Chapter 1

Introduction

Internet of Things (IoT) has become an expeditious phenomenon integrating physical systems with digital technologies. It establishes a network ecosystem that can gather, process, and act upon data at the point of generation. Thus, there is a need for effective and intelligent edge device integration in the IoT environment. IoT deployment has grown in various fields such as industrial automation, smart cities, healthcare systems *etc.* Examples of edge devices include smartphones, IoT devices, edge servers, and network infrastructure components. Data processing closer to the source or edge computing has become more popular, resolving the bandwidth limitation, latency and data privacy raised in cloud-centric computing systems. The edge devices in long-range IoT applications may require data transmission over long distances in remote locations, which is challenging. The capabilities and efficiency of conventional edge devices may not be sufficient to satisfy the demands of such scenarios. Hence, we need a Machine Learning (ML) solution to improve edge device capability and efficiency by analysing data locally. It also improves responsiveness and quality of service. To create a large network of entities that share information, the IoT primarily requires the installation of networked devices, sensors, edge devices and other items that collect, send, and process data.

We also define edge devices by employing key parameters, including processing power, memory (RAM), power consumption, and connectivity. The edge devices interface for connecting sensors and actuators and inference latency for real-time applications. Furthermore, cost, scalability, AI/ML capabilities, battery life, and edge analytics capacity further influence their suitability for specific applications.

IoT applications rely on low-power devices that communicate and interact together and transfer data from the environment to the Network Server (NS). Long Range Wireless Area Networks (LoRaWAN) is a new emerging technology for long-range and low-

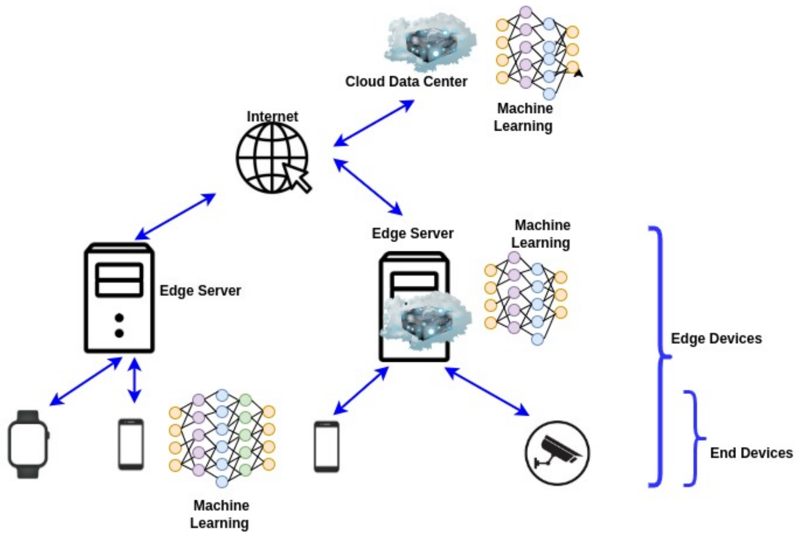

Figure 1.1: Edge devices framework with edge server and cloud layers employing machine learning.

power communication paradigms. The popularity of LoRaWAN in IoT applications is a strength of its low power consumption, cost-effectiveness, and high range. There are many Low Power Wireless Area Networks (LPWAN) communication protocols such as NBIoT, Sigfox, LoRaWAN, *etc.* Among these protocols, LoRaWAN is the most popular because of its features. Sensor devices are connected with a small LoRa chipset to enable various IoT applications by transmitting LoRa packets. The Chirp Spread Spectrum (CSS) modulation technique is used in LoRa to send data from the LoRa Node (LN) to the LoRa Gateway (LG). LG sends data to NS, which manages the entire network and responds to allocate resources (like Spreading factor (SF), Bandwidth (BW), *etc.*) to LN. NS sends data to the application server to get information for further interpretation. LoRaWAN technologies find their applications in various domains, including smart homes, smart agriculture, smart metering, smart parking, pollution monitoring, fire detection, industries, health care, *etc.*

