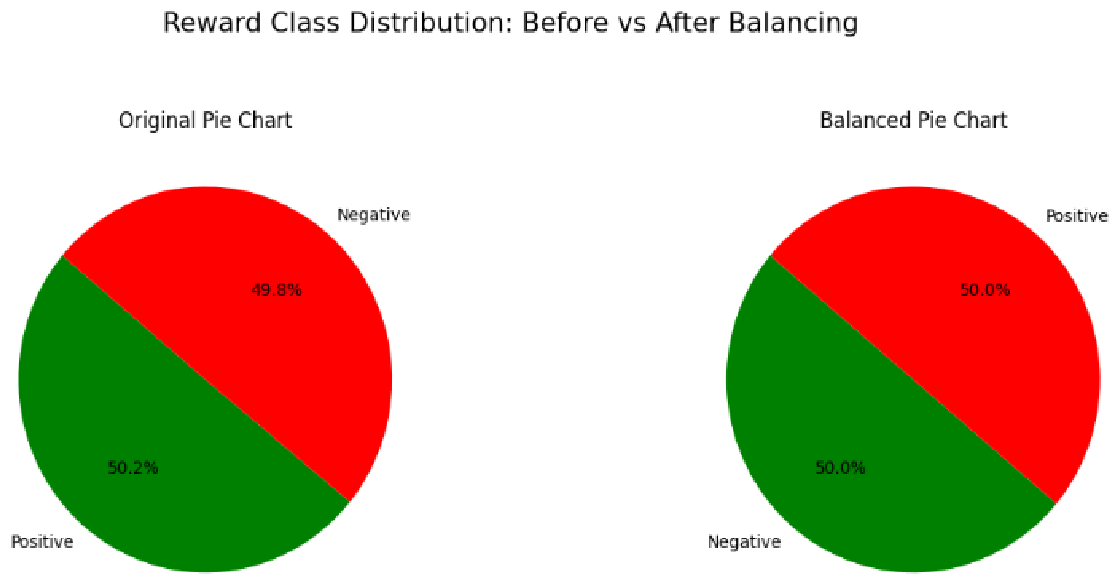

Figure 5.3: Class Distribution: Before Vs After Balancing

5.5.2 Proposed Model Implementation

In this section, we detailed an implementation that is a vital scenario in any research topic. It covers and implementation of the selected model for selecting the optimal route path in WSNs routing protocols, as well as importing essential libraries, and optimization parameters we used for model training. Before implementing the DQN RL, we imported the necessary libraries.

```python
# === IMPORTS ===
import pandas as pd
import numpy as np
import torch
import torch.nn as nn
from torch.utils.data import DataLoader, Dataset, random_split
import matplotlib.pyplot as plt
import random
from sklearn.metrics import mean_squared_error, r2_score

# === CONFIGURATION ===
DATASET = '/kaggle/input/aeerprl-normalized-ds/Aeerprl_normalized_dataset.csv'
BATCH_SIZE = 64
EPOCHS = 100
LR = 1e-5

INPUT_FEATURES = [
    'Node_energy_levels', 'Residual_energy', 'Link_quality', 'Channel_condition',
    'Node_distance', 'Queue_length', 'Hop_count_to_sink', 'Neighboring_node_states',
    'Next_hop_selection', 'Reward', 'Successful_packet_delivery',
    'PDR', 'Throughput'
]
TARGET = 'Energy_consumption'

# === SEED SETTING FOR REPRODUCIBILITY ===
torch.manual_seed(42)
np.random.seed(42)
random.seed(42)

# === CUSTOM DATASET CLASS ===
```

Figure 5.4 Importing necessary libraries

In our study, we have used the Model hyperparameters of Table 5.1 and the training hyperparameters of Table 5.2 below.

Table 5.1 Model Hyperparameters

Model Hyperparameters		
Hyper parameter	Value	Description
Input_size	13	Number of input features
Hidden_layer_1	64	Number of neurons in the first hidden layer.
Hidden_layer_2	128	Number of neurons in the first hidden layer.
Activation_function	ReLU	Activation used between layers.
Output_Size	1	Predicting a single target

Table 5.2: Training Hyperparameters

Training Hyperparameters		
Hyper parameter	Value	Description
BATCH_SIZE	64	Mini-batch size for training
EPOCHS	100	Total number of training iterations over the full dataset.
LR (Learning Rate)	1e-5	Learning rate for the Adam optimizer
Loss_function MSELoss	MSELoss	Use to train the network
optimizer	Adam	Optimization algorithm used for updating weights
Shuffel	True (train)	Randomizes training data at each epoch.

CHAPTER SIX

RESULTS AND DISCUSSIONS

In this section, we examine a comprehensive evaluation of the provided Adaptive Energy Efficient Routing Protocol for WSNs, which leverages Deep Q-Network (DQN) reinforcement learning to optimize routing decisions, optimize energy utilization, and prolong overall network lifetime. The experimental output assesses the effectiveness of the proposed DQN-based model in comparison to existing routing approaches, focusing on key performance metrics. From our analysis, a detailed description of feature importance, highlighting the relative contribution of various network parameters to the target performance outcomes, is detailed to evaluate the performance of our model.

6.1 Feature Importance Analysis

To improve the clarity of the proposed adaptive energy-efficient routing protocol and assess the influence of individual input parameters, a feature importance analysis was performed using the Random Forest Regressor. Our analysis details the key relevant features that critically impact the model's prediction of energy utilization, and an essential performance indicator in Wireless Sensor Networks (WSNs). Feature selection is a basic aspect of machine learning for efficiency optimization, as not all features contribute equally to model performance. Identifying and selecting these features that are most relevant for our study is crucial for enhancing the model's accuracy and efficiency, particularly in energy-efficient routing protocols. In this context, the most impactful features from a total of 19, as outlined in section 3.3 and detailed in the RandomForestRegressor were used in this study to compute the importance scores of each feature.

Table 6.1: Selected feature list based on RandomForestRegressor scores

Rank	Feature	Score Importance
1	Network_lifetime	0.101576
2	Throughput	0.101505

3	Node_distance	0.100924
4	Latency	0.100552
5	Energy_consumption	0.085007
6	Residual_energy	0.084544
7	Node_energy_levels	0.081756
8	Link_quality	0.076854
9	Queue_length	0.071088
10	PDR	0.069642
11	Hop_count_to_sink	0.057841
12	Neighboring_node_states	0.044718
13	Action	0.010590
14	Channel_condition	0.009321
15	Successful_packet_delivery	0.004081

Table 6.1 describes the feature importance rankings, which were determined using the RandomForestRegressor experiments. As a result, these attributes are deemed more optimal than others for training our models and achieving strong performance.