With the exponential growth in the LoRa Network, the researchers explore and analyze issues like scalability, collision, capture effect, and LoRa parameter allocation. A potential tool for solving problems with data processing is ML. The IoT framework has improved the adaptability, intelligence, versatility, and decision-making capabilities of ML. One of the main advantages of ML in IoT is its capacity to get valuable insights from the enormous data streams produced by networked devices. Combining IoT and ML may greatly increase the effectiveness, dependability, and functionality of IoT applications, creating a more connected and intelligent society. Additionally, IoT ML models' versatility is a game-changer. IoT systems will continue to be sensitive to shifting conditions and requirements thanks to these algorithms' ability to learn and change along with the changing data landscape.

The convergence of ML IoT models is moving us toward a future in which our urban and personal settings are not just more connected but also more intelligent. It is a future in which data is actively used to improve our lives, fuel innovation, and address complicated challenges rather than only being collected. An era of IoT systems that not only perceive but also grasp the world is ushered in by this synergy, making the world for all of us more responsive, intelligent, and effective. Nowadays, IoT application uses new machine learning techniques like federated learning and lightweight deep neural networks because it implemented in resource-constrained devices. In the time of big data, data science, and machine learning, there is demand for robust and data privacy-preserving techniques has become more important. In IoT applications, we collected data from different sources to analyze, process, and provide patterns or results. The process of transferring huge amounts of data from different servers raises issues of security and bandwidth wastage. Tradition ML technique uses the concept of collecting and analyzing at the centralized server. Federated Learning (FL) is the part of machine learning is comes with promising solutions to these issues by enabling a decentralized device training model. After training the decentralized model, each device sends only the training parameter instead of the complete data set.

FL is a system that allows several parties to train models cooperatively without sharing raw data, making it appealing to both industry and research. Despite its high promise, FL faces several challenges, such as model synchronization, communication efficiency, security concerns, and less availability of label data points. Furthermore, it needed algorithm and systems architecture advancements to utilize its full capacity. IoT applications are used in a variety of fields, including smart buildings, industrial automation, healthcare, smart cities, and Intelligent Transportation Systems (ITS). The IoT application, however, confronts numerous hurdles in the areas of security, scalability, energy efficiency, connectivity, and data processing as it develops an ITS that is revolutionizing transportation by seamlessly integrating sensing, analysis, control, and communication technologies to increase mobility. However, it also faces challenges with data privacy and datasets with inaccurate labels, which can reduce the efficiency of machine learning models. The adoption of LoRaWAN, which provides long-range and energy-efficient connectivity, solves ITS communication issues, sensors, acoustics, and networked devices are used in smart buildings to automatically gather and analyze data for increased energy efficiency, cost savings, and occupant comfort. Due to sensor installation and maintenance challenges, particularly in spaces with high ceilings, energy-efficient and long-range wireless communication solutions are necessary. LoRaWAN emerges as a practical option thanks to its star-of-stars network architecture and low-power, long-range communication capabilities.

Augmented IoT (AIoT), improves decision-making processes by combining human and machine intelligence. In cases involving data collecting, mathematical calculations, medical crises, and environmental monitoring, AIoT is critical to UAV-enabled aerial networks. However, effective solutions for UAVs with low resources are required. AIoT is a unique method that focuses on shrinking the size of Deep Neural Networks (DNNs) to enable successful training and deployment based on variables such as remaining battery life, available memory, and task frequency. These paradigms—ITS, Smart Buildings, and AIoT—emphasize technology's revolutionary capacity by providing creative solutions and addressing the specific difficulties that arise in each industry. Some challenges with existing techniques are mentioned as follows: Long distances, erratic wireless connectivity, and restricted bandwidth can all make it difficult for vehicles (participants) to send huge volumes of transportation data to the central server. This is addressed by long-range communication technologies such as LoRaWAN, however they still have interference issues. The dynamic nature of vehicular environments faces challenges to manage stable and reliable communication between moving vehicles and the central system. As vehicles move, network connectivity may become unstable, leading to data loss or delays. In ITS, the data collected from different vehicles and infrastructure sensors may be highly heterogeneous (Non-IID), as each vehicle may encounter different traffic conditions, routes, or environments. ITS Systems require real-time data processing to ensure prompt action, such as managing incidents, modifying traffic signals, or providing route suggestions. In the FL system, sharing data among devices and participants may concern of privacy and security problems, particularly when it comes to sensitive transportation data like speed, location, and travel habits.