6..1 Correlation Analysis Using Barplot Visualization

Barplot-based correlation analysis offers a straightforward and visual way to understand how input attributes relate to the variable. By using feature importance scores from a trained Random Forest Regressor, the barplot arranges each feature based on its importance on the model's performance. This visual tool makes it easier for researchers to interpret the data and identify the most crucial parameters for optimizing routing protocols. In essence, barplot visualization plays a vital role in confirming feature relevance and supporting the development of smart, energy-efficient routing strategies in Wireless Sensor Networks.

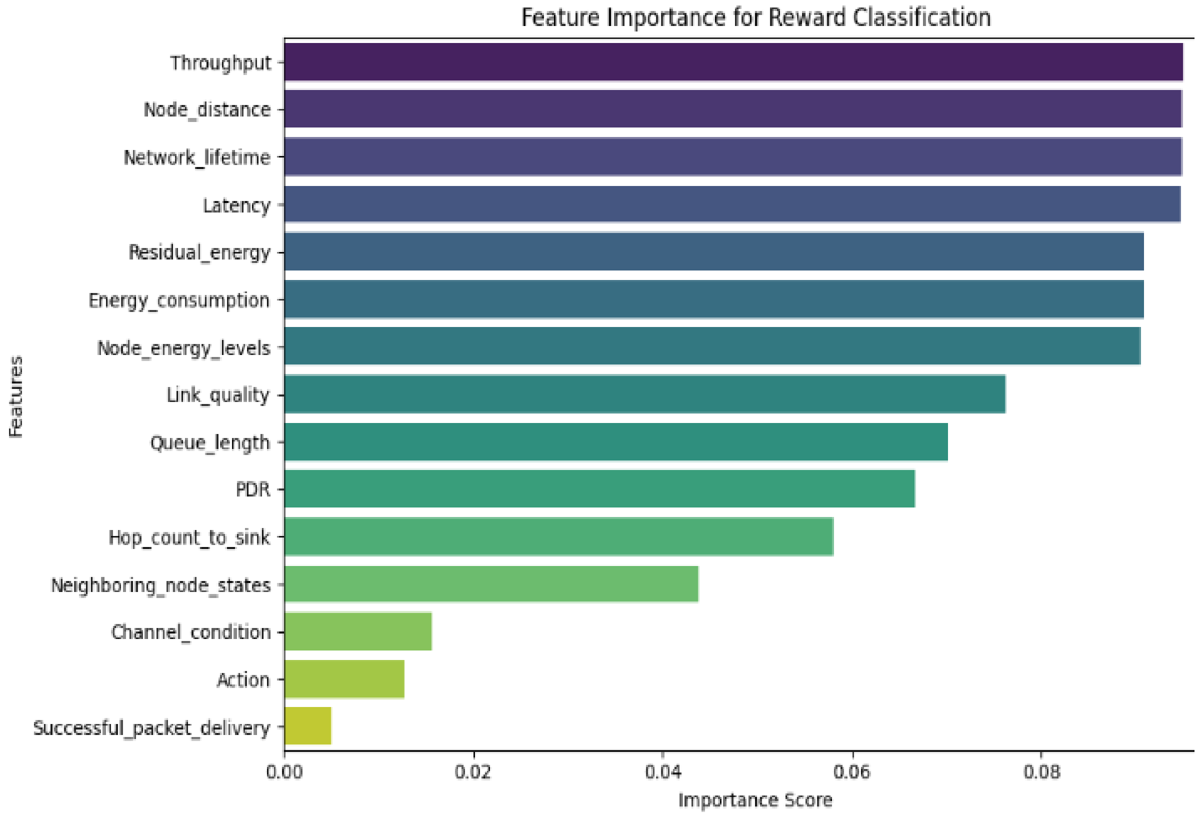

Figure 6.1: Feature Importance for Reward Classification

Figure 6.1 displays a feature relevance analysis for reward classification in the context of our study. The horizontal chart ranks various attributes based on their importance score. Invaluably, “throughput”, “Node_distance”, “Network_Lifetime”, and “Latency” imply the optimal important scores, suggesting these factors are most important in determining the reward signal for the Reinforcement Learning agent. Conversely, “Successful_delivery” and “Action” show the least importance. This relevance provides vital insights into which network parameters and node characteristics most significantly affect the defined reward function, directing the design of the RL agent’s state space and action policy to maximize these important features for enhancing energy efficiency, a key objective highlighted in the RLBEEP paper.

6.3 Outcomes of the Proposed Reinforcement Learning Models

Our study utilizes Deep Q-Learning, a machine learning-based technique, which in this work to optimize routing choices in WSNs. The main goals are to enhance packet delivery performance in changing network conditions, increase energy efficiency, and extend network lifetime. Here, we have to analyze the output of applying the DQN-based reinforcement learning model and emphasize its advantages over traditional routing mechanisms. It is expected that the model will provide the most significant benefits by utilizing DQN, an adaptive and energy-efficient routing algorithm for WSNs.

6.3.1 Experiment Setup

Table 6.2: Performance Comparison Table

SNo.	Protocol	Avg_Reward	Avg_Throughput	Avg_PDR
1	RLBEEP	0.000023	0.000006	0.000016
2	AEERPRL	0.499455	0.501446	0.509905

The AEERPRL protocol performs invaluable better than RLBEEP in all assessed parameters, including Average Reward, throughput, Packet Delivery Ratio (PDR), and Alive Nodes, according to the experiment, performance metrics, and visual comparison. According to the bar graph, AEERPRL demonstrated a significantly greater number of live nodes, which suggests optimizing energy efficiency and a longer network lifetime. This was described by the numerical data: With an Average Reward of 0.499455 as opposed to RLBEEP's nearly insignificant 0.000023, AEERPRL outperforms RLBEEP in reinforcement learning when it comes to striking a balance between energy consumption and data delivery. compared to RLBEEP, which only outperforms these instances to hit 0.000016, AEERPRL records a significantly greater Average Throughput (0.501446) and Average PDR (0.509905). These measurements demonstrate that AEERPRL not only more effectively conserves network resources but also guarantees.