1.1 Research objectives

The primary objective of this research is to investigate and develop the ITS framework, which improves vehicle health, driver safety, and data safety, ensures efficient data communication and improves training time. This framework aims to address the following key research objectives:

- Vehicle Health Improvement: ITS framework continuously improves monitoring and assessing vehicle health. To develop a model to predict the need for mainte-

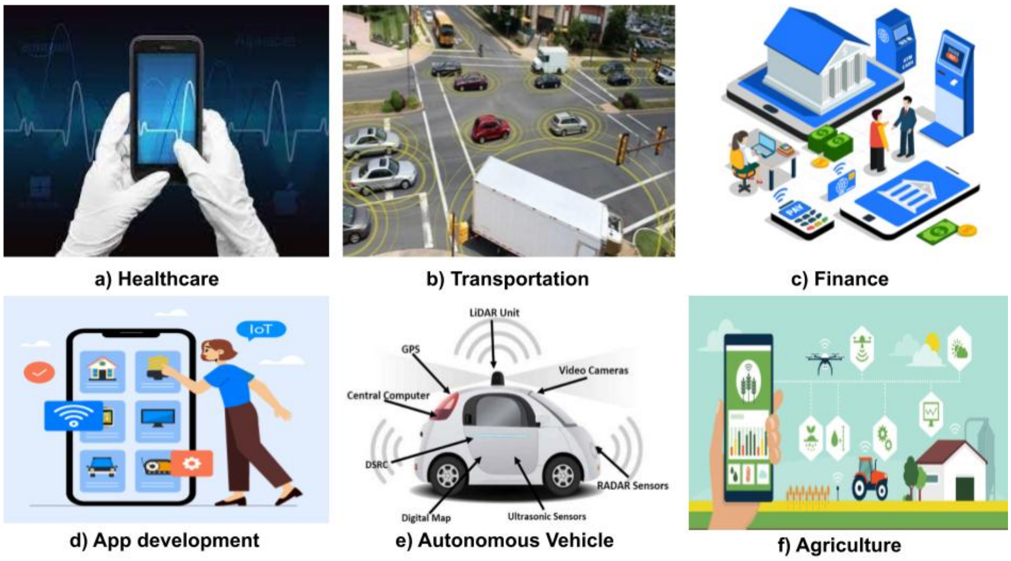

Figure 1.2: An application of edge devices.

nance and prevent potential failure.

- Driver Safety improvement: ITS framework continuously improves driver behaviour monitoring and environmental condition monitoring. To design a system to generate alert drivers about the wrong driving and suggest the corrective major.
- Data Privacy and Security in ITS: Address privacy and security concerns associated with sharing ITS data remotely. To secure these data by using federated learning techniques.
- Networking Resource Optimization: The network resources requirement sharing parameter while training ITS data, which focused on high data rate, low latency, low cost and low packet loss. LoRaWAN network is best suited for ITS framework for transmitting the parameter while training the model.
- Handling Imperfect Labels: To develop an algorithm to estimate the class-wise centroid of the dataset at both the client and the server. To design the mechanism to identify the client with imperfect labels using similarity score.
- Performance Improvement: Evaluate the performance improvements achieved through the proposed framework, including enhanced communication efficiency, and improved accuracy after reducing imperfect labelling at clients in federated learning.