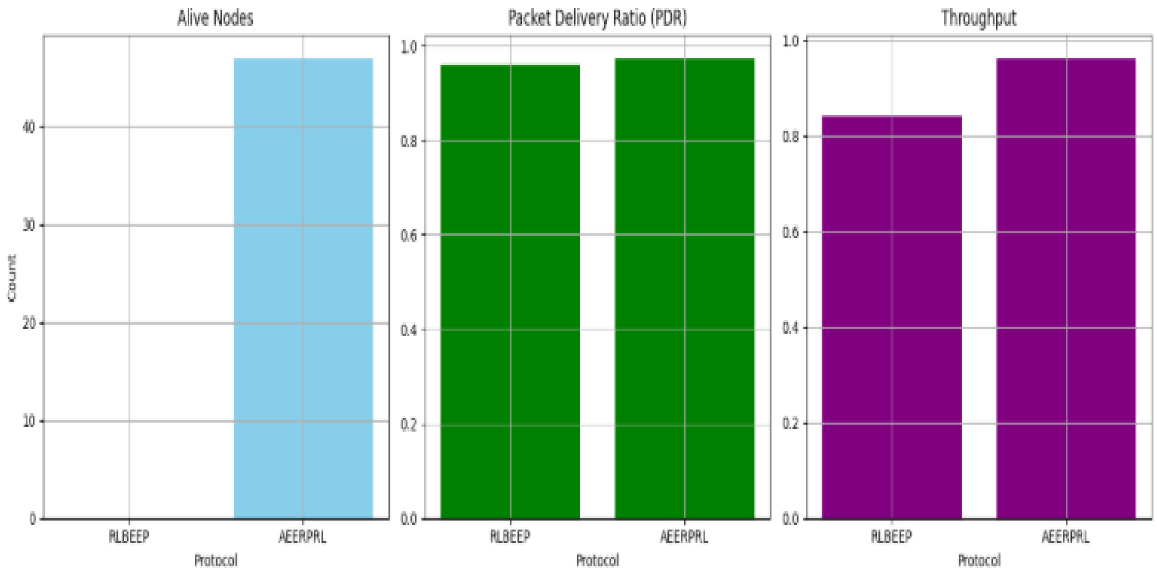

Figure 6.2: Protocol evaluation

Figure 6.2 performs a comparison between RLBEEP and AEERPRL protocols across the three most important key metrics relevant to Wireless Sensor Networks: Alive Nodes, Packet Delivery Ratio (PDR), and Throughput. The “Alive Nodes” plot implies that after a certain simulation time, AEERPRL maintains a significantly larger number of operational nodes compared to RLBEEP, implying higher energy efficiency in prolonging network lifetime. The "Packet Delivery Ratio" graph shows that both protocols provide a very high PDR, which suggests dependable data transfer. Lastly, the "Throughput" graph illustrates that AEERPRL also outperforms RLBEEP in the amount of data successfully transmitted per unit of time.

These results indicate that although both protocols successfully achieve reliable data transmission, AEERPRL shows better performance in terms of network lifetime and data transmission in the analyzed scenario, describing the future gains associated with adaptive energy management methods used by AEERPRL, which have been detailed in the first section of the the RLBEEP research illustrates the promising advantages associated with adaptive energy management approaches which are employed by AEERPRL, as outlined in the first section of the RLBEEP.

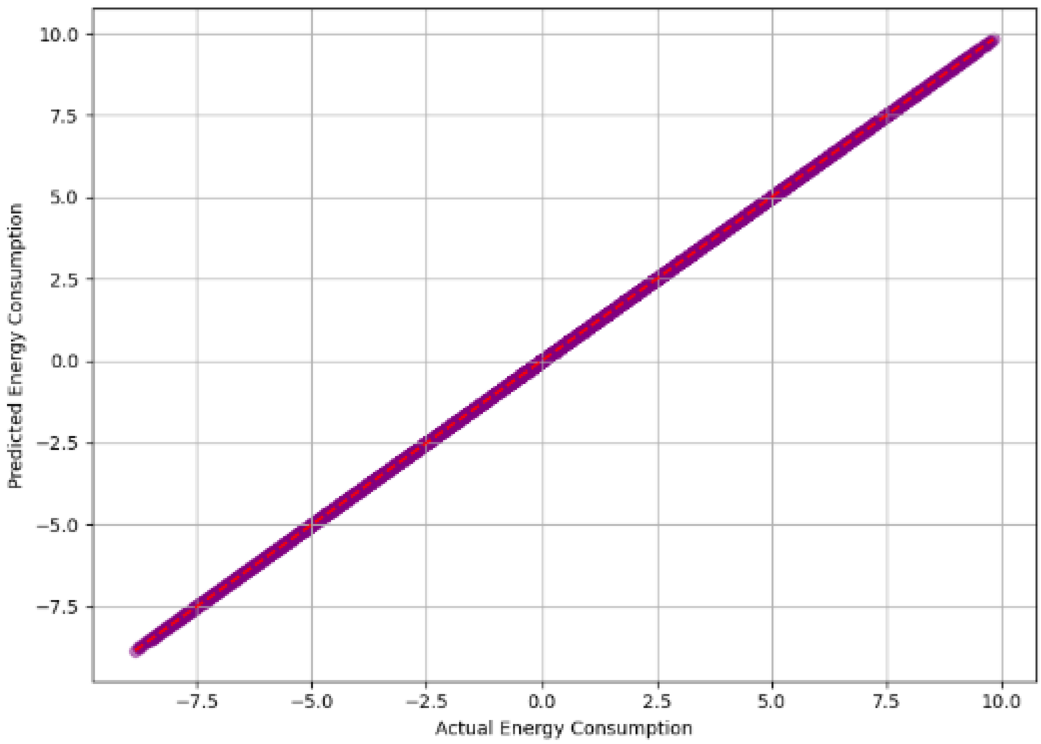

Figure 6.3: DQN Model: Actual vs Predicted Energy Consumption

6.3.2 DQN Training Performance Analysis

Figure 6.4 illustrates the training performance of a Deep Q-Network (DQN) model, which is a key component within our study with Reinforcement Learning. The plot shows the training loss decreasing steadily over a large number of epochs (around 100). The rapid initial reduction in loss shows that the DQN is quickly learning to cooperate with optimal Q-values, which implies the expected future reward for performing specific actions in those states. The sequence of gradual flattening of the loss curve shows that the model is converging and stabilizing its learning. A low and stable training loss is a positive indicator, implying that the DQN has effectively learned a policy that can lead the routing decisions for optimizing the defined reward, which in our context would likely collaborate with factors like optimizing energy consumption, optimizing packet delivery. This successful convergence of our model is a vital step in enabling the development of our routing protocols.