1.2 Motivation of the research work

ITS are emerging areas of recent research. Literature indicates several important contributions toward FL and LoRaWAN in intelligent applications such as transportation systems, smart buildings, and UAVs. The critical issues facing the field of ITS and the urgent demand for a workable solution serve as the driving forces behind this research. Modern transportation has seen the emergence of ITS as a critical paradigm, offering safer and more effective movement for both products and passengers. However, several serious problems have hampered ITS's potential, necessitating creative fixes. The most important of these concerns is the privacy and security of personal data. The substantial volume of data gathered and processed by ITS includes sensitive data relating to cars, drivers, and passengers. How can we preserve this data while maintaining the continuing development of ITS technology given the conventional method of remotely sharing this data for training machine and deep learning models? Information integrity must be protected because data breaches and security risks are so prevalent.

Furthermore, a significant problem within the FL technique is the existence of imperfect labels in participant datasets. Label imperfections, which frequently result from subpar annotation procedures or other causes, have the potential to seriously lower the calibre of models. Due to the importance of model accuracy and utility, the success of FL in the context of ITS is closely related to how this problem is solved.

The effectiveness of communication in ITS, a crucial element for real-time operations, is a worry in addition to these problems. Vehicle mobility, great separations between them and the processing units, and the erratic nature of wireless connectivity all add complications that must be resolved. We aim to maximise communication efficiency to ensure that data transmission continues to be both dependable and energy-efficient. We are aware of the potential of LoRaWAN technology as a solution LoRaWAN, there are many problems and solutions given in existing work. We are facing interface problems in the LoRaWAN network in real-time applications such as ITS, smart building, industry automation, traffic management, etc. These applications motivate my research toward LoRaWAN interface mitigation systems in intelligent real-time systems. The relevance of ITS in the real world, where efficiency, security, and data management are crucial, is the driving force behind this research. The accomplishment of the below-mentioned issues will not only improve the capabilities of ITS but also advance the larger objective of improving transportation security, mobility, and safety.

Finally, the growing need to handle data privacy and security issues, reduce the effects of inaccurate labels, and maximize communication efficiency, advancing the field and improving transportation systems security and effectiveness. We also developed a method for an interference-free efficient communication system.

1.3 Contributions and organization of the thesis

In this thesis, we investigated the different challenges encounter while enhancing the edge devices applicability for IoT applications. We first considered the problem of imperfect labels or poor annotation in the federated learning environment. We solved the problem using centroid estimation of participant vehicle data and compared the similarity score of the server centroid and participant centroid. We use the LoRa communication protocol that provides efficient, low cost and long-range data transmission. We next solved the interference problem in the LoRa network using SNR-based analysis techniques to solve the issue. Finally, we develop an approach to enhance the lifetime of UAV-enabled aerial networks via Augmented IoT (AIoT). The main objective of the AIoT approach is to improve the lifetime of aerial networks by replacing DNN on battery-operated UAVs with lighter versions, incurring minimal accuracy compromise. We utilized the concept of knowledge distillation, where we adopted layer sharing among teachers and students, followed by selective back-propagation of shared layers during student training.

- The first contribution is to transferred the information from vehicles to the processing unit in a communication-efficient manner. Each vehicle has an LN to collect and transfer the information to the LG using LoRa communication. We propose Fed-LoRa approach to perform secure and reliable data transfer. Initially it estimates the class-wise centroids of the local datasets at participants and servers. Fed-LoRa approach generated feature vectors from the raw data on the participants and server. Afterwards, the similarity score is estimated between the received and server's centroids. The participants with any of the class similarity scores less than a threshold are inferior. The inferior participants demand a fraction of the server dataset to improve their similarity score and, thus, performance. We proposed novel data reduction and inclusion mechanisms that ensure that the number of instances in the local dataset does not exceed the existing ones. We further proposed an optimization problem and derive an expression for the optimal fraction of the server dataset transferred to the inferior participants.
- In the next contribution, we introduced a novel approach that utilizes a classification model to identify interfering LNs in high-ceiling smart buildings. Our approached involves collecting and analyzing network parameters such as signal-