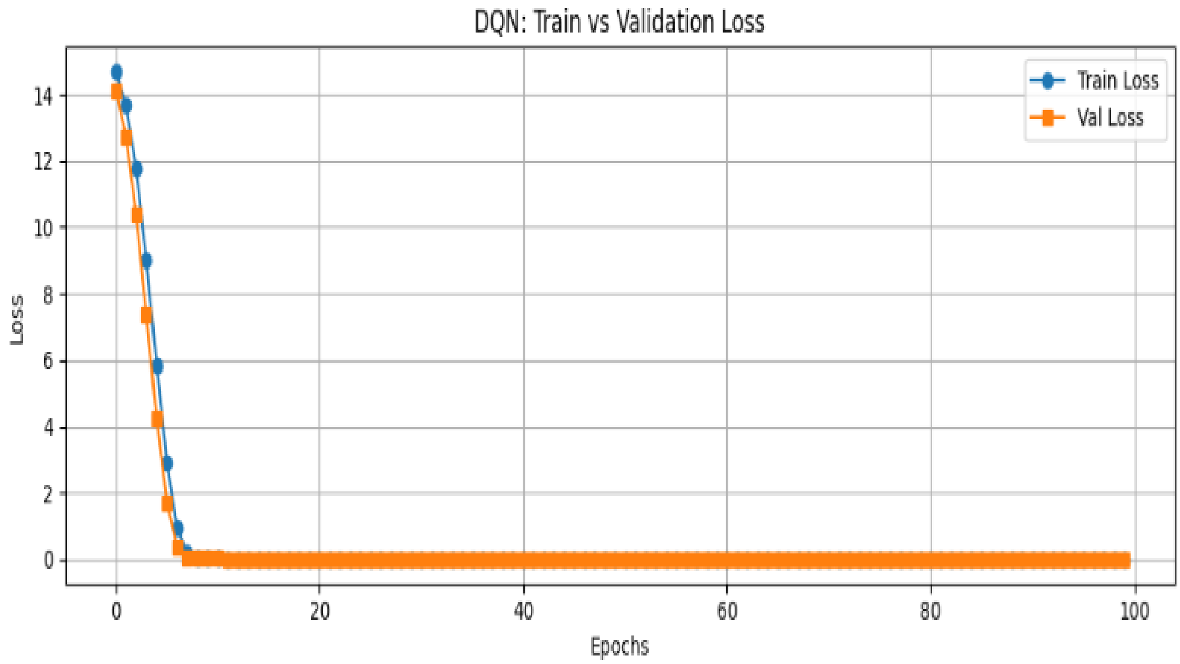

Figure 6.4: DQN-Train vs Validation Loss

From the graph, the scatter plot visualizes the relationship between the actual energy consumption and the predicted energy consumption by our developed model. Each data instance on the graph details a specific instance, with the x-coordinate indicating the actual energy consumption level and its y-coordinate representing the corresponding predicted energy consumption level of the routing protocol. The tight clustering of the data points along a diagonal line, which would ideally have a slope of 1 and pass through the origin, strongly suggests a high degree of correlation between the actual energy consumption and predicted values.

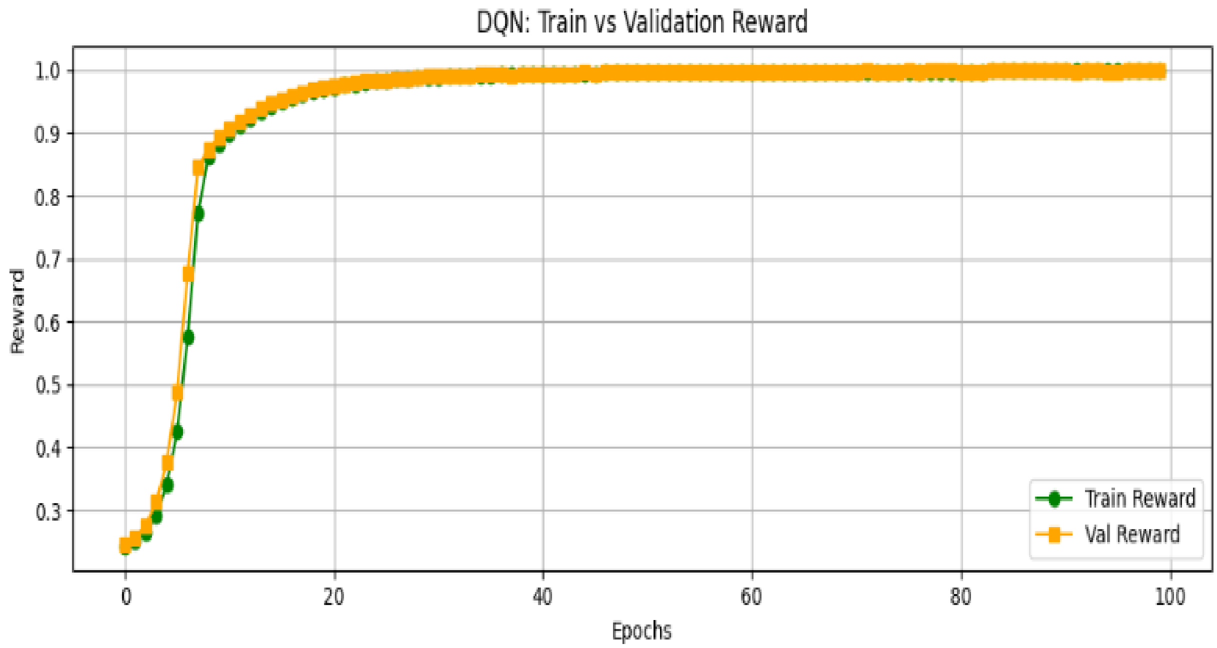

Figure 6.5: DQN Train vs Validation Reward

The figure plots the learning process of a Deep Q-Network (DQN) agent in terms of how well it can develop an adaptive energy-efficient routing protocol for Wireless Sensor Networks (WSNs). The x-axis is the epochs of training, i.e., the number of times that the agent has passed over the training data, and the y-axis is the reward, which is a numerical quantity that indicates the proficiency of the agent at making energy-efficient routing choices. There are two curves plotted separately: the "Train Reward," plotted in green with circular markers, is the agent's performance on the training set, and the "Val Reward," plotted in orange with square markers, is the agent's performance on a separate validation set held out to evaluate the agent's ability to generalize to new situations. Both rewards start off increasing rapidly, indicating the agent is learning improvements to its routing policy quickly. The payoff from training typically achieves a higher degree, a standard phenomenon within the machine learning field, since the agent is optimized for this particular dataset.

But the validation reward does experience a strong peak, indicating the agent is learning a generalizable policy. That the validation reward levels off just under the training reward indicates that, although the agent is learning well, there may be some degree of overfitting to the training set. Ideally, both reward curves would meet at the same high value, indicating strong and generalizable learning. The curve shows that the DQN is successfully learning to optimize energy efficiency over time.

6.4 Discussion of Research Questions

This study has 3 research questions in the first chapter of section 1.4. Hence, our study has answered these questions according to their order as follows.

RQ1: How can reinforcement learning adopt dynamic routing paths?

We applied appropriate data preprocessing, feature selection, and suitable parameter settings to develop the model architecture for routing networks. Once datasets are generated using NS-3 based on the given number of nodes, traffic rates, and dataset size in a controlled environment, data preprocessing is conducted for the model to fit with the selected feature flows. After this is achieved, the DQN RL-based is trained on the training set through the combination of appropriate parameter settings to identify parameters that deviate vitally from normal behavior. Accordingly, dropout = 0.002 and 0.001, learning rate = 0.001 and 0.0001, activation function ReLU in the input layer and Sigmoid in the output layer, and callbacks of 64 batch size over 300 episodes are used to achieve a high performance of greater than 0.000023 average reward.

After that we evaluated the model performance on the test set using metrics like average reward, number of alive nodes after routing, packet delivery ratio (PDR), and throughput to gauge its efficacy in developing an adaptive energy-efficient routing protocol using reinforcement learning for wireless sensor networks, followed by comprehensive testing to confirm its ability to generalize and mange an adaptive energy-efficient routing protocol.

RQ2: What are the key limitations of existing energy-efficient routing protocols in WSNs?

The current energy-efficient routing protocols in WSNs have many major limitations that point to their usability in volatile, resource-constrained environments, namely: poor adaptivity to dynamic network conditions, a high communication overhead, difficulty in achieving scalability, and node heterogeneity. Since the energy conservation objective is important, the largely critical parameters such as throughput, average reward, and packet delivery ratio (PDR), which are necessary for real-time controlling or emergency applications, have been ignored.

However, deterministic routing strategies lack learning capability. Unlike reinforcement learning (RL) approaches, which systematically learn the optimal routes based on environmental information from the network, conventional protocols have no such way to learn automatically concerning network changes or find additional energy-efficient path connections under potentially unpredictable circumstances. Both of these shortcomings together provide a rationale for adapting to improve energy efficiency without losing sight of energy constraints in WSNs.

RQ3: How can WSNs be modeled using reinforcement learning?

To achieve the objectives of our study, we used a Markov Decision Process (MDP) to model an energy-efficient routing protocol. The state space encompasses the changing conditions of the network, including individual node energy levels, queue lengths, channel conditions, and potentially the situations of neighboring nodes. The actions available to an RL agent, depending on each node (or a central controller), are the choices of the next hop for data packets.

The reward function is carefully designed to incentivize energy-efficient routing, basically by allocating positive rewards for successful packet delivery with low energy usage and penalties for energy depletion or packet loss. Through interconnection with the WSN environment, the RL agent learns an optimal policy that maps states to actions (next hops) to maximize the cumulative reward over time, thereby achieving adaptive and energy-efficient routing. Through interactions with the WSN environment, the RL agent learns an optimal policy that maps states to actions (next hops) to maximize the cumulative reward over time, thereby achieving adaptive and energy-efficient routing.

CHAPTER SEVEN

CONCLUSION and FUTURE WORKS

7.1 Conclusion

The problem associated with WSN energy efficient routing protocol is caused by an optimal route selection of sensor nodes to transmit data and communicate with the neighbor nodes, as well as the sink nodes, to achieve energy efficient routing. This study discusses the dynamic nature of wireless sensor networks to address the limitations of energy constraints in an adaptive energy-efficient routing. Heuristics based on traditional WSNs routing protocols often face challenges in maintaining energy efficiency in such dynamic environments because of their resource limitations. Our study proposes an adaptive routing protocol that employs Deep Q Network reinforcement learning to make intelligent and contextually aware decisions in various network settings and dynamic network environments.

7.2 Contribution of the Study

Our study's fundamental contribution is developing an adaptive, energy-efficient routing protocol for Wireless Sensor Networks (WSNs) that utilizes Deep Q-Network (DQN), a reinforcement learning technique. The proposed protocol minimizes energy consumption across the network by dynamically modifying routing options based on simulated network metrics such as residual energy, channel condition, next hop selection, and throughput.

Extensive simulation can demonstrate that our adaptive energy-efficient routing protocol using reinforcement learning for WSNs outperforms conventional systems like RLBEEP and AODV, achieving valuable contributions in packet delivery ratio (PDR), energy efficiency, and overall network performance. To facilitate ongoing research and development in RL-based WSN routing protocols, the paper also presents a replicable framework for creating NS-3 network-simulated datasets and an implementation methodology.

7.3 Future Work and Recommendations

While Energy-Efficient routing is a basic challenge in WSNs, we come up with DQN-RL feature selection mechanism for better energy-efficient routing. Future work on adaptive energy-efficient

RL framework-based routing protocols in WSNs under controlled settings may emphasize hybrid RL models that couple centralized monitoring with decentralized learning to improve scalability. Energy-conscious reward frameworks, which facilitate convergence and adaptation, should emphasize the incorporation of dynamic considerations like mobility, energy harvesting opportunities. Apply reinforcement learning-based adaptive routing protocols like AEERPRL to best adapt energy usage and prolong network lifetime in energy-constrained wireless sensor networks. Assign top priority to the use of the performance measures of Packet Delivery Ratio, Throughput, and Alive Nodes in the evaluation of routing protocols to ensure efficient and accurate data delivery. Promote ongoing research in scalable machine learning models for WSNs to address shifting environmental conditions and application needs in applications like precision agriculture, environmental monitoring, and IoT systems.

SPECIAL ACKNOWLEDGMENT

This research was funded by Adama Science and Technology University with grant number: **ASTU/ SM-R/1115/25**

Adama, Ethiopia

REFERENCES

A. Khan, Z., & Sampalli, S. (2012). AZR-LEACH: An Energy Efficient Routing Protocol for Wireless Sensor Networks. *International Journal of Communications, Network and System Sciences*, *05*(11), 785–795. https://doi.org/10.4236/ijcns.2012.511082

Abadi, A. F. E., Asghari, S. A., Marvasti, M. B., Abaei, G., Nabavi, M., & Savaria, Y. (2022). RLBEEP: Reinforcement-Learning-Based Energy Efficient Control and Routing Protocol for Wireless Sensor Networks. *IEEE Access*, *10*, 44123–44135. https://doi.org/10.1109/ACCESS.2022.3167058

Ahmad, R., Alhasan, W., Wazirali, R., & Aleisa, N. (2024). Optimization Algorithms for Wireless Sensor Networks Node Localization: An Overview. *IEEE Access*, *12*(February), 50459–50488. https://doi.org/10.1109/ACCESS.2024.3385487

Akinola, O. I. (2024). Adaptive Location-based Routing Protocols for Dynamic Wireless Sensor Networks in Urban Cyber-physical Systems. *Journal of Engineering Research and Reports*, *26*(7), 424–443. https://doi.org/10.9734/jerr/2024/v26i71220

Al-Karaki, J. N., & Kamal, A. E. (2004). A taxonomy of routing techniques in wireless sensor networks. *Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems*, 6-1-6–24. https://doi.org/10.1201/9780203489635

Al-Okby, M. F. R., Junginger, S., Roddelkopf, T., Huang, J., & Thurow, K. (2024). Ambient Monitoring Portable Sensor Node for Robot-Based Applications. *Sensors*, *24*(4). https://doi.org/10.3390/s24041295

Al Aghbari, Z., Khedr, A. M., Osamy, W., Arif, I., & Agrawal, D. P. (2020). Routing in Wireless Sensor Networks Using Optimization Techniques: A Survey. *Wireless Personal Communications*, *111*(4), 2407–2434. https://doi.org/10.1007/s11277-019-06993-9

Alate, M., Godase, U. R., & Kumbhar, U. T. (2024). *Secure and scalable data aggregation techniques for healthcare monitoring in WSN Karad Maharashtra Pune Maharashtra Karad Maharashtra India Sheela Hundekari * MIT Art , Design & Technology University Loni Kalbhor Maharashtra Araddhana Arvind Deshmukh @ Pu*. *27*(2), 441–452.