to-noise ratio (SNR) and received signal strength indicator (RSSI) from the signals to extract features for the classifier. The approach classified interference based on the number of interfering LNs, where each class represents a different number of interfering LNs. Additionally, we proposed a push-based mechanism for detecting and adjusting the power levels of faulty LNs to mitigate interference. Our approached is hardware-independent, making it a cost-effective solution that can be implemented on the LG platform. We contribute to the field in two ways: (1) a novel classification approach for identifying interfering LNs, and (2) a push-based mechanism for identifying faulty LNs and adjusting their power levels.

- Further, in the third contribution, we transformed a large DNN into a lightweight, which achieves high accuracy using available residual energy α and memory β on task frequency γ. The operator decides γ using AIoT. We refer to this transformation as (α, β, γ)-DNN transformation problem. We utilized the knowledge distillation technique to simultaneously train the student (lightweight DNN) and teacher (large-size DNN). We introduced the concept of clever training, which incorporates layer sharing among teachers and students with selective back-propagation. The shared layers of the student undergo training via selective back-propagation. We present an algorithm that uses the proposed solution of (α, β, γ)-DNN transformation problem and quick training of obtained lightweight DNN. The algorithm plays a vital role in scenarios where the resources of UAVs change dynamically and demand different versions of DNN.
- Our final contribution is the creation of an open dataset of LoRa interference. This labeled dataset includes various deployment scenarios with multiple LNs, both with and without interference. For each message sent by the LN to the LG, the dataset includes important features such as SNR, RSSI, payload, and time stamps. The free availability of this dataset will enable researchers to compare the performance of different interference mitigation techniques and encourage the development of new approaches for LoRa interference.

The rest of the thesis is organized as follows:

Chapter 2: This chapter presents the preliminary techniques used in this thesis, including federated learning techniques, LoRaWAN, knowledge distillation, and federated learning. The chapter also includes recent research or prior studies on A Federated Learning Approach with Imperfect Labels in LoRa-based Transportation Systems, Machine Learning-Based Interference Mitigation in Long-Range Networks for High-Ceiling Smart Buildings, and Enhance Lifetime of UAV-Enabled Aerial Networks.

Chapter 3: This chapter covers Fed-LoRa, a federated learning approach for Intelligent

Transportation Systems (ITS). This work aims to improve data security and communication efficiency while handling the issue of imperfect label data in participant device datasets.

Chapter 4: This chapter covers the second problem of identifying the interference nodes in LoRa networks. Our approach utilizes machine learning techniques, using SNR and RSSI values to identify nodes that cause network interference accurately. We conducted experiments in different deployment scenarios, including a high-ceiling smart building, and observed that network parameter-based interference solutions work well in scenarios with few obstacles and static LNs and LG.

Chapter 5: This chapter proposes an approach to enhance the lifetime of UAV-enabled aerial networks via AIoT. The main objective of the EU-AIoT approach is to improve the lifetime of aerial networks by replacing DNN on battery-operated UAVs with lighter versions, incurring minimal accuracy compromise. We further utilized the concept of knowledge distillation, where we adopted layer sharing among teachers and students, followed by selective back-propagation of shared layers during student training. We finally carried out various experiments to validate the effectiveness of the EU-AIoT approach.

Chapter 6: In this chapter, we summarize the key findings of the research work emphasized in this thesis. We also point out the promising futuristic research direction in the domain of LoRa network sensing and FL for intelligent applications.

Technical posters, research papers, and books are listed in References. The publications resulting from the research work in this thesis are included in the List of Publications.