Alem, I., Yessad, S., Bouallouche-medjkoune, L., Alem, I., Yessad, S., Reinforcement, L. B., Based, L., & Alem, I. (2023). *Protocols for Industrial WSNs : A Critical Mini-Review . To cite this version : HAL Id : hal-04219667 Reinforcement Learning Based Communication Protocols for Industrial WSNs : A Critical*.

Alkhatib, A. A. A., & Baicher, G. S. (2012). *Wireless sensor network architecture*. *35*(Cncs), 11–15. https://pdfs.semanticscholar.org/ba9d/7436c266a8b309485fcaa126f661f2665dd3.pdf%0A http://www.ipcsit.com/vol35/003-CNCS2012-N010.pdf

Alsalmi, N., Navaie, K., & Rahmani, H. (2024a). Energy and throughput efficient mobile wireless sensor networks: A deep reinforcement learning approach. *IET Networks*, *March*. https://doi.org/10.1049/ntw2.12126

Alsalmi, N., Navaie, K., & Rahmani, H. (2024b). Energy and throughput efficient mobile wireless sensor networks: A deep reinforcement learning approach. *IET Networks*, *February*, 413–433. https://doi.org/10.1049/ntw2.12126

Aoudia, F. A., Gautier, M., & Berder, O. (2017). Learning to survive: Achieving energy neutrality in wireless sensor networks using reinforcement learning. *IEEE International Conference on Communications*. https://doi.org/10.1109/ICC.2017.7996978

Arafat, M. Y., Pan, S., & Bak, E. S. (2024). An Adaptive Reinforcement Learning-Based Mobility-Aware Routing for Heterogeneous Wireless Body Area Networks. *IEEE Sensors Journal*, *24*(19), 31201–31214. https://doi.org/10.1109/JSEN.2024.3440412

Batool, T., Ahmed, A., & Gaiti, D. (2024). Two level data centric aggregation scheme for wireless sensor networks. *Annales Des Telecommunications/Annals of Telecommunications*, *79*(5–6), 289–299. https://doi.org/10.1007/s12243-023-00984-5

Bhattacharyya, D., Kim, T. hoon, & Pal, S. (2010). A comparative study of wireless sensor networks and their routing protocols. *Sensors*, *10*(12), 10506–10523. https://doi.org/10.3390/s101210506

Bhimshetty, S., & Ikechukwu, A. V. (2024). Energy-efficient deep Q-network: reinforcement learning for efficient routing protocol in wireless internet of things. *Indonesian Journal of*

Electrical Engineering and Computer Science, *33*(2), 971–980. https://doi.org/10.11591/ijeecs.v33.i2.pp971-980

Buratti, C., Conti, A., Dardari, D., & Verdone, R. (2009). An overview on wireless sensor networks technology and evolution. *Sensors*, *9*(9), 6869–6896. https://doi.org/10.3390/s90906869

Chaurasia, S., & Kumar, K. (2024). MBASE: Meta-heuristic Based optimized location allocation algorithm for baSE station in IoT assist wireless sensor networks. *Multimedia Tools and Applications*, *83*(18), 53383–53415. https://doi.org/10.1007/s11042-023-17541-w

Chen, X., & Yu, P. (2010). Research on hierarchical mobile wireless sensor network architecture with mobile sensor nodes. *Proceedings - 2010 3rd International Conference on Biomedical Engineering and Informatics, BMEI 2010*, *7*(Bmei), 2863–2867. https://doi.org/10.1109/BMEI.2010.5639549

D, U. S. R., P, S., Arunkumar, K., R, S., & P, M. (2024). A HSEERP—Hierarchical secured energy efficient routing protocol for wireless sensor networks. *Peer-to-Peer Networking and Applications*, *17*(1), 163–175. https://doi.org/10.1007/s12083-023-01575-w

Daousis, S., Peladarinos, N., Cheimaras, V., Papageorgas, P., Piromalis, D. D., & Munteanu, R. A. (2024). Overview of Protocols and Standards for Wireless Sensor Networks in Critical Infrastructures. *Future Internet*, *16*(1), 1–27. https://doi.org/10.3390/fi16010033

Ding, Q., Zhu, R., Liu, H., & Ma, M. (2021). An overview of machine learning-based energy-efficient routing algorithms in wireless sensor networks. *Electronics (Switzerland)*, *10*(13). https://doi.org/10.3390/electronics10131539

Dvir, E., Shifrin, M., & Gurewitz, O. (2024). Cooperative Multi-Agent Reinforcement Learning for Data Gathering in Energy-Harvesting Wireless Sensor Networks. *Mathematics*, *12*(13). https://doi.org/10.3390/math12132102

Georgeon, O. L., Casado, R. C., & Matignon, L. A. (2015). Modeling Biological Agents beyond the Reinforcement-learning Paradigm. *Procedia Computer Science*, *71*, 17–22. https://doi.org/10.1016/j.procs.2015.12.179

Ghamry, W. K., & Shukry, S. (2024). Multi-objective intelligent clustering routing schema for internet of things enabled wireless sensor networks using deep reinforcement learning. *Cluster Computing*, *27*(4), 4941–4961. https://doi.org/10.1007/s10586-023-04218-0

Godfrey, D., Suh, B. K., Lim, B. H., Lee, K. C., & Kim, K. Il. (2023). An Energy-Efficient Routing Protocol with Reinforcement Learning in Software-Defined Wireless Sensor Networks. *Sensors (Basel, Switzerland)*, *23*(20). https://doi.org/10.3390/s23208435

Guo, Z., Chen, H., & Li, S. (2023). Deep Reinforcement Learning-Based UAV Path Planning for Energy-Efficient Multitier Cooperative Computing in Wireless Sensor Networks. *Journal of Sensors*, *2023*. https://doi.org/10.1155/2023/2804943

Hu, L., Han, C., Wang, X., Zhu, H., & Ouyang, J. (2024). Security Enhancement for Deep Reinforcement Learning-Based Strategy in Energy-Efficient Wireless Sensor Networks. *Sensors*, *24*(6), 1–14. https://doi.org/10.3390/s24061993

Institute of Electrical and Electronics Engineers. (2013). *2013 13th Mediterranean Microwave Symposium (MMS) : Saida, Lebanon, 2-5 September, 2013.*

Iqbal, M., Naeem, M., Anpalagan, A., Ahmed, A., & Azam, M. (2015). Wireless sensor network optimization: Multi-objective paradigm. *Sensors (Switzerland)*, *15*(7), 17572–17620. https://doi.org/10.3390/s150717572

Jain, A. C. (2024). *Strategy Relate to Congestion Control Protocol for Wireless Sensor Networks*. *1*(1), 26–31.

Jeyakarthic, M., & Selvakumar, T. (2024). Smart Cluster Head Selection Using Novel Duty Cycling and Sleep Schedule Optimization with Data-Aggregation for Wireless Sensor Network. *Proceedings of 2024 International Conference on Science, Technology, Engineering and Management, ICSTEM 2024*, *April*, 1–7. https://doi.org/10.1109/ICSTEM61137.2024.10561073

Journal, I. (2024). *An Introduction of Cluster Based Energy Efficient WSN Routing Protocols*. *December*.

Karunkuzhali, D., Meenakshi, B., & Lingam, K. (2024). A QoS-aware routing approach for Internet of Things-enabled wireless sensor networks in smart cities. *Multimedia Tools and*

Applications, *0123456789*. https://doi.org/10.1007/s11042-024-18125-y

Keerthika, A., & Berlin Hency, V. (2022). Reinforcement-Learning based energy efficient optimized routing protocol for WSN. *Peer-to-Peer Networking and Applications*, *15*(3), 1685–1704. https://doi.org/10.1007/s12083-022-01315-6

Lakshmi, M. S., Ramana, K. S., Ramu, G., Shyam Sunder Reddy, K., Sasikala, C., & Ramesh, G. (2024). Computational intelligence techniques for energy efficient routing protocols in wireless sensor networks: A critique. *Transactions on Emerging Telecommunications Technologies*, *35*(1), 1–27. https://doi.org/10.1002/ett.4888

López-Ardao, J. C., Rodríguez-Rubio, R. F., Suárez-González, A., Rodríguez-Pérez, M., & Sousa-Vieira, M. E. (2021). Current trends on green wireless sensor networks. *Sensors*, *21*(13). https://doi.org/10.3390/s21134281

Mann, P. S., & Singh, S. (2017). Energy-Efficient Hierarchical Routing for Wireless Sensor Networks: A Swarm Intelligence Approach. *Wireless Personal Communications*, *92*(2), 785–805. https://doi.org/10.1007/s11277-016-3577-1

Mubarak, Z., Zulkifl, M., Mustafa, M., Hussain, M. Z., Siddiqui, A., & Yaqub, M. A. (2024). *Application Layer Issues and Challenges in Supply Chain 4 . 0*. *2*(1), 1–14.

Nadeem, T., & Agrawala, A. (2005). Performance of IEEE 802.11 based wireless sensor networks in noisy environments. *Conference Proceedings of the IEEE International Performance, Computing, and Communications Conference*, *May 2005*, 471–476. https://doi.org/10.1109/pccc.2005.1460615

Nakas, C., Kandris, D., & Visvardis, G. (2020). Energy efficient routing in wireless sensor networks: A comprehensive survey. *Algorithms*, *13*(3), 1–65. https://doi.org/10.3390/a13030072

Nandyala, C. S., Kim, H., & Cho, H. (2022). *Pr ep rin ot pe er r Pr ep rin t n ed*.

Okine, A. A., Adam, N., Naeem, F., & Kaddoum, G. (2024). Multi-Agent Deep Reinforcement Learning for Packet Routing in Tactical Mobile Sensor Networks. *IEEE Transactions on Network and Service Management*, *21*(2), 2155–2169. https://doi.org/10.1109/TNSM.2024.3352014

Pantazis, N. A., Nikolidakis, S. A., & Vergados, D. D. (2013). Energy-efficient routing protocols in wireless sensor networks: A survey. *IEEE Communications Surveys and Tutorials*, *15*(2), 551–591. https://doi.org/10.1109/SURV.2012.062612.00084

Prabhu, D., Alageswaran, R., & Miruna Joe Amali, S. (2023). Multiple agent based reinforcement learning for energy efficient routing in WSN. *Wireless Networks*, *29*(4), 1787–1797. https://doi.org/10.1007/s11276-022-03198-0

Priyadarshi, R. (2024a). Energy-Efficient Routing in Wireless Sensor Networks: A Meta-heuristic and Artificial Intelligence-based Approach: A Comprehensive Review. *Archives of Computational Methods in Engineering*, *31*(4), 2109–2137. https://doi.org/10.1007/s11831-023-10039-6

Priyadarshi, R. (2024b). Exploring machine learning solutions for overcoming challenges in IoT-based wireless sensor network routing: a comprehensive review. *Wireless Networks*, *30*(4), 2647–2673. https://doi.org/10.1007/s11276-024-03697-2

Qiu, Y., Ma, L., & Priyadarshi, R. (2024). Deep Learning Challenges and Prospects in Wireless Sensor Network Deployment. *Archives of Computational Methods in Engineering*, *31*(6), 3231–3254. https://doi.org/10.1007/s11831-024-10079-6

Rajaravivarma, V., Yang, Y., & Yang, T. (2003). An overview of Wireless Sensor Network and applications. *Proceedings of the Annual Southeastern Symposium on System Theory*, *2003-Janua*, 432–436. https://doi.org/10.1109/SSST.2003.1194607

Rodoshi, R. T., Song, Y., & Choi, W. (2021). Reinforcement Learning-Based Routing Protocol for Underwater Wireless Sensor Networks: A Comparative Survey. *IEEE Access*, *9*, 154578–154599. https://doi.org/10.1109/ACCESS.2021.3128516

Saadallah, N. R., & Alabady, S. A. (2024). A comprehensive study on energy-efficient-based routing protocols in the Internet of things Part I: definition and classification. *Iran Journal of Computer Science*, *7*(3), 631–661. https://doi.org/10.1007/s42044-024-00180-1

Safiyan, A., Safiyan Ibrahim, A., Abubakar, M., & Iliyasu, U. (2022). Optimization of Adaptive Method for Data Reduction in Wireless Sensor Networks. *American Academic Scientific*

Research Journal for Engineering, *90*(1), 27–40.

Sardar, M., & Majumder, K. (2013). *A Survey on Trust Based Secure Routing in Manet*. *May*, 243–253. https://doi.org/10.5121/csit.2013.3525

Selvakumar, T., & Jeyakarthic, M. (2024). Adaptive Receiver-Centric Multicast Routing Protocol Using Dynamic Deep Q-Networks. *Nanotechnology Perceptions*, *20*(S9), 438–453. https://doi.org/10.62441/nano-ntp.v20is9.33

Shah, B. R., & Yashwanth, N. (2022). Energy-Efficient Routing Protocols for Wireless Sensor Networks. *4th International Conference on Emerging Research in Electronics, Computer Science and Technology, ICERECT 2022*. https://doi.org/10.1109/ICERECT56837.2022.10060003

Sharma, T., Balyan, A., Nair, R., Jain, P., Arora, S., & Ahmadi, F. (2022). ReLeC: A Reinforcement Learning-Based Clustering-Enhanced Protocol for Efficient Energy Optimization in Wireless Sensor Networks. *Wireless Communications and Mobile Computing*, *2022*. https://doi.org/10.1155/2022/3337831

Shi, F., Tuo, X., Ran, L., Ren, Z., & Yang, S. X. (2020). Fast Convergence Time Synchronization in Wireless Sensor Networks Based on Average Consensus. *IEEE Transactions on Industrial Informatics*, *16*(2), 1120–1129. https://doi.org/10.1109/TII.2019.2936518

Singh, V., & Rastogi, A. (2019). Various energy efficient routing protocols for wireless sensor networks. *International Journal of Engineering and Advanced Technology*, *8*(5), 2619–2623.

Singh, Y., & Walingo, T. (2024). Smart Water Quality Monitoring with IoT Wireless Sensor Networks. *Sensors*, *24*(9). https://doi.org/10.3390/s24092871

Sivadasan, S., & Nagarajan, G. (2025). *Optimized Energy-Efficient Routing and Adaptive Coverage in Wireless Sensor Networks for Multi-Access Edge Environments*. *32*(2), 435–446.

Somula, R., Cho, Y., & Mohanta, B. K. (2024). SWARAM: Osprey Optimization Algorithm-Based Energy-Efficient Cluster Head Selection for Wireless Sensor Network-Based

Internet of Things. *Sensors*, *24*(2). https://doi.org/10.3390/s24020521

Suresh, S. S., Prabhu, V., Parthasarathy, V., Senthilkumar, G., & Gundu, V. (2024a). Intelligent data routing strategy based on federated deep reinforcement learning for IOT-enabled wireless sensor networks. *Measurement: Sensors*, *31*(November 2023), 101012. https://doi.org/10.1016/j.measen.2023.101012

Suresh, S. S., Prabhu, V., Parthasarathy, V., Senthilkumar, G., & Gundu, V. (2024b). Intelligent data routing strategy based on federated deep reinforcement learning for IOT-enabled wireless sensor networks. *Measurement: Sensors*, *31*(January), 101012. https://doi.org/10.1016/j.measen.2023.101012

Suresh, S. S., Prabhu, V., Parthasarathy, V., Senthilkumar, G., & Gundu, V. (2024c). Intelligent data routing strategy based on federated deep reinforcement learning for IOT-enabled wireless sensor networks. *Measurement: Sensors*, *31*(December 2023). https://doi.org/10.1016/j.measen.2023.101012

Thein, M. C. M., & Thein, T. (2009). A framework for secure and survivable wireless sensor networks. *Annals of Dunărea de Jos University. Fascicle I : Economics and Applied Informatics*, *1*(1584–0409), 243–252.

Ullah Khan, S., Ulalh Khan, Z., Alkhowaiter, M., Khan, J., & Ullah, S. (2024). Energy-efficient routing protocols for UWSNs: A comprehensive review of taxonomy, challenges, opportunities, future research directions, and machine learning perspectives. *Journal of King Saud University - Computer and Information Sciences*, *36*(7), 102128. https://doi.org/10.1016/j.jksuci.2024.102128

Üveges, B. Á., & Oláh, A. (2024). Enhanced reliability in hazardous event detection: A resilient multipath routing protocol for wireless sensor networks. *IET Wireless Sensor Systems*, *14*(4), 111–131. https://doi.org/10.1049/wss2.12078

Vassilaras, S. (2012). An Overview of Wireless Sensor Networks , their Applications and Technical Challenges. *International Journal of Soft Computing and Engineering (IJSCE), ISSN: 2231-2307*, *2*, 538–540.

Veena, R. S., Khambra, G., Sayal, R., Singh, N., Dhandayuthapani, V. B., & Sar, S. K. (2024).

An Energy Efficient IPv6 Packet Delivery Outline for Industrial IoT Over G.9959 Protocol Based Wireless Sensor Network (WSN). *SN Computer Science*, *5*(7). https://doi.org/10.1007/s42979-024-03195-4

Venkateswara Rao, M., & Malladi, S. (2021). Improving Packet Delivery Ratio in Wireless Sensor Network with Multi Factor Strategies. *International Journal of Advanced Computer Science and Applications*, *12*(5), 627–634. https://doi.org/10.14569/IJACSA.2021.0120575

Wakili, A., Bakkali, S., & Alaoui, A. E. H. (2024). Machine learning for QoS and security enhancement of RPL in IoT-Enabled wireless sensors. *Sensors International*, *5*(July). https://doi.org/10.1016/j.sintl.2024.100289

Wang, D., Liu, J., & Yao, D. (2020). An energy-efficient distributed adaptive cooperative routing based on reinforcement learning in wireless multimedia sensor networks. *Computer Networks*, *178*(October 2019). https://doi.org/10.1016/j.comnet.2020.107313

Yadawad, S., & Joshi, S. M. (2024). Efficient energy consumption and fault tolerant method for clustering and reliable routing in wireless sensor network. *Peer-to-Peer Networking and Applications*, *17*(3), 1552–1568. https://doi.org/10.1007/s12083-024-01664-4

Yu, X., Liu, Y., & Liu, Y. (2024). WSN routing algorithm based on node classification and multi-layer transport. *Wireless Networks*, *30*(2), 737–747. https://doi.org/10.1007/s11276-023-03497-0

Yun, W. K., & Yoo, S. J. (2021). Q-Learning-based data-aggregation-aware energy-efficient routing protocol for wireless sensor networks. *IEEE Access*, *9*, 10737–10750. https://doi.org/10.1109/ACCESS.2021.3051360

Zibakalam, V., & HosseinKahaei, M. (2012). Increasing Throughput and Reducing Delay in Wireless Sensor Networks Using Interference Alignment. *International Journal of Communications, Network and System Sciences*, *05*(02), 90–97. https://doi.org/10.4236/ijcns.2012.52